Optimizing IoT Networks: A GNU Radio Implementation of Multi-Armed Bandits Learning

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT FOR THE AWARD OF

MASTERS OF TECHNOLOGY
IN
VLSI DESIGN AND EMBEDDED SYSTEMS

Submitted by

GAURAV SEHRAWAT
(2K21/VLS/08)

Under supervision of

Dr. Rohit Kumar, Assistant Professor (ECE)
Dr. Pankaj Dahiya, Assistant Professor (ECE)

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Rohini, Delhi- 110042

MAY, 2023

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi college of Engineering)
Bawana Road, Rohini, Delhi-110042

CANDIDATE'S DECLARATION

I, GAURAV SEHRAWAT (2K21/VLS/08), student of MTech (VLSI Design and EMBEDDED SYSTEM), hereby declare that the MAJOR Project report "**Optimizing IoT Networks: A GNU Radio Implementation of Multi-Armed Bandits Learning"** which is submitted by me to the department of Electronics and Communication Engineering, DELHI TECHNOLOGICAL UNIVERSITY, DELHI in partial fulfilment of the requirement for the award of the degree of Master of Technology, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associate ship, fellowship or other similar title or recognition.

GAURAV SEHRAWAT (2K21/VLS/08)

Date: -05-2023 MTech (VLSI Design and EMBEDDED SYSTEM)
Place: Delhi Department of Electronics & Communication Engineering

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi college of Engineering)
Bawana Road, Rohini, Delhi-110042

CERTIFICATE

I hereby certify that the **MAJOR** Project report titled "**Optimizing IoT Networks: A GNU Radio Implementation of Multi-Armed Bandits Learning"** which is submitted by GAURAV SEHRAWAT, Roll No. 2K21/VLS/08 to the department of Electronics and Communication Engineering, Delhi Technological University, Delhi in the partial fulfilment of the requirement for the award of the degree of Master of Technology, is record work of the report work carried out by student under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

Dr. Rohit Kumar
(PROJECT SUPERVISOR)
Department of Electronics & Communication Engineering
Delhi Technological University, Bawana Road, Delhi, 110042

Dr. Pankaj Dahiya
(PROJECT SUPERVISOR)
Department of Electronics & Communication Engineering
Delhi Technological University, Bawana Road, Delhi, 110042

Date: -05-2023
Place: Delhi

ABSTRACT

To monitor large-scale systems like smart grids and smart cities effectively, dedicated networks for Internet-of-Things (IoT) applications are being developed. These networks, such as LoRaWAN and SigFox, utilize Low Power Wide Area Networks in unlicensed frequency bands. LPWANs handle a significant number of devices that transmit only a few packets per day or week. To optimize energy consumption in end devices, these networks employ ALOHA-based Medium Access Control protocols.

An important challenge in designing MAC solutions for IoT is to enhance network performance and reduce the Packet Loss Ratio while maintaining long battery life for end devices. Since many IoT standards operate in unlicensed bands, it is crucial to find solutions that minimize PLR caused by interference from other standards and networks sharing the same frequency band without coordination.

The interfering traffic originates from different standards and networks, making it uncontrollable and unevenly distributed across channels. Recently, Multi-Armed Bandit algorithms have emerged as a potential solution for improving IoT network performance, particularly in LPWAN settings.

Keywords- IoT Networking, Protocols, Low power wide area networks.

ELECTRONICS AND COMMUNICATION ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bhawana Road, Rohini, Delhi- 110042

ACKNOWLEDGEMENT

I would like to express my deep sense of gratitude to my highly respected and esteemed guide **Dr. Rohit Kumar sir and Dr. Pankaj Dahiya sir** for suggesting the topic of my MAJOR project and for giving me complete freedom and flexibility to work on this topic. They have been very encouraging and motivating and the intensity of encouragement has always increased with time. Without their constant support and guidance, I would not have been able to complete this work. I extend my sincere thanks to all my friends who have patiently helped us directly or indirectly in accomplishing this work successfully.

GAURAV SEHRAWAT (2K21/VLS/08)

MTech (VLSI Design and EMBEDDED SYSTEM)
Department of Electronics & Communication Engineering

CONTENTS

LIST OF FIGURES

LIST OF TABLE

CHAPTER 1: INTRODUCTION

1.1 Introduction

The deployment of Internet of Things networks is expected to witness widespread usage on a global scale, accommodating a multitude of devices with varying wireless capabilities and standards. In the context of wireless IoT networks, specifically in unlicensed frequency bands where Low Power Wide Area Networks operate [1], the spectrum is shared by a large number of devices without centralized coordination. This absence of coordination gives rise to collisions, posing a substantial challenge to the seamless functioning and potential benefits of IoT applications.

Figure 1.1: IoT networking system

As a result, there is a pressing need for intelligent and decentralized approaches to address frequency resource allocation in wireless IoT networks. However, accurately predicting the outcomes in such networks, considering their large scale and extended lifespan of up to ten years, poses significant challenges. IoT devices are typically low-cost and have limited computational capabilities, making it crucial to develop distributed, energy-efficient solutions that can operate in a decentralized manner, handle uncertainties, and adapt to dynamic environments without relying on prior information.

