



Learning Automata Based Long Range Wide Area Network Channel Selection

Submitted By

ATADET Luka Aime

Registration Number: 222019234

College of Science and Technology

School of Information and Communication Technology (SoICT)

Master of Science in Information and Communication Technology (Option: Operational
Communication)

July, 2024



UNIVERSITY of
RWANDA

***Research and Postgraduate Studies
(RPGS) Unit***

Learning Automata-Based Long Range Wide Area Network Channel Selection By

Name: ATADET Luka Aime

Registration Number: 222019234

A dissertation submitted in partial fulfillment of the requirement for the degree of
MASTER of ICT (Option: Operational Communication)

In the College of Science and Technology

Supervisors: Assoc. Prof. Richard MUSABE

Dr. Omar GATERA

Dr. Eric HITIMANA

July 2024

DECLARATION

I declare that this project entitled: «Learning Automata-Based Long-Range Wide Area Network Channel Selection» is original and has never been submitted to any University or other Institution of Higher Learning. It is my research whereby other scholar's writing was cited and referenced. I thus declare this work is mine and was completed successfully under Associate Professor Richard MUSABE, Dr. Omar GATERA, and Dr. Eric HITIMANA.

ATADET Luka Aimé

Reg: 222019234

Signature:

Date:

CERTIFICATE

This is to certify that the project work entitled « Learning Automata-Based Long-range Wide Area Network Channel Selection » is the record of my original work done by ATADET Luka Aimé with Reg No: 222019234 in partial fulfillment of the requirement for the award of Master of Information and Communication Technology, option Operational Communication of College of Science and Technology, University of Rwanda during the academic year 2024.

Supervisors

Assoc. Prof. Richard MUSABE

Signature:

Date:

Dr. Omar GATERA

Signature:

Date:

Dr. Eric HITIMANA

Signature:

Date:

ACKNOWLEDGEMENT

First and foremost, I would like to thank God for his countless blessings much more than I deserve. I am deeply grateful to everyone who has supported and guided me throughout the process of completing this master's thesis.

I would like to express my sincere gratitude to my supervisors: Assoc. Prof Richard MUSABE, Dr. Omar GATERA, and Dr. Eric HITIMANA for their continuous support, they have been invaluable in this thesis. Their outstanding support and contributions have made this work possible.

I am also thankful to the faculty and staff of the School of Information Communication Technology faculty and staff at the University of Rwanda for providing a stimulating academic environment and for their assistance throughout my studies. Special thanks to Dr. Frederick NZANYWAYINGOMA for his advice and support.

I extend my appreciation to my family and friends for their unwavering support and encouragement. To my parents, Ing. ATADET Azarak Mogro and ATADET Djedida Baifouta, thank you for believing in me and providing the foundation for my academic pursuit.

Lastly, I am grateful to all those who contributed directly or indirectly to the successful completion of this thesis. Your support and encouragement have been invaluable.

Thank you all.

ATADET LUKA AIME

ABSTRACT

The advancement in technology has led to the requirement of wireless communication at longer distances in areas such as wide range environmental monitoring, automated meter reading and wide area sensors networks. The principle of Long-Range WAN is to provide reliable communication shifting the focus of wireless systems from static spectrum to dynamic spectrum access. Variation of channel state is the key to designing the channel selection mechanism in LoRaWAN. The mechanism must be predictive and must consider the multi-step decision-making based on the reward and penalty response for channel selection. High efficiency and reliability in decision making concerning various channel states can be determined by sequential decision processes. For LoRaWAN Learning Automata is the ideal process to integrate the channel selection mechanism.

Learning Automata has been used as AI powerful tool in different area to solve different problems including the one for channel selection. Despite its variant version Learning automata improved its power invoking the concept of Pursuit. This research aims to integrate a Learning Automata-based channel selection framework for LongRange Wide Area Network (LoRaWAN) deployments. The framework leverages Hierarchical Discrete Pursuit Learning Automata (HDPA) to select optimal channels among multiple channels, aiming to enhance network performance and mitigate interference. The study investigates LoRaWAN communication characteristics, integrates the HDPA model, evaluates its performance through simulations, and compares it with existing methods. The research methodology involves theoretical modeling, simulation, to evaluate the proposed Hierarchical Discrete Pursuit Learning Automata-based channel selection algorithm's effectiveness. The outcomes are expected to advance the efficiency and reliability of LoRaWAN networks, benefiting applications across smart cities, IoT, and industrial sectors. The simulation shows that our HDPA outperformed the HCPA in terms of speed and accuracy.

Keywords: Channel Selection, Pursuit, Learning automata, Estimator based – LA, LoRaWAN, HCPA, HDPA.

Table of Contents

DECLARATION	ii
CERTIFICATE	iii
ACKNOWLEDGEMENT	iv
ABSTRACT	v
CHAPTER I: GENERAL INTRODUCTION	1
1.0 Introduction.....	1
1.1 Background of the study	1
1.1.1 Motivation.....	3
1.1.2 Problem statement	3
1.2 General Objective	4
1.2.1 Specific Objectives	4
1.3 Scope and Limitation of the Research	4
1.4 Research Questions.....	5
1.5 Hypothesis	5
1.6 Significance of the study	5
1.7 Organization of the Study	5
1.8 Conclusion of Chapter One	6
CHAPTER 2: LITERATURE REVIEW	7
2.0 Introduction.....	7
2.1 LoRa and Long-Range Wide Area Network (LoRaWAN)	7
2.2 Machine Learning	9
2.3 Learning Automata	10
2.3.1 The pursuit.....	11
2.3.2 Linear VSSA Schemes	11
2.4 Conclusion	26
CHAPTER 3: RESEARCH METHODOLOGY	27
3.0 Introduction.....	27
3.1 System model.....	27
3.2 Proposed algorithm flowchart.....	28

3.3	Development Approaches.....	29
3.4	Mathematical Development.....	29
3.5	Software Environment	31
3.6	System Simulation	31
3.7	Simulation Variables	32
3.8	Conclusion	32
CHAPTER 4: RESULTS AND DISCUSSION		33
4.0	Introduction.....	33
CHAPTER 5: CONCLUSION, RECOMMENDATION AND FUTURE WORK		39
5.1	Conclusions.....	39
5.2	Recommendation	39
5.3	Future Work.....	39
References.....		41

ACRONYMS

AI: Artificial Intelligence.

BW: Bandwidth.

FSSA: Fixed structure Stochastic Automata.

IoT: Internet of Things.

LA: Learning Automata.

LoRaWAN: Long Range Wide Area Network.

MAB: Multi-Armed-Bandit.

MAC: Medium Access Control.

RL: Reinforcement Learning.

TM: Tsetlin Machine.

VSSA: Variable Structure Stochastic Automata.

LIP: Linear Inaction-Penalty.

LRI: Linear Reward-Inaction.

LRP: Linear Reward -Penalty.

LIST OF FIGURES

Figure 1: LoRaWAN Network Architecture.....	8
Figure 2: LoRaWAN Network Architecture.....	9
Figure 3: Learning- Automata and Environment model.	11
Figure 4: Linear Reward- Penalty scheme.	12
Figure 5: Proposed LoRa Network Model	25
Figure 6: Simulation flow chart	26
Figure 7: Research Development Approach	27
Figure 8: Reward probabilities for 8 channels	33
Figure 9: Throughput	34
Figure 11: Channel selection updating probability for successful transmission.	35
Figure 12: Number of iterations for convergence for 200 experiments.	36

LIST OF TABLES

Table 1: List of successful data transmission probability for the 8 channels.	32
Table 2: Result of our simulations for 8 channels	33
Table 3: Comparison between HDPA and HCPA	36

LIST OF NOTATIONS

\mathcal{A} : Learning Automata

α : Action

β : Response from the environment

Δ : Learning parameter T :

Convergence criterion Threshold d :

Estimate reward probability.

CHAPTER I: GENERAL INTRODUCTION

This Section intends to deal with the background of the study which will detail the statement of the problem, scope and Purpose of the study, objectives of the study, research questions, significance of the study, and outline the organization of the work.

1.0 Introduction

The rapid evolution of Wireless Communication Technology has necessitated the development of efficient and reliable data transmission methods [1]. Long Range Wide Area Networks have emerged as a prominent solution for IoT applications due to their low power consumption and Long-Range capabilities. As the number of IoT devices increases there is a demand for better Throughput, performance, and reliable communication [2]. Finding the best optimal path for data transmission is essential to avoid collision and improve the performance in LoRaWAN. However, the task of choosing a channel has gotten more challenging with the introduction of a Heterogeneity network in LoRaWAN where data rate changes for each channel [3].

The utilization of a technique for frequency allocation referred to as Pseudo-random channel hopping is intrinsically integrated into the LoRaWAN protocol to uniformly spread data transmissions across the range of available channels, thereby reducing the probability of signal interference occurrences [4]. However, this approach is constrained in its capacity to dynamically adjust to the fluctuating levels of noise that may be detected on communication channels. To address this issue we propose a Channel selection method base on Learning Automata which converge to the optimal channel that have the maximum probability of successful transmission.

1.1 Background of the study

In the paradigm of Machine Learning, we can find the field of Reinforcement Learning which has a very interesting subset known as learning automata (LA). In this subset, an agent learns intending to solve a task through the computer program. The discipline of Learning Automata (LA) is focused on the creation of intricate adaptive learning systems that seek resolutions for problems characterized by stochastic properties

through repetitive interactions with an entity known as the Environment. The term automaton was used for the first time by Tsetlin in the 1960s. After several years this term was used to describe deterministic and stochastic schemes to improve the performance of discrete systems which was adopted in the field of Artificial Intelligence [5].

The roots of LA can be traced back to a combination of behavioral studies, statistical analysis of decision-making based on past information, the resolution of the multiarmed bandit problem, and system theories that guide rational decision-making in stochastic environments [6].

On the other hand, Long Range Wide Area Network (LoRaWAN) refers to the communication protocol and the system architecture for the network that must be built on the already defined LoRa technology for the physical layer. A medium access control (MAC) layer protocol makes communication between IoT devices and the network server through secure, bidirectional communication. LoRaWAN architecture may be characterized by star-of-stars topology since each gateway forwards messages to and from end devices to a central network server and defends three classes of end devices [7].

Several aspects enable the LoRaWAN signals to mitigate some interference which uses CSS modulation and allows the receiver to separate the overlapping signals by having the property to capture an effect that is useful in the context of a crowded network. Additionally, LoRaWAN uses ADR algorithms that optimize the data rate and RF power according to network conditions. It then reduces the possibilities of risk or interference within the channel [8]. Although LoRaWAN is highly substantial in easing interference, as the number of IoT devices grows exponentially, LoRaWAN leads to a reduced performance that requires a continuous effort to be taken to improve its interference mitigation.

The recent developments in the LoRaWAN protocol provide more complex approaches to reduce interference such as Spread Spectrum Modulation, Adaptive Data Rate (ADR), Listen Before Talk (LBT), Channel Hopping, Adaptive Frequency Agility (AFA), Gateway Placement and Network Planning and Clear Channel Assessment (CCA). This will lead to a living together of a greater number of networks with a larger

duty cycle if possible while minimizing the interference that would be afflicted on each of these networks.

LoRaWAN does the interference reduction by spreading the transmissions around different frequencies and one channel does not become too busy in frequency hopping. Conversely, duty-cycle management regulates the time when the device is transmitting in the given period to reduce the probability of simultaneous transmissions that can bring interference. The proposed solution in this thesis is to select a single channel that has a high channel selection probability which is the best among a set of multiple channels and that channel will statistically give the maximum reward for successfully transmitting a message. At the starting point, all the channels are the same and the best one should be selected after a random process by the Learning Automata which will learn and converge to the best channel that can produce a maximum throughput.

1.1.1 Motivation

The motivation behind this research is to address challenges in dynamic and congested radio environment where factors like interference poses significant hurdles in the expanding landscape of IoT deployments. With IoT devices finding their application in multiple sectors, LoRaWAN has emerged as a popular choice due to its long-range capability and low-power operation. Existing work often fails to adapt to the dynamic nature of the radio environment, leading to reduced network reliability. Learning Automata-based approach offers a promising solution by learning and converging to the best channel among a set of channels for the best communication.

1.1.2 Problem statement

The central challenge in Long-Range Wide Area Networks (LoRaWANs) lies in maintaining high-quality communication the midst of varying radio channel conditions. The limited capacity of the shared radio channel, compounded by multiple end devices sharing the same channel, leads to an unpredictable environment characterized by interference, signal attenuation, and collisions. These factors necessitate a more dynamic approach as traditional static methods often fall short in adapting to such changing conditions.

In response to these challenges, there is a pressing need for an intelligent solution capable of classifying radio channels according to environmental characteristics, ensuring optimal network performance. This thesis aims to address this need by proposing the integration of a sophisticated Learning Automata system known as Hierarchical Discrete Pursuit Learning Automata (HDPA). The HDPA system offers the capability to intelligently select the best channel among multiple options, thereby optimizing communication reliability and throughput in LoRaWAN.

By leveraging the adaptive and learning capabilities of HDPA, this research seeks to enhance the resilience and efficiency of Long-Range Wide Area Networks, ultimately contributing to the advancement of wireless communication technologies in dynamic and challenging environments.

1.2 General Objective

The general Objective of this thesis is to integrate and validate a Learning Automata Based solution for Long-Range Wide Area Network radio channel classification. The aim is to enhance the adaptability and efficiency of LoRaWAN in a dynamically changing environment.

1.2.1 Specific Objectives

- 1) Review the weakness in the existing literature on Learning Automata for their and limitation.
- 2) Integrating Learning Automata with LoRaWAN to enhance learning and convergence for throughput maximization.
- 3) To evaluate the performance of HDPA in terms of accuracy in Long-Range Wide Area Networks and compare it to HCPA.

1.3 Scope and Limitation of the Research

The research will focus on the integration of Hierarchical Continuous Pursuit Learning Automata in LoRaWAN using simulations. The scope includes integration of channel selection algorithms, performance analysis and comparison with HCPA solutions. However, due to resource constraints, the study may not implement the scenario in hard devices.

1.4 Research Questions.

1. What are the parameters to consider when deploying Learning Automata into LoRaWAN communication system?
2. How can HDPA improve the data transmission quality in Long Range Wide Area Networks?
3. How fast and accurately is the HDPA compared to HCPA in terms of converging to the optimal Channel?

1.5 Hypothesis

We hypothesize that the integration of Learning Automata will significantly improve the efficiency and reliability of channel selection in LoRaWAN networks, leading to enhanced network performance and maximize the throughput.

1.6 Significance of the study

The outcomes of this study are expected to contribute significantly to the field of Communication by providing insights into effective channel selection for LoRaWAN networks. The proposed HPA-based approach has the potential to enhance the overall performance of IoT deployment relying on LoRaWAN technology.

1.7 Organization of the Study

The remainder of this Thesis is organized as follows:

- ✦ Chapter 2: Takes the readers through the previous research, identifying the gaps in the Previous Literature and explaining the objectives of this thesis endeavor in tackling those identified deficiencies.
- ✦ Chapter 3: This chapter provides an overview of the approaches and methodologies utilized in this thesis, along with a detailed description of how they were implemented.
- ✦ Chapter 4: Examining the experiments and outcomes derived from the proposed methods, this chapter will also provide a summary of the work and findings of this thesis.

- ✦ Chapter 5: This is the final chapter where the Thesis is concluded and outline the ideas regarding future work for potential improvements.

1.8 Conclusion of Chapter One

This chapter provides an overview of the research emphasis and setting of the study, highlighting the integration of Learning Automata into LoRaWAN to increase transmission Performance. The backdrop of the study has been provided, and the motivation for it has been defined. The issue statement has been well delineated, and the study objectives, encompassing both broad and detailed aims, have been explicitly stated. The hypothesis has been articulated, and the study's scope and significance have been emphasized. The thesis has been structured, delineating a clear path for the forthcoming chapters. The upcoming chapter will provide an extensive examination of pertinent literature in the field, encompassing research methodology, system design, and findings that contribute to the best performance of Learning Automata in terms of convergence to the best channel for good communication.

CHAPTER 2: LITERATURE REVIEW

2.0 Introduction

This chapter provides a relevant literature review of previous works done in this thesis. It will be critically examined in alignment with the specific objectives of this study. This review aims to encapsulate the breadth of scholarly research on these topics, identifying key theories, and findings. By doing so, it will provide a thorough understanding of the current landscape overview.

2.1 LoRa and Long-Range Wide Area Network (LoRaWAN)

Originally developed by Cycleo and acquired by Semtech, LoRa (Long-Range) has grown to be the building block in communication with the Internet of Things (IoT) from its unique configuration: long-range connectivity and efficiency in ratio power consumption using Chirp Spread Spectrum (CSS) signaling to overcome both interference and capture of the signaling when it is below the noise floor.

Operating in 868 MHz in Europe and 915 MHz in USA, it is here the core technology for deployment of IoT applications, which would require a minimum energy consuming and cost-effective connectivity from urban to countryside with large area footprints. More recent studies therefore indicate the growing importance of LoRa in the IoT ecosystem through applications in various fields such as smart agriculture, urban planning, and environmental monitoring [9].

The underlying LoRaWAN network uses LoRa technology for Chirp Spread Spectrum (CSS) modulation, which tolerates the interference and fading of signals. Actually, one of the central reasons that make LoRa ideal for all long-range communications, pretty much over any kind of terrain, is its very inherent resiliency. But with the rising popularity of LoRaWAN, especially in dense urban areas where a lot of devices are being added, the interference question is starting to raise a few eyebrows. The sources of interference are co-channel from the outside and within

LoRaWAN, among others, existing between the other LoRaWAN devices and external sources when they are operating within the same frequency band, for example, ISM band devices [10].

The architecture of the Long-Range Wide Area Network (LoRaWAN) is designed to support vast networks of low-power devices over large geographic areas. It is structured to optimize both the communication range and battery life of these devices, which is essential for the various applications of the Internet of Things (IoT). This structure is particularly effective for deployments where devices need to send small amounts of data over long distances intermittently, rather than streaming data continuously.

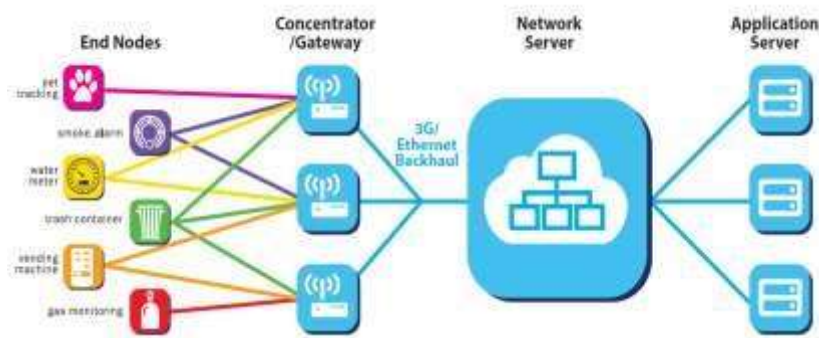


Figure 1: LoRaWAN Network Architecture.

The network infrastructure is built around a star-of-stars topology where end nodes (devices) transmit data to gateways. These gateways act as bridges, receiving messages from multiple end devices and forwarding them to a central network server. The gateways do not process the data; instead, they convert the frequency of the radio message to a suitable format for the backhaul connection to the network server.

The network server plays a crucial role in the architecture. It manages the network's operation, including the handling of join requests from devices, the scheduling of downlinks (messages from the server to the devices), ensuring message security, and performing adaptive data rate adjustments to optimize communication efficiency. The adaptive data rate (ADR) mechanism is vital for managing the data rate and RF output of each device, depending on network conditions and device proximity to the gateway, which helps in conserving the device's battery life.

LoRaWAN specifies three distinct types of devices based on their communication capabilities and power requirements: Class A, Class B, and Class C. Class A devices have the lowest power consumption and communicate based on a schedule of uplink followed by two downlink windows, thus maximizing battery life. Class B devices have scheduled receive slots that offer more receive windows at the cost of increased power consumption. Class C devices are almost always on and can receive messages at any time (except when transmitting), which leads to higher power usage but allows for real-time communication.

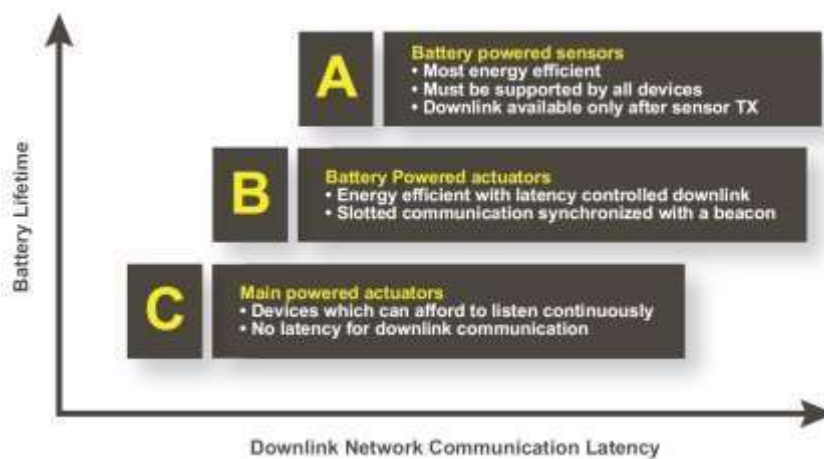


Figure 2: LoRaWAN Network Architecture

2.2 Machine Learning

Machine Learning (ML) is a crucial subdivision of artificial intelligence that concentrates on creating systems that can acquire knowledge from data and make judgments without the need for explicit programming. These systems acquire the ability to carry out tasks by identifying patterns and drawing conclusions from data. Machine Learning can be classified into three primary categories: Supervised learning, Unsupervised learning, and Reinforcement learning. These categories are designed to address various tasks and applications.