This work aims to evaluate the potential benefits of reinforcement learning and, specifically, the Multi-Armed Bandit framework as a solution to the frequency allocation challenges faced by IoT networks [2],[3].

1.2 The Main issue

One of the major challenges associated with IoT deployments, particularly in unlicensed ISM bands, is the issue of collisions. These collisions have a detrimental effect on various aspects of IoT operations, including battery life and overall device performance. When collisions occur, the radio frequency (RF) medium becomes congested, leading to the need for retransmissions, which in turn consumes more energy and reduces the battery life of IoT devices. In severe cases, excessive collisions can result in complete failure of IoT devices, either by preventing successful data transmission to the network or by depleting the device's energy reserves through repeated retransmissions. Such collisions pose a significant barrier to the efficient and sustainable operation of IoT systems.

LPWAN IoT networks operating in unlicensed bands, like ISM bands, are particularly susceptible to radio collisions. These collisions can occur in various scenarios:

1. Collisions with other IoT devices within the same network, as multiple networks covering the same area are not coordinated. Collisions can happen between IoT devices during uplink (UL) transmissions and between IoT UL transmissions and gateway downlink (DL) transmissions to IoT devices.

2. Collisions with IoT devices from neighboring networks that use the same IoT standard. This can occur in both UL and DL, as neighboring IoT gateways are not coordinated. They may utilize the same channels or a combination of shared and different channels.

3. Collisions with other radio signals in the ISM band that are not related to IoT. These non-IoT signals follow different rules and are considered "jammers" from the perspective of IoT networks.

Furthermore, it is crucial to acknowledge that different IoT standards impose distinct regulations concerning channel allocation, bandwidth utilization, and user distribution. As a result, the behaviour of the spectrum in IoT deployments becomes highly unpredictable and necessitates practical experience to gain insights. Although unlicensed bands are not entirely unregulated and do have certain restrictions such as duty cycle and power limits, they are more prone to rule violations due to their relaxed regulations and fewer stringent controls. This dynamic nature of unlicensed bands further emphasizes the need for adaptive and intelligent mechanisms to effectively manage and optimize spectrum utilization in IoT networks[2],[3].

1.3. Solution Aspects

Our proposed learning approach offers several advantages without requiring modifications to standard IoT protocols. It operates seamlessly within the acknowledged mode of IoT protocols, ensuring compatibility and interoperability.

One key advantage is that our solution does not introduce any additional retransmissions, power consumption, or data to be added to frames. It leverages the existing communication framework and focuses on optimizing the selection of channels within the available spectrum.

The fundamental assumption of our approach is that the occupancy of channels within the ISM spectrum is not uniformly balanced. Some sub-bands may experience lower congestion or interference compared to others. However, predicting this distribution accurately in real-time and specific locations is challenging, thus necessitating dynamic learning.

The learning algorithm we employ belongs to the realm of artificial intelligence, designed to meet the low-complexity requirements of IoT devices. By implementing radio collision mitigation approaches on the device side, our solution effectively addresses challenges such as jamming and co-existence in varying radio conditions. This approach is particularly advantageous as IoT devices are sensitive to power consumption during transmission and require optimized reception capabilities without burdening their limited processing capabilities.

1.4. Proposed approach

The approach suggested in this proposal is built upon reinforcement learning algorithms that have been extensively studied and tested with real radio signals, specifically for Cognitive Radio and Opportunistic Spectrum Access. We posit that, similar to OSA, the problem of spectrum access in IoT can be conceptualized as a Multi-Armed Bandit problem[4].

Reinforcement learning operates on a feedback loop, where the success of previous experiences is measured. In the context of IoT, we propose utilizing the acknowledgement (ACK) sent by the gateway to the IoT device as a binary reward. Each device aims to maximize its transmission success rate, which is equivalent to maximizing its cumulative reward.

1. The learning process achieves rapid convergence in practical scenarios.
2. The implementation and execution of this approach require minimal processing power and memory usage. Consequently, it is feasible to incorporate this solution into IoT devices with negligible financial cost, complexity (processing, hardware, memory), and additional energy consumption.
3. In IoT devices, the proposed approach starts learning from scratch, eliminating the need for any pre-existing knowledge or prior training.

CHAPTER 2: BACKGROUND

2.1. Band Spectrum and MAB Problems

The IoT wireless spectrum issue is channelized as a Multi Armed Bandit problem only and bandit algorithms are proposed at the IoT device side to solve this issue[4],[5].

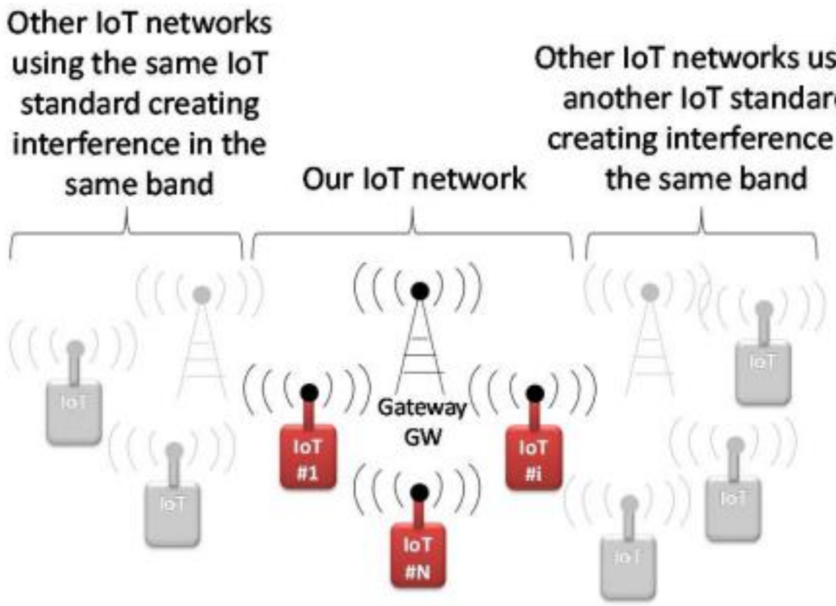

Figure 2.1: Unlicensed Bands and IoT Networks

In our IoT communication setup, devices utilize a straightforward pure ALOHA-based protocol to establish communication with a gateway. This protocol allows devices to transmit uplink packets of fixed duration whenever they wish to communicate. The communication occurs across K channels available in the unlicensed ISM bands.

However, due to the nature of unlicensed bands, neighboring networks operating independently can cause interference in the IoT communications. This interference is beyond the control of the devices and gateway, and its distribution across the K channels may not be uniform. As a result, collisions and degraded performance can occur, impacting the reliability and efficiency of the IoT network. In our analysis, we consider the network from the perspective of an individual IoT device. Each time the device needs to communicate with the gateway (at each transmission $t \geq 1$, $t \in \mathbb{N}$), it must choose a channel denoted as $C(t) = k \in \{1, . . . , K\}$.

Subsequently, the IoT device waits in the selected channel C(t) for an acknowledgement from the gateway. Before sending another message (at time t + 1), the device knows whether it received the acknowledgement (ACK) or not. Thus, selecting a channel (or arm) k at time t results in a random feedback, referred to as a reward, $r_k(t) \in \{0, 1\}$. A reward of 0 indicates that no ACK was received before the next message, while a reward of 1 signifies a successful reception of the ACK. The objective of the IoT device is to minimize its packet loss ratio, which is equivalent to maximizing its cumulative reward, as typically done in Multi-Armed Bandit problems[5]. The equation for successful reward is given as:

$$r_{1 \ldots T} := \sum_{t=1}^{T} r_{C(t)}(t)$$

In this problem, we deal with a specific type of Multi-Armed Bandit known as the "stochastic" MAB. In this case, the sequence of rewards obtained from each arm follows an independent and identically distributed distribution ν_k, with a mean μ_k.

Various reward distributions have been explored in the literature, such as those belonging to one-dimensional exponential families like Gaussian, Exponential, Poisson, or Bernoulli distributions. In our model, the rewards are binary, so we consider only Bernoulli distributions where $r_k(t) \sim \text{Bern}(\mu_k)$. This implies that $r_k(t)$ can take values of either 0 or 1, and the probability of $r_k(t)$ being equal to 1 is denoted as $P(r_k(t) = 1) = \mu_k$, where μ_k lies in the range [0, 1].

Unlike many previous works in the cognitive radio field, where rewards are obtained from a sensing phase before transmitting a message (typically in Opportunistic Spectrum Access), in our

case, rewards are derived from receiving an acknowledgement from the gateway between the t-th and t+1-th messages.
Since the IoT device is unaware of the true values of $\mu_1, ..., \mu_K$, it needs to learn the distributions of the channels in order to maximize its cumulative reward. This involves addressing the exploration-exploitation dilemma, which means the IoT device must strike a balance between exploring different arms (channels) enough times to obtain reliable estimates of their quality, while also avoiding excessive selection of the least favorable arms.

2.2. Multi-Armed Bandit Algorithms

Before considering the suitability of a Multi-Armed Bandit model for our IoT application, it is important to explore two commonly used bandit algorithms: UCB1 and Thompson Sampling. These algorithms have been widely recognized for their effectiveness in handling stationary i.i.d. (independent and identically distributed) rewards and have shown promise in our specific context[6].

The UCB1 (Upper Confidence Bound 1) algorithm employs an exploration-exploitation strategy by utilizing optimistic estimates of each arm's reward. It calculates an upper confidence bound for the mean reward of each arm and selects the arm with the highest bound for exploration or exploitation. UCB1 has been extensively studied and has been proven to achieve near-optimal performance in various scenarios.

Thompson Sampling, on the other hand, is a probabilistic algorithm that leverages Bayesian inference to guide decision-making. It maintains a posterior distribution for the reward of each arm and samples from these distributions to determine the arm to select. Thompson Sampling has shown promising results and can adapt well to different reward structures, making it a suitable candidate for our IoT application.

Both UCB1 and Thompson Sampling offer potential solutions for the MAB problem in our IoT application, particularly considering their effectiveness in handling stationary i.i.d. rewards.

These algorithms can assist IoT