Supervised learning, the most prevalent form of machine learning, involves training a model on a dataset where each example is labeled with the correct outcome [11]. The model learns to predict the label from the input features, to make accurate predictions on new, unseen data. This type of learning is commonly applied in scenarios where the

historical data can predict future events, such as in image recognition, speech processing, and predicting consumer behavior [12]. Techniques used include linear regression, logistic regression, support vector machines, and neural networks. Unsupervised learning, in contrast, does not utilize labeled outputs. Instead, it identifies patterns and structures in data independently. The main applications of unsupervised learning include clustering data into groups based on similarities, anomaly detection for identifying unusual data points, and association mining to discover rules that describe large portions of data. Common algorithms employed in unsupervised learning include k-means clustering, hierarchical clustering, and principal component analysis [13].

Reinforcement learning is distinct in that it involves training an agent through the process of trial and error, using rewards and penalties as signals. It is particularly useful in dynamic environments where the algorithm must adapt to changing conditions and make decisions that maximize some notion of cumulative reward. This method is widely used in robotics, gaming, autonomous vehicle navigation, and other real-time decision-making applications. Techniques such as Q-learning, policy gradient, and deep reinforcement learning are often employed to develop sophisticated models that can learn complex behaviors [14].

2.3 Learning Automata

The concept of LA revolves around being an agent that makes decisions within a random environment, selecting optimal actions based on feedback received from the environment in the form of rewards or penalties. Its effectiveness lies in the swift and precise convergence achieved at a low computational cost by leveraging collaborative decision-making among groups of LA, surpassing neural network accuracy, as demonstrated by the Tsetlin Machine. Consequently, LA is demonstrating significant international promise within the realm of machine learning. Various types of LA can be categorized into two main groups: Fixed Structure Stochastic Automata (FSSA) and Variable Structure Stochastic Automata (VSSA). Initially, LA functioned as an FSSA, as proposed by Tsetlin. The category of VSSA encompasses schemes such as Linear Reward-Penalty LRP , Linear Reward-Inaction LRI , and Linear Inaction-Penalty

LIP , where reinforcement learning (RL) mechanisms are integrated into LA's feedback and updating processes, resulting in linear increments in action probabilities. Both continuous and discretized updating methods can be identified in [15].

As mentioned, we can model the LA as seen in this figure where LA is denoted by \mathcal{A} selects actions among a set of Actions giving it to the environment and receiving in return feedback as a reward or penalty denoted by β , where $\beta = 0$ is a Reward and $\beta = 1$ is a Penalty.

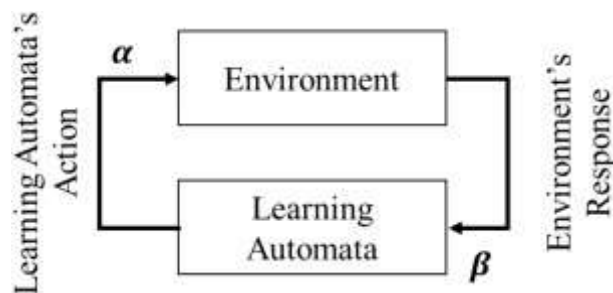


Figure 3: Learning- Automata and Environment model.

2.3.1 The pursuit

The concept of pursuit refers to a method where the automata actively seek to optimize its performance by focusing on actions that are currently estimated to yield the highest rewards. This strategy diverges from traditional LA approaches that typically update their action probabilities based on direct feedback from the environment. In pursuit schemes, the automata employ statistical estimators to forecast the potential reward of each action, thereby guiding the learning process toward the best action by dynamically adjusting its action selection probabilities in favor of those that are estimated to be optimal, significantly speeding up the learning process and enhancing the accuracy. This concept involves using the Maximum Likelihood Estimators to approximate the reward probabilities of various actions within the environment. Once the probabilities are estimated, the LA updates its action probabilities not merely in response to the actual rewards received, but also the estimate [16].

2.3.2 Linear VSSA Schemes

In the variable structure stochastic Automata, we have the Linear Reward- Penalty which takes in account both the reward and the penalty then updates the action

probabilities, and the Linear Reward - Inaction in which we ignore the penalties. The updating scheme for this paper we consider the Linear Reward - Penalty that can be formulated as:

$$P_{\{k\bar{j}\bar{h}(\bar{t})\}}(t + 1) = (1 - \Delta)P_{\{k\bar{j}\bar{h}(\bar{t})\}} \quad (1)$$

$$P_{\{k\bar{j}\bar{h}(\bar{t})\}}(t + 1) = 1 - P_{\{k\bar{j}\bar{h}(\bar{t})\}}(t - 1) \quad (2)$$

For the choosing action that received a Reward as feedback.

$$P_{\{k\bar{j}\bar{h}(\bar{t})\}}(t + 1) = P_{\{k\bar{j}\bar{h}(\bar{t})\}}(t) \quad (3)$$

$$P_{\{k\bar{j}\bar{h}(\bar{t})\}}(t + 1) = P_{\{k\bar{j}\bar{h}(\bar{t})\}}(t) \quad (4)$$

For the action that received a penalty as feedback it remains the as the previous.

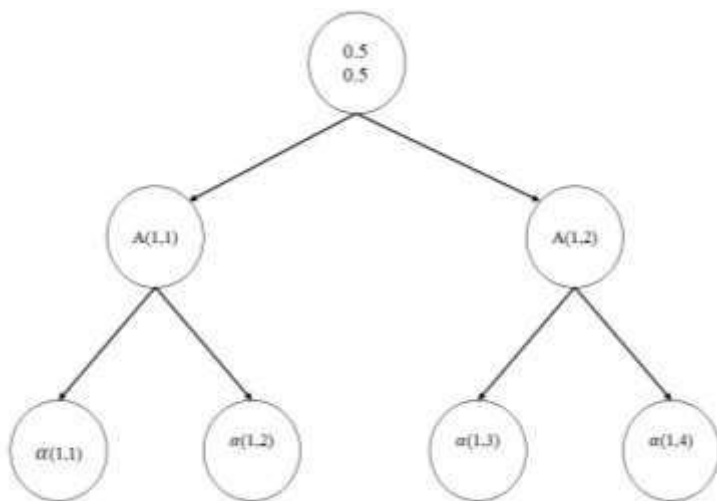


Figure 4: Linear Reward- Penalty scheme.

Traditionally, the channel selection problem is tackled using a variety of methods that are designed to optimize the performance of the network and minimize interference following this perspective the authors in [17] made a significant contribution to the field of cognitive Radio Networks by enhancing the efficiency and reliability of secondary users. They first introduce a novel approach that integrates the concept of Bayesian Learning Automata with the ability to switch the channel. Their mechanism allows the system to adaptively select channels based on updated probabilistic of channel

availability, thereby improving the likelihood of successful communication without interference from primary users. The is contribution not only addresses the practical challenges of channel selection in densely populated network environments but also enhances the throughput and overall performance of cognitive radio networks. This approach could potentially influence future designs and implementations of cognitive radio systems, marking a significant step forward in the evolution of communication technology in dynamic and unpredictable environments.

Learning Automata has also been used for cyber-physical systems in [18]. To dynamically determine the number of gateway level in the network, thereby optimizing network performance by reducing packet collisions and delay. The algorithm assigns priority and time limits to each packet, ensuring efficient and prioritized data handling. The proposed algorithm with – two-level MAC systems and demonstrates improved network throughput, reduced collision rates, and fewer packet drops. Learning Automata is also used in LoRaWAN for Event Detection in [19] to enhance the efficiency of event packet transmission enabling the network to switch intelligently between different MAC protocols TDMA and Slotted ALOHA.

Additionally, in [20] they proposed a method that uses a «Lightweight decentralized reinforcement Learning approach» to dynamically select communication channels aiming to reduce collision and enhance the overall network efficiency. The technique leverages the concept of multi-armed bandit (MAB) algorithms they used the ‘arms’ as the available communication channels and the ‘reward as the successful transmission of data without collision. We employ the same method in our paper.

In [21] the authors proposed a novel approach using Reinforcement Learning to enable end devices in LoRaWAN to autonomously determine their transmission times, in their system each device operates as an independent agent that learns optimal transmission times based on their interaction with the network environment, characterized by the state of the network at any given time to maximize a defined reward function. The model in this work assumes a decentralized framework where no central controller is required, allowing adaptability to different network conditions, and highlighting how devices can adapt their behavior based on the outcomes of past actions to improve

network efficiency. The results indicate that their methods outperform the traditional ALOHA-based protocol by reducing the packet collision and improving the data throughput.

A Reinforcement based channel selection in Industrial IoT was furthermore proposed in [22]. The authors proposed a spectrum handoff (SH) strategy integrated with Reinforcement Learning (RL) algorithms to improve the efficiency of channel selection when a cognitive radio user must vacate a channel upon the return of the primary user. Their contribution lies on the dual reinforcement learning model having the first one with the purpose of predicting the likelihood of future channel availability while the second accesses the channel quality providing a view of the landscape.

They detail the implementation of their work in a candidate channel list and a backup channel list, both crucial for reducing the selection time, improving the overall throughput of the network, and minimizing the handoff delay typically experienced in CR networks. The work does not only ensure efficiency but also enhances the quality of service by reducing the likelihood of selecting poor-quality channels.

In addition, the authors in [23] focus on the practical application of Tug-of-War (TOW) dynamics, inspired by the multi-armed bandit problem, allowing IoT devices to dynamically select the most optimal communication channels in environments where interference from various sources can degrade performance. They used an algorithm that adapt to changing network condition without requiring extensive computational resources, particularly for IoT devices that have limited processing capabilities. The experiments conducted in a real environment with coexisting IoT systems demonstrate that the algorithm can improve packet delivery rate compared to static channel assignment strategies.

In [24] proposed an optimization of channel access methods for LoRaWAN focusing on the slotted Aloha protocol, they present a details simulation aiming at reducing collision probabilities by introducing structured time slots for transmissions and examining the impact of network parameters like slot length, guard times and cross traffic to improve the data transmission reliability and efficiency of LoRaWAN networks. The study shows that the most effective slot lengths are the ones the exceed the time on air (ToA) of messages. Additionally, the paper investigates the role of guard times which are crucial in mitigating the risk of packet collisions due to timing errors

and synchronization issues. By contrasting slotted Aloha's performance against pure Aloha methods under identical conditions, they demonstrate that slotted Aloha consistently results in lower collision rates and supports a higher device count per network. Slotted Aloha is also proposed in [25] focusing on how the Spreading factor influences the network's data rate and range aiming to achieve an optimal balance between throughput and delay of the communication, they employ the Markov chain to evaluate different SF allocation strategies and their impact on network performance, and introduce a framework for predicting and improving the performance of Slotted Aloha-based LoRaWAN system by analyzing and optimizing re-transmission probabilities.

In [26] the authors proposed a game-theoretic approach to optimize channel selection among various operators in LoRaWAN aiming to maximize the total normalized throughput, where each operator is independently selfishly to select the channel that maximizes their throughput. To establish stability and efficiency in channel selection they used both Nash equilibrium and correlated equilibrium concepts. They utilized both centralized methods as linear optimization and decentralized method using reinforcement learning then the algorithm was evaluated for effectiveness through simulation and the comparison shows that it improves the throughput and packet delivery ratios.

Massive multi-armed bandit is proposed in [27] to optimize the network in IoT, especially introducing Decreasing -Order- Reward-Greedy (DORG) and Decreasing Order- Fair-Greedy (DOFG) strategies to enhance network performance by optimizing the balance between energy consumption and packet loss. The simulation demonstrates superior results than the traditional Adaptive Data Rate (ADR). However, their strategies do not consider the variability in the mean rewards of arms between different players nor account for non-orthogonal channels.

To address the challenge of inter-network interference in Long Range (LoRa) communication the authors in [28] proposed a novel solution known as Online Current Transmission (OTC) at LoRa gateway. They designed the system to enhance network throughput by enabling the gateway to handle and recover multiple collided packets simultaneously and achieving this through detecting the preamble, packet decoding, and using LoRa modulation and demodulation information.

The authors in [29] introduce a novel Time Division Duplex (TDD) enhanced by reinforcement Learning to address reliability and scalability challenges inherent in LoRa networks. The Time Division duplex is employed to separate uplink and downlink transmission where each device can access the channel's state before transmitting using Channel Activity Detection. Following the idea of avoiding interference, the authors in [30] examined the issues of – adjacent-channel interference by providing a comprehensive analysis of how they affect the LoRa network, especially in scenarios with a large number of transmitted devices. They proposed cellular communication networks as a strategy for channel allocation and a duration division mode to handle the limitations in channel availability. The proposed solution was validated through simulation and shown that it can be applied in IoT applications with large-scale sensors. In [31] the authors proposed a new approach to analyze the impact of combined expert and distribution on transmission parameters by introducing the Modified EXP3(M-EXP3) algorithm as a novel method to optimize the transmission parameter selection in LoRaWAN networks aiming to reduce the energy consumption and packet losses allowing for arm switches during transmission and ensuring controlled regret with respect that to policies that permit parameter change. The algorithm outperforms the traditional EXP3 algorithm in terms of energy efficiency and successful packet reception ratio.

To mitigate collisions some authors, present an innovative approach for mitigating radio collisions in IoT devices operating in unlicensed bands.

Introducing low-cost machine learning algorithms directly into the devices to enable intelligent real-time decision-making regarding spectrum use without requiring a change to the existing LoRaWAN protocol. The experiment proved that the algorithm could successfully learn and adapt to optimize their radio frequency usage and reduce the collision [33]. Dynamic parameter allocations are proposed using several strategies for managing transmission and improving the performance in terms of energy consumption and Network throughput with the use of RL [34].

To address the challenges of packet collisions, power consumption, and network performance in LoRa Networks a novel Deep reinforcement learning for adaptive PHY layer transmission parameter selection is proposed. The authors designed the algorithm to be computationally efficient, ensuring that it can operate effectively on resource-

constrained devices. Their simulation demonstrates high improvement in packet delivery ratio (PDR) for some scenarios in comparison to existing state of art [38]. In this same perspective paper [39] proposes a hybrid approach that makes a combination of random and scheduled access strategies to enhance the network ability and handle the event-triggered traffic. They introduce a novel Medium Access Control (MAC) layer that facilitates better synchronization and dynamically schedules post-event-detection allowing for a better-organized data transmission process and improving overall data delivery performance [40].

Reinforcement learning based – adaptive is proposed to optimize transmission parameters on the network server side, addressing the trade-offs between energy consumption and packet delivery rate by dynamically adjusting the transmission parameters such as SF, Transmissions power, CF, and coding rate and the evaluation shows that the LP-MAB algorithm outperforms traditional methods in terms of energy consumption while maintaining competitive packet delivery ratio [41].

A novel application of the multi-armed-bandit algorithm as a solution to boost network performance and energy efficiency is proposed by authors in [43]. They evaluate a modified Discount Upper Confidence Bound (DUCB) approach to optimize transmission parameters like SF in the LoRa Network by dynamically adapting to the changing environment then compare the proposed solution with the existing solution and the proposed algorithm shows significant improvement in terms of the number of successful transmissions in LoRa Networks.

The Authors in [44]. explored advanced channel selection techniques to enhance communication efficiency in large-scale IoT, addressing the challenge of network congestion in environments populated by a vast number of devices by implementing a multi-armed Bandit approach, they proposed a dynamic method for channel allocation that promises to optimize network traffic distribution and minimize congestion without the requirement of centralized control. The evaluation was done with 1000 of IoT devices focusing on performance metrics such as Frame Success Rate (FSR). To balance the trade-offs between energy consumption and packet loss they demonstrate that traditional Adaptive Data Rate (ADR) can be significantly enhanced by utilizing lightweight learning algorithms and employing multi-armed-bandit, nodes that can autonomously adjust their communications parameters such as spreading factor and transmissions power allowing nodes to learn from their environment [45].

In [46]. The authors investigate on how irregular traffic patterns can impact the functionality of LoRaWAN networks. Delving into the network’s ability to manage high levels of traffic bustiness, for ensuring reliability and timely data transmission. The study contributes significantly by providing a detailed understanding of traffic patterns and their effects on network performance including data transmission rates, packet delivery ratios, and overall network reliability.

Identified Gaps.

Authors	Topic/Title	Contribution	Limitations
Shahzeb Ahsan; Syed Ali Hassan; Ahsan Adeel; Hassan Khaliq Qureshi [35].	‘Improving channel utilization of LoRaWAN by using Novel Channel Access Mechanism’	The authors proposed three novel channel access protocols: LoRa Binary Exponential Delay (LoRa- BED), LoRa Binary Exponential Back-off (LoRa- BEB), LoRa- Binary Exponential Hybrid (LoRa-BEH). Aiming to sense the channel before transmitting data, their approach allows the system to decide whether the channel is free or busy. They modified the existing LoRaWAN approach by integrating a modified listen-before-talk (LBT).	One significant limitation in the paper is the assumption that all nodes are stationary and equally spread around the gateway, which does not account for dynamic and varied environments where IoT devices might be mobile.

		Through simulation they evaluate the performance of these protocols in terms of channel utilization, packet delivery ratios, and collision rates; the results indicate significant improvement compared to traditional ALOHA based access	
--	--	---	--

<p>Lykov Yuri; Lykova Anna; Savenko Stepan [36].</p>	<p>‘Research on throughput Capacity of LoRaWAN Communication Channel’</p>	<p>This paper provides an understanding of how different configurations impact the performance of LoRaWAN channels.</p> <p>The paper demonstrates how SF and Payload influence the airtime, which in turn affects the throughput capacity of the network.</p> <p>The results provide practical guidance showing how to optimize device settings for improved data transmission efficiency.</p>	<p>One major limitation of this paper is the exclusion of ADR which is a crucial feature for optimizing network performance.</p> <p>They do not also consider other influential factors like channel interference and network congestion that impact the network efficiency and reliability.</p>
--	--	--	--

<p>Lilian Besson; Remi Nnefoi; Christophe Moy [37].</p>	<p>GNU Radio Implementation of MALIN: Multi-Multi-Armed Bandits Learning for Internet of Things Networks</p>	<p>The authors demonstrate and evaluate multi-armed bandit algorithm to optimize network access in IoT using a simple ALOHA-based protocol. Using UCB1 and Thompson Sampling intelligent devices learn and make decisions on channel selection based on past iterations. Their work reduced packet loss and improved the efficiency of network communications, particularly in environments with multiple interference networks</p>	<p>The assumption of stationary traffic scenarios that do not always reflect dynamic IoT network conditions in real-world</p>
<p>Azizi Farzad, Benyamin Teymuri, Rojin Aslani, Mehdi Rasti, Jesse Tolyaneny, and Pedro HJ Nardelli</p>	<p>Mix-MAB: Reinforcement Learning based Resource Allocation Algorithm for LoRaWAN</p>	<p>To optimize the packet delivery ratio, the authors in [42]. combined Exploration and exploitation (EXP3) AND Successive Elimination (SE) allowing end devices to autonomously select the optimal transmission settings</p>	<p>The scalability issues arise; the algorithm needs to manage a large set of iterations as more devices join the network. The lack of centralized control enables the algorithm to handle data efficiently</p>

		without central coordination, the	
--	--	-----------------------------------	--

		simulation demonstrates its superiority over the previous approach and traditional Adaptive Data Rate (ADR). MIX-MAB achieves faster convergence and a higher Packet delivery ratio.	
--	--	--	--

<p>Li, Xiaowu, Junjie Xu, Runxin Li, Lianyin Jia, and Jinguo You (2024)</p>	<p>Advancing Performance in LoRaWAN Networks: The circular Region Grouped Bit-Slot LoRa MAC Protocol</p>	<p>The authors proposed a novel MAC Protocol, which segments the network area into circular regions based on the distance between end devices and the gateway allocating different spreading factors accordingly aiming to reduce collisions and enhance overall network performance, significantly improving scalability and reducing latency in dense LoRaWAN deployments.</p> <p>Experimental results show that the proposed solution improves packet delivery rate and</p>	<p>The limitation of this paper is the assumption of static conditions for both the gateway and the end devices which may not capture the dynamic nature where devices often move.</p>
---	---	--	--

		<p>reduces collision compared to pure ALOHA.</p>	
--	--	--	--

<p>Frank Loh, Noah Mehling and Tobias Hobfeld</p>	<p>Towards LoRaWAN without data loss: Studying the performance of Different Channel Access Approaches</p>	<p>The paper contributes to improving data transmission reliability in LoRaWAN networks by delving into different channel access to mitigate data loss due to message collisions, the authors present proposed a time-scheduled channel access method designed to nearly eliminate collisions. The solution was proved efficacy than traditional random access.</p>	<p>The limitation of this paper is the constraint imposed by LoRaWAN duty cycle regulation which significantly impacts the feasibility and scalability of the proposed approach. Another limitation revolves around the challenges related to device synchronization and clock drifts.</p>
<p>Ameer Ivoghlian, Zoran Salcic, and Kevin I-Kai Wang</p>	<p>Adaptive Wireless Network Management with Multi-Agent Reinforcement Learning</p>	<p>The paper introduces an innovative framework using multi-agent deep reinforcement learning to dynamically and efficiently manage network resources. The novelty of the</p>	<p>One significant limitation is the lack of experience sharing among individual nodes in the network. Each node operates independently with the reinforcement</p>

		<p>framework lies in its utilization of a custom reward function that incorporates application awareness and fairness, ensuring both node-specific and network-wide objectives are met. The paper includes the development of an autonomous network management system that adjusts to network condition in real-time and promote fair access to network resources and ensure optimal performance across various applications. The paper demonstrates through simulation that their multi-agent approach significantly outperforms traditional single agent systems in terms of responsiveness and network resource allocation.</p>	<p>learning agent.</p>
--	--	--	------------------------

2.4 Conclusion

In this chapter, the detailed literature was given. Methodologies and challenges in channel selection overview are provided through an examination of various approaches ranging from traditional fixed algorithms to more dynamic machine learning techniques. It becomes evident that while considerable advancements have been made, significant challenges remain.

CHAPTER 3: RESEARCH METHODOLOGY

3.0 Introduction

This chapter delve into the overview of the research methods that are used for the work. It highlights the methods and approaches to be used and how results are presented. Mainly scientific methods are used for this project, however quantitative approach is appearing because results from experiments are presented. Our approach lies more on scientific research methods, and it covers almost the whole project.

3.1 System model

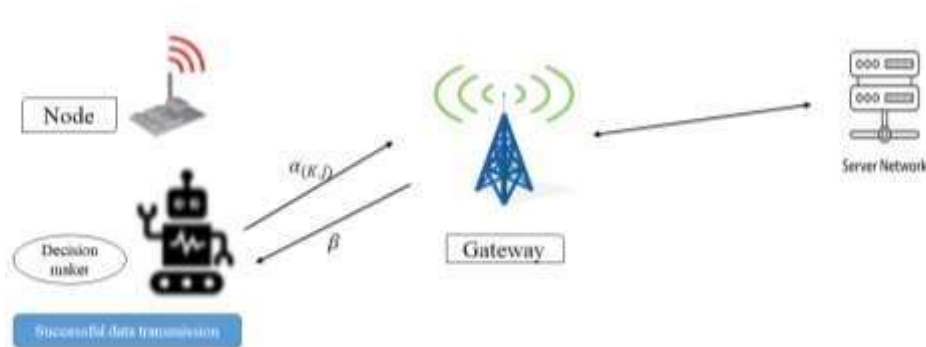


Figure 5: Proposed LoRa Network Model

This model consists of several key components that work together to optimize data transmission through learning and feedback mechanisms.

The node is an end device equipped with a radio transmitter that sends data packets. The decision maker represented by the robot, implements the Hierarchical Discrete Pursuit Learning Automata (HDPA) algorithms. Its role is to select the optimal radio channel for data transmission based on past transmission success rate feedback from the gateway. The decision maker updates the probability distribution of channel selection using the number of times it received a reward over the number of times it was selected.

The feedback loop represents the action taken by the decision maker regarding which channel to use for the next data transmission. Beta represents the feedback received from the gateway. If the transmission is successful, the decision maker receives positive

feedback reinforcing the choosing channel. If unsuccessful, the decision maker receives negative feedback, decreasing the likelihood of selecting that channel again.

3.2 Proposed algorithm flowchart

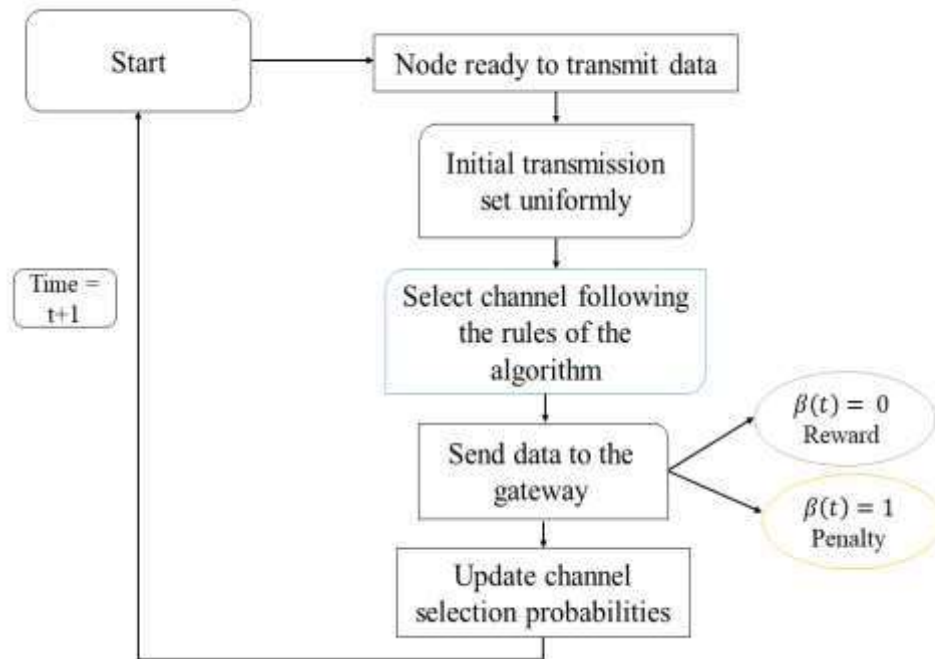


Figure 6: Simulation flow chart

In terms of process flow, the node transmits data to the gateway using a channel selected by the decision-maker. The gateway provides feedback on the transmission's success. Positive feedback increases the probability of selecting the successful channel in the future while negative feedback decreases the probability of encouraging the decision maker to explore other channels. The decision maker then continuously updates its channel selection probabilities based on the feedback, adapting to the dynamic network environment to optimize data transmission reliability.

3.3 Development Approaches

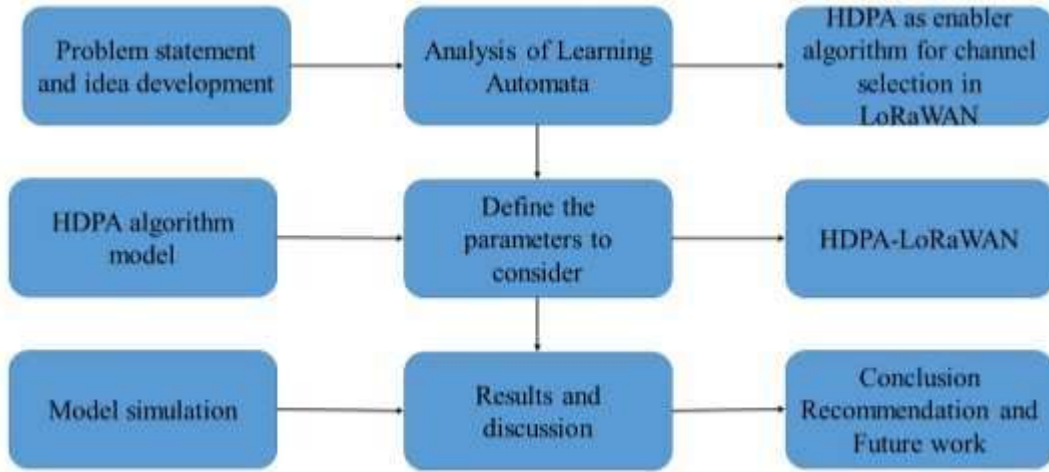


Figure 7: Research Development Approach

3.4 Mathematical Development

Algorithm $t=0$

Loop

1. Depths from 0 to $K-1$:

$\mathcal{A}_{[0,1]}$ selects an action by randomly sampling as per its action probability vector $[p_{\{1,1\}}(t), p_{\{1,2\}}(t)]$.

We denote $j_1(t)$ as the chosen action at depth 0 with $j_1(t) \in \{1,2\}$.

$\mathcal{A}_{\{1, j_1(t)\}}$, chooses an action and activates the next LA at depth «2».

The process continues until $K-1$ which is the level that chooses the action.

2. Depth K :

✧ The index of the action chosen at depth K denoted $j_K(t) \in \{1, \dots, 2^K\}$.

✧ Update the estimated chance of reward based on the response received from the environment at leaf depth K :

$$u_{\{K, j_K(t)\}}(t+1) = u_{\{K, j_K(t)\}}(t) + (1 - \beta(t))$$

$$v_{\{K, j_K(t)\}}(t+1) = v_{\{K, j_K(t)\}}(t) + 1$$

$$\hat{d}_{\{K, j_K(t)\}}(t+1) = u_{\{K, j_K(t)\}}(t+1) / v_{\{K, j_K(t)\}}(t+1) .$$

For the other action at the leaf, where $j \in \{1, \dots, 2^K\}$ and $j \neq j_K(t)$:

$$u_{\{K,J\}}(t + 1) = u_{\{K,J\}}(t)$$

$$v_{\{K,J\}}(t + 1) = v_{\{K,J\}}(t)$$

$$\hat{d}_{\{K,J\}}(t + 1) = u_{\{K,J\}}(t + 1) + \gamma \sum_{\{K',J'\} \in \mathcal{A}_{\{K,J\}}(t)} v_{\{K',J'\}}(t) \hat{d}_{\{K',J'\}}(t)$$

3. Define the reward estimate recursively for all subsequent actions along the path to the root, $k \in \{0, \dots, K - 1\}$, where \mathcal{A} at any one level inherits the feedback from the \mathcal{A} at the level below:

$$\hat{d}_{\{k,j\}}(t) = \max(\hat{d}_{\{k+1,2j-1\}}(t), \hat{d}_{\{k+1,2j\}}(t))$$

4. Update the action probability vectors along the path to the leaf with the current maximum reward estimate:

Each $\mathcal{A} j \in \{1, \dots, 2^k\}$ at depth k where $k \in \{0, \dots, K - 1\}$ has two actions $\alpha \{k + 1, 2j - 1\}$ and $\alpha \{k + 1, 2j\}$.

We denote the larger element between $\hat{d}_{\{k+1,2j-1\}}(t)$ and $\hat{d}_{\{k+1,2j\}}(t)$ as $j_k^{h+1}(t) \in \{2j - 1, 2j\}$ and the lower reward estimate as $\bar{j}_k^{h+1}(t) = \{2j - 1, 2j\} \setminus j_k^{h+1}(t)$.

✧ Update $P_{\{k+1, j_k^{h+1}(t)\}}(t)$ and $P_{\{k+1, \bar{j}_k^{h+1}(t)\}}(t)$ using the estimate $\hat{d}_{\{k+1,2j-1\}}(t)$ and

$\hat{d}_{\{k+1,2j\}}(t)$ for all $k \in \{0, \dots, K - 1\}$ as:

If $\beta(t) = 0$ Then

$$P_{\{k+1, j_k^{h+1}(t)\}}(t + 1) = \min(P_{\{k+1, j_k^{h+1}(t)\}}(t) + \Delta, 1),$$

$$P_{\{k+1, \bar{j}_k^{h+1}(t)\}}(t + 1) = 1 - P_{\{k+1, j_k^{h+1}(t)\}}(t + 1),$$

Else

$$P_{\{k+1, \bar{j}_k^{h+1}(t)\}}(t + 1) = P_{\{k+1, \bar{j}_k^{h+1}(t)\}}(t), \quad k+1$$

$$P_{\{k+1, j_k^{h+1}(t)\}}(t + 1) = P_{\{k+1, j_k^{h+1}(t)\}}(t), \quad k+1$$

Endif

5. For each Learning Automata, if either of its action selection probabilities surpasses a threshold T , with T being a positive number close to the unity, the action probability will stop updating, meaning the convergence is achieved.

6. $t = t + 1$

End Loop

3.5 Software Environment

MATLAB is chosen for its robust capabilities and extensive support for simulations involving complex algorithms and network models.

3.6 System Simulation

This process combines theoretical modeling with practical experiments to validate the hypothesis that HDPA can enhance the efficiency and reliability of channel selection in LoRaWAN networks.

The simulation setup begins with the configuration of the node to transmit data packets. This node represents an end device in the LoRaWAN network, equipped with a radio transmitter. The node interacts with the gateway, responsible for receiving the transmitted data. The gateway acts as an intermediary, forwarding the data to a server network for processing and storage.

At the heart of the simulation is the decision maker, which implements the HDPA algorithm. The decision maker selects the optimal radio channel for data transmission based on the feedback from the previous transmissions. This feedback involves the gateway providing success or failure notifications for each transmission, which the decision maker uses to update its channel selection probabilities.

The simulation is conducted in a Matlab software environment, chosen for its robust capabilities in handling complex algorithms, providing a platform for running extensive simulations to assess performance under various conditions. The simulation parameters include the node, the channel available the successful data transmission. Throughout the simulation, key performance metrics are monitored. Including accuracy, the overall network throughput, Std, and speed. By analyzing these metrics, the effectiveness of the HDPA can be evaluated.

3.7 Simulation Variables

Variable	Symbol	Description
Number of channels	N	Total number of available channels in the network
Initial action probability	$P(0)$	Initial probability vector for channel selection, $P(0) = [p_1, p_2, \dots, p_N]$
Reward	R	Reward metric for successful transmission on a channel such as PDR, SNR
Learning Rate	δ	Step size for probability updates.
Hierarchical levels	L	Number of levels in the HDPA hierarchy
Convergence threshold	B	Threshold for convergence, indicating when the algorithm has likely found the optimal channel
Maximum iteration	T	Maximum number of iterations for the simulation
Action selection probability	P_i	Probability of selecting channels at iteration
Reward estimate	d_i	Estimate of the reward for channel i
Channel State	S	The state of each channel is either idle or busy

3.8 Conclusion

In this chapter, we outlined the research methodology employed to investigate the efficacy of the HDPA algorithm in optimizing channel selection for LoRaWAN. Combining theoretical modeling with practical simulation to validate the proposed algorithm's performance. We first detailed the system model, which includes key components such as the node configuration, gateway, and decision maker, each plays a critical role in the HDPA-based channel selection process. Then we discussed the development approach emphasizing the mathematical foundations and algorithmic steps. Through this methodology, we aimed to achieve a comprehensive understanding of how HDPA can enhance the adaptability and efficiency of LoRaWAN networks. The subsequent chapter will present the results obtained from these simulations.

CHAPTER 4: RESULTS AND DISCUSSION

4.0 Introduction

This chapter presents the outcomes of employing Learning Automata into LoRaWAN, giving the critical importance of efficient channel selection network performance aiming to test the effectiveness of HDPA. We present comprehensive results from the simulation that was conducted and discuss the implications of these findings. This analysis not only highlights the strengths of HDPA but also compares it with HCPA.

$$\text{Mean} = \frac{\sum_{ni=1} X_i}{n} \quad (4)$$

Where X_i is the number of iterations in the i -th trial. n :
the total number of iterations.

$$\text{Variance} = \frac{\sum_{ni=1} (X_i - \text{Mean})^2}{n} \quad (5)$$

$$\text{Standard deviation} = \sqrt{\text{Variance}} \quad (6)$$

$$\text{Accuracy} = \frac{\text{Number of successful trials}}{\text{Total number of trials}} \quad (7)$$

The performance of HDPA is demonstrated here using the formula above and carrying out the simulation results to ensure the effectiveness of our simulations; we set our number of iterations to be 9000 and 10,000, using 200 experiments expecting that the HDPA with the highest successful transmission probability would converge faster and select the best channel

A ₁	A ₂	A ₃	A ₄	A ₅	A ₆	A ₇	A ₈
0.199	0.282	0.394	0.499	0.681	0.698	0.971	0.999

Table 1: List of successful data transmission probability for the 8 channels.

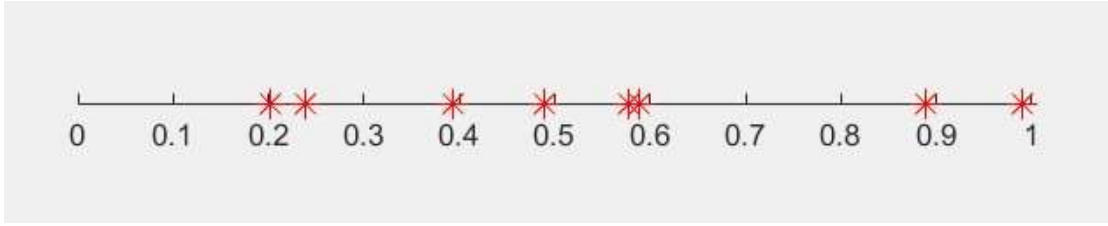


Figure 8: Reward probabilities for 8 channels

The simulation was done for the environment with 8 channels on a benchmark successful transmission probabilities list in the figure showing the probability of the actions with successful transmission meaning the action with $\beta = 0$ which is a reward from the environment.

Our simulation shows that the HDPA with a small Learning parameter can converge to the optimal channel with highly successful data transmission, and a higher learning parameter leads to fast convergence to the optimal channel, however, when we set the Learning parameter higher than 0.00087 the algorithm did not converge to the best channel with successful transmission probability. Therefore, to find the optimal channel with a higher speed of convergence we decreased the Learning parameter step by step until we achieved 98.78% accuracy from this value the algorithm converged to the optimal channel but took all the iterations that were set. The Mean value to converge to the optimal channel for the 200 experiments with the convergence criterion of 0.99 was 6,279.64 confirming that the HDPA has achieved a 0.99 probability of choosing one of the channels with the std of 131.36% on the benchmark probabilities.

Parameter	HDPA
Mean	6,279.64
Std	131.36
Accuracy	98.78%
Learning parameter	$8.7e^{-4}$

Table 2: Result of our simulations for 8 channels

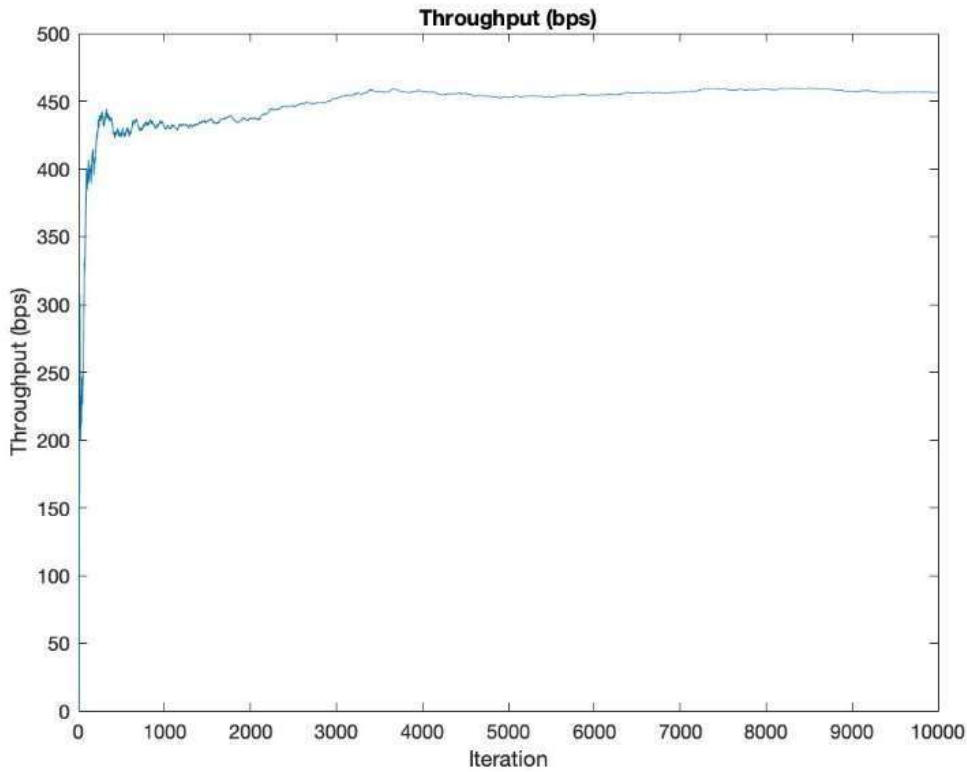


Figure 9: Throughput

The graph represents the throughput as a function of iterations for the simulation. At the beginning of the iterations, there is a rapid increase in throughput indicating the HDPA quickly starts identifying and selecting more optimal channels leading to an immediate improvement in the efficiency of data transmission, after the initial rise we observe that it stabilizes around 450bps. Showing that the HDPA has converged to a set of optimal channels, and the channel selection process become stable maintaining a consistent high throughput. The graph demonstrates the effectiveness of the algorithm in quickly learning and converging to the best channel for data transmission in a dynamic environment with the ability to adapt and optimize the network performance efficiently.

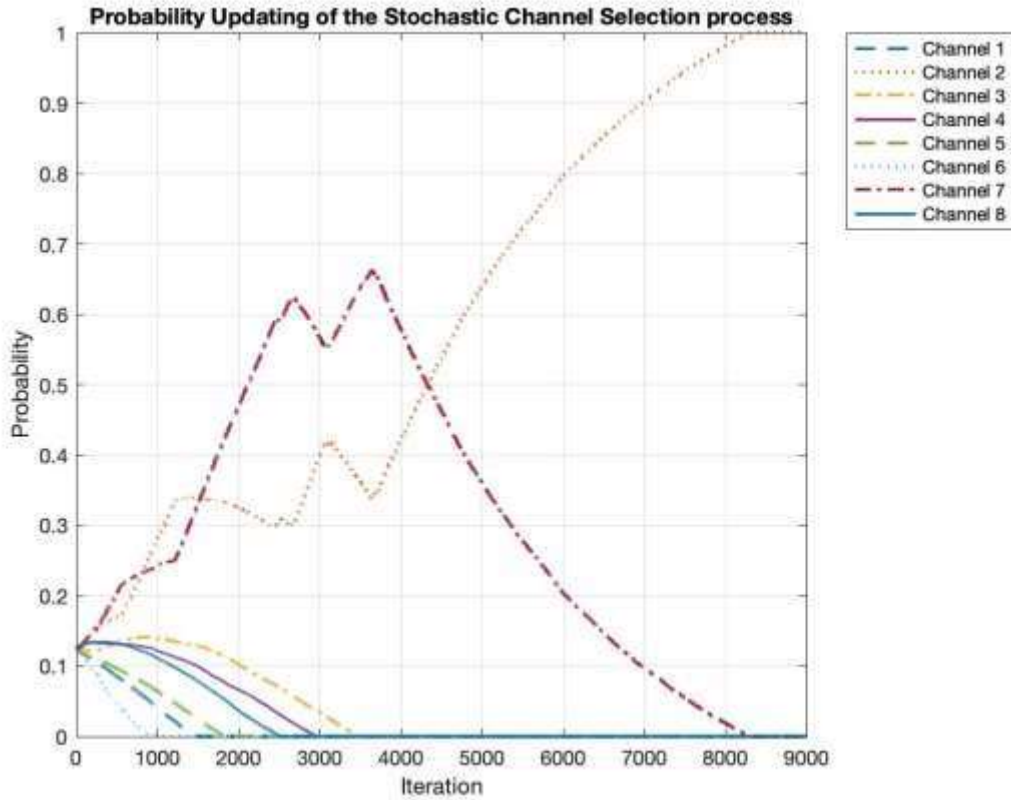


Figure 10: Channel selection updating probability for successful transmission.

This figure directs the reader into the process of selecting the best channel among the set of our 8 channels, we try them for communication, one channel gives statically higher successful probability for transmitting messages others have lower probabilities. In the beginning, we had no idea which channel was the best Learning automata help to converge to the best channel to be used so that we have maximum throughput. We notice that channels 2 and 7 have been preserved by the algorithm HDPA as the best and the second-best channels. In a certain iteration around 3000, we observe that channel 7 outperforms channel 2, however, due to the randomness of channel selection the decision-making process leads to the selection of channel 2 as the optimal one among the 8.

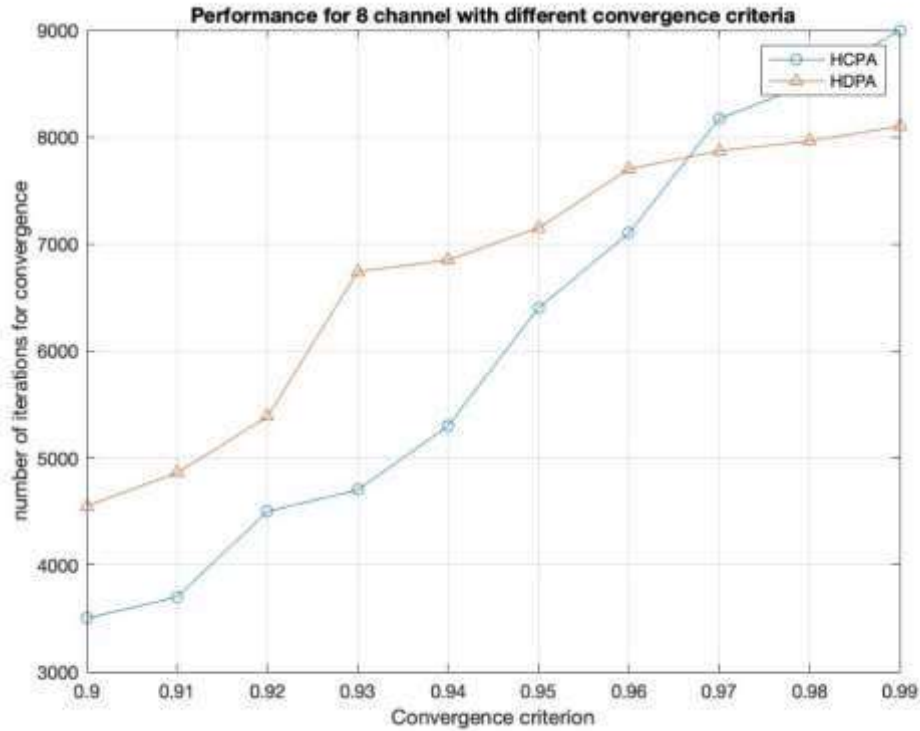


Figure 11: Number of iterations for convergence for 200 experiments.

Parameters	HDPA	HCPA
Mean	6,279.64	6,778.34
STD	131.36	117.12
Accuracy	98.78%	93.89%
Learning parameter	$8.7e^{-4}$	$6.9e^{-4}$

Table 3: Comparison between HDPA and HCPA

From the graph above and the table, we have both comparisons of HDPA and HCPA. The table resorts to the Mean, the Std, and the accuracy of both algorithms. On this graph, it appears that when we set a convergence criterion to 0.9 HCPA outperformed HDPA because it took about 3500 to converge while HDPA took over 4500 iterations to converge as long as we converge for the 200 experiments and when the convergence criteria were set to 0.99 which is the criteria for our successful data transmission, we observed that before converging to the optimal channel HCPA is faster than HDPA as we get closer to the optimal channel for example 0.97 HDPA become faster than the HCPA and converge to the best channel with about 8000 iterations while HCPA required 9000 and none of them was less than 9000. The best learning parameters for both HDPA and HCPA were respectively 0.00087 and 0.00069. Considering the table

one can observe that HDPA required 6279.64 on the mean average in terms of std HDPA has more variation on the other hand HCPA required 6778.34 on the mean average with a std of 117.12. In terms of accuracy, HDPA is more accurate than HCPA as we can see in the table HDPA has almost 99% accuracy making our system better than HCPA in terms of channel selection.

CHAPTER 5: CONCLUSION, RECOMMENDATION AND FUTURE WORK

5.1 Conclusions

This dissertation has successfully explored the efficacy of Hierarchical Discrete Pursuit Learning Automata (HDPA) in the context of channel selection for Long Range Wide Area Networks (LoRaWAN). The study aimed to address the complexities and dynamic nature of radio channel conditions by employing an advanced machine learning strategy, which outperformed the Hierarchical Continuous Pursuit Learning Automata (HCPA) in terms of accuracy and convergence speed. The results convincingly demonstrate that HDPA can effectively enhance the channel selection process, thereby improving the reliability and efficiency of LoRaWAN deployments.

The significance of integrating Learning Automata into LoRaWAN has been substantiated by simulations and comparative analyses. Shows that HDPA not only adapts more adeptly to fluctuating environmental conditions but also ensures high throughput with the ability to adapt to dynamically optimize channel selection making it pivotable in minimizing transmission failures and interferences, which are critical factors in the scalability and performance of IoT networks.

5.2 Recommendation

Considering the findings from this research, it is recommended that HDPA algorithms be considered for integration into existing and future LoRaWAN infrastructures.

Network engineers are encouraged to adopt this advanced channel selection mechanism to significantly enhance the operational efficiency of their networks. Implementing HDPA can lead to more reliable and consistent network performance, especially in environments characterized by high device density and variable channel conditions.

5.3 Future Work

The successful integration of HDPA in this study opens an avenue for future research. One potential area is expanding the scope of Learning Automata to manage interference in densely populated IoT networks represents a critical research frontier. Future studies

could also focus on real-world trials to better understand the practical challenges and opportunities of implementing these algorithms in diverse environments.

References

- [1] E. S. R. A. M. S. S. R. Imane Cheikh, "Throughput-Delay Tradeoffs for SlottedAloha-based LoRaWAN Networks," *International Wireless Communications and Mobile Computing (IWCMC)*, pp. 2020-2025, 2021.
- [2] Honggang Wang, , Peidong Pei, Ruoyu Pan, Kai Wu, Yu Zhang, Jinchao Xiao, and Jingfeng Yang, "A Collision Reduction Adaptive Data Rate Algorithm Based on the FSVM for a Low-Cost LoRa Gateway," *Mathematics*, vol. 10, no. 3920, 2022.
- [3] Xuan Zhang, Lei Jiao, Ole-Christoffer Granmo, B. John Oommen, "Channel Selection in Cognitive Radio Networks: A Switchable Bayesian Learning Automata Approach," *2013 IEEE 24th International Symposium on Personal, Indoor and Mobile Radio Communications: Mobile and Wireless Networks*, pp. 2362-2367, 2013.
- [4] X. V. P. T.-P. B. M. J. M.-S. T. W. Ferran Adelantado, "Understanding the Limits of LoRaWAN," *IEEE Communications Magazine in January 2017*.
Please, cite the IEEE, vol. 55, no. 9, pp. 34-40, 2017.
- [5] L. J. X. Z. A. Y. B. J. O. Rebekka Olsson Omslandseter, "The Hierarchical Discrete Pursuit Learning Automaton: A Novel Scheme With Fast Convergence and Epsilon-Optimality," *IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS*, 2022.
- [6] L. J. B. J. O. Rebekka Olsson Omslandseter, "Pioneering approaches for enhancing the speed of hierarchical LA by ordering the actions," *Information Science*, vol. 119487, p. 649, 2023.
- [7] "A Technical Overview of LoRa and LoRaWAN," LoRa Alliance, [Online].
Available: <https://lora-alliance.org/>.
- [8] M. L. A. Z. Alexandre Figueiredo, "A Fair Channel Hopping Scheme for LoRa Networks with Multiple Single-Channel Gateways," *Sensors*, vol. 22, no. 14, p. 5260, 2022.
- [9] K. C. Wiklundh, "Understanding the IoT technology LoRa and its interference vulnerability," *International Symposium on Electromagnetic Compatibility*, pp.

533-538, 2019.

- [10] A. Kelly, "NEXT Gen," GTA, 24 August 2023. [Online]. Available: <https://www.kellytelecom.com/what-is-lora-technology/>.
- [11] J. N. A. .. a. K. Alzubi, "Machine Learning from theory to algorithm: an Overview," *Journal of Physics: Conference series*, vol. 1142, p. 012012, 2018.
- [12] V. Nasteski, "An overview of the supervised machine learning methods," *Horizons*, vol. 4, no. 51, p. 56, 2017.
- [13] S. A. A. A. S. a. A. M. M. Naeem, "Unsupervised machine learning algorithms: Comprehensive review," *Journal Computing and Digital Systems*, 2023.
- [14] H. C. R. a. J. Bai, "Evolutionary reinforcement learning: A survey," *Intelligent computing*, vol. 2, p. 0025, 2023.
- [15] X. Z. L. J. a. B. J. O. Anis Yazidi, "The Hierarchical Continuous Pursuit Learning Automation: A Novel Scheme for Environments with Large Numbers of Actions," *IEEE Transactions on neural networks and learning systems*, , vol. 31, pp. 512-526, 2019.
- [16] A. S. A. B. J. Oommen, "The Power of the “Pursuit” Learning Paradigm in the Partitioning of Data," in *International Conference on Engineering Applications of Neural Networks champ*, 2019.
- [17] B. John Oommen, Ole-Christoffer Granmo, Lei Jiao, Xuan Zhang, "Channel Selection in Cognitive Radio Networks: A Switchable Bayesian Learning Automata Approach," *IEEE 24th International Symposium on Personal, Indoor and Mobile Radio Communications: Mobile and Wireless Networks*, pp. 23622367, 2013.
- [18] Pounambal Muthukumar, Venkata Krishna P, Kavitha S, "Learning automatabased multilevel medium access control in cyber-physical system-based smart wireless networks," *International Journal of Communication Systems*, vol. 33, no. 9, 2020.
- [19] Georgios Papadimitriou, Petros Nicopolitidis, Konstantinos Kantelis, Georgia Beletsioti, Anastasios Valkanis, Athanasios Tsakmakis, "Learning-AutomataBased Hybrid Model for Event Detection in LoRaWAN Networks," *IEEE 21st Mediterranean Electrotechnical Conference (MELECON)*, pp. 1183-1188, 2022.

- [20] Mikio Hasegawa, Song-Ju Kim, Ryoma Kitagawa, Ikumi Urabe, Minoru Fujisawa, Aohan Li, "A Lightweight Decentralized Reinforcement Learning Based Channel Selection Approach for High-Density LoRaWAN," *IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, pp. 9 - 14, 2021.
- [21] J. J. S.-H. Y. Y. D. Xu Huang, "A Reinforcement Learning Based Medium Access Control Method for LoRa Networks," Huang, X., Jiang, J., Yang, S. H., & Ding, Y. (2020, October). A reinforcement learning IEEE International Conference on Networking, Sensing, and Control (ICNSC), pp. 1 -6, 2020.
- [22] "An Effective Spectrum Handoff Based on Reinforcement Learning for Target Channel Selection in the Industrial Internet of Things," *Sensors*, vol. 19, no. 6, p. 1395, 2019.
- [23] "Performance Evaluation of Machine Learning Based Channel Selection Algorithm Implemented on IoT Sensor Devices in Coexisting IoT Networks,"
Hasegawa, S., Kim, S. J., Shoji, Y., & Hasegawa, M. (2020, January). Performance evaluation of machine learning base IEEE 17th Annual Consumer Communication, pp. 1 -5, 2020.
- [24] "Simulative Performance Study of Slotted Aloha for LoRaWAN Channel Access," Loh, F., Mehling, N., Geißler, S., & Hoßfeld, T. (2022, April). Simulative Performance Study IEEE/IFIP Network Operations and Management Symposium, pp. 1 - 9, 2022.
- [25] Sebastien Roy, Mohamed Sadik, Rachid Aouami, Essaid Sabir, Imane Cheikh, "Throughput-Delay Tradeoffs for Slotted-Aloha-based LoRaWAN Networks,"
Cheikh, I., Sabir, E., Aouami, R., Sadik, M., & Roy, S. (2021, June). Throughputdelay tradeoffs for S International Wireless Communications and Mobile Computing (IWCMC), pp. 2020 -2025, 2021.
- [26] Steven Martin, Odalric-Ambrym Maillard, Samer Lahoud, Kinda Khawam, Hassan Fawaz, "A channel selection game for multi-operator LoRaWAN deployments," *Computer Networks*, vol. 16, p. 109185, 2022.
- [27] R. F. N. V. P. M. R. L. Hiba Dakdouk, "Massive Multi-Player Multi-Armed Bandits for IoT Networks: An Application on LoRa Networks," *Ad Hoc Networks*, vol. 151, p. 103283, 2023.

- [28] Guihai Chen, Kaishun Wu, Liang He, Kangjie Xu, Linghe Kong, Zhe Wang, "Online Concurrent Transmissions at LoRa Gateway," Wang, Z., Kong, L., Xu, K., He, L., Wu, K., & Chen, G. (2020, July). Onl IEEE INFOCOM 2020-IEEE Conference on Computer Communications, vol. 151, pp. 2331 -2340, 2020.
- [29] P. N. K. K. G. A. B. A. V. Georgios Papadimitriou, "Ensuring Reliability for LoRa Networks using a Reinforcement Learning Assisted Time Division Duplex Protocol," *IEEE Internet of Things Journal.*, 2023.
- [30] J. L. Atsushi Ito, "Adjacent Interference of LoRa for Large-scale Livestock Monitoring," *International Journal of Networking and Computing*, vol. 11, no. 2, pp. 283 -298, 2021.
- [31] Ibrahim Gomaa, Zaki Matar, Ahmed Yahya, Samar Adel Almarzoqi, "ReLearning EXP3 Multi-Armed Bandit Algorithm for Enhancing the Massive IoTLoRaWAN Network Performance," *Sensors*, vol. 22, no. 4, p. 1603, 2022.
- [32] A. K. P. G. A. C. Shakya, "Reinforcement Learning algorithms: A survey," *Expert system with applications*, p. 120495, 2023.
- [33] C. Moy, L. Besson, G. Delbarre and L. Toutain, "Decentralized spectrum learning for radio collision mitigation in ultra-dense IoT networks: LoRaWAN case study and experiments," *Annals of Telecommunications*, no. 75, pp. 711727, 2020.
- [34] Mi Chen; Jalel Ben-Othman; Jean-Michel Fourneau, "Dynamic Parameter Allocation with Reinforcement Learning for LoRaWAN," *IEEE INTERNET OF THINGS JOURNAL*, vol. 10, no. 12, 2023.
- [35] S. Ahsan, S. A. Hassan, A. Adeel, and H. K. Qureshi, "Improving Channel Utilization of LoRaWAN by using Novel Channel Access Mechanism," 15th International Wireless Communications and Mobile Computing Conference, pp. 1656-1661, 2019.
- [36] L. Yurii, L. Anna, and S. Stepan, "Research on the Throughput Capacity of LoRaWAN Communication Channel," *IEEE East-West Design & Test Symposium (EWDTS)*, pp. 1-6, 2023.
- [37] Lilian Besson; Rémi Bonnefoi; Christophe MoyGNU Radio Implementation of MALIN: "Multi-Armed bandits Learning for Internet-of-things Networks" *IEEE Wireless Communications and Networking Conference (WCNC)*pp. 1-62019
- [38] Junaid Qadir; Muhammad Umar Janjua; Muhammad Omer Farooq; Muhammad Usama; Inaam IlahiLoRaDRL: Deep Reinforcement Learning Based Adaptive

- PHY Layer Transmission Parameters Selection for LoRaWAN *IEEE 45th Conference on Local Computer Networks (LCN)* pp. 457-460 2020
- [40] Vasileios Astreriou, Anaastasios Valkanis; Georgia Beletsioti; Konstantinos Kantels, Georgios Papadimitriou; LoRaWAN-Based Adaptive MACs for Event Response applications *ACCESS* 10 pp. 97465-97480 2022
- [41] Teymuri, Benyamin, Reza Serati, Nikolaos Athanasios Anagnostopoulos, and Mehdi Rasti. "LP-MAB: Improving the energy efficiency of LORAWAN using a reinforcement-learning-based adaptive configuration algorithm." *Sensors* , pp 2363 no. 4 ,2023.
- [42] Farzad Azizi; Benyamin Teymuri; Rojin Aslani; Mehdi Rasti; Jesse Tolvanen; , "MIX-MAB: Reinforcement Learning-based Resource Allocation Algorithm for LoRaWAN," *IEEE 95th Vehicular Technology Conference (VTC2022-Spring)*, pp. 1-6, 2022.
- [43] A. R. A. a. B. S. Chaudhari, "Multi-Armed Bandit Algorithm Policy for LoRa Network Performance Enhancement," *Sensors and Actuators Networks*, vol. 12, no. 3, p. 38, 2023.
- [44] R. K. A. L. S. J. K. Y. W. Y. S. a. M. H. So Hasegawa, "Multi-Armed-Bandit Based Channel Selection Algorithm for Massive Heterogeneous of Things Networks," *MDPI*, p. 7424, 2022.
- [45] R. A. R. F. N. V. P. M. Raouf Kerkouche, "Node-based optimization of LoRa Transmissions without Multi-Armed Bandit algorithms," *International Conference on Telecommunications (ICT)*, pp. 521-526, 2018.
- [46] A. Tsakmakis, A. Valkanis, G. A. Beletsioti, P. Nicopolitidis, and G. Papadimitriou, "On the Effect of Traffic Burstiness in LoRaWAN Networks' Performance," *International Conference on Communications, Computing, Cyber and Informatics (CCCI)*, pp. 1-4, 2020.
- [47] J. X. L., J. a. J. Y. Xiaowu Li, "Advancing Performance in LoRaWAN Networks: The circular Region Grouped Bit – Slot LoRa MAC Protocol," *Electronics*, vol. 13, no. 3, p. 621, 2024.
- [48] N. M. A. T. H. Frank Loh, "Towards LoRaWAN without data loss: Stuying the performance of Different Channel Access Approaches," *Electronics*, p. 691, 2022.

- [49] Z. S. a. K. I.-K. W. Ameer Ivoghlian, "Adaptive Wireless Network Management with Multi-Agent Reinforcement Learning," *Sensors*, vol. 22, no. 3, p. 1019, 2022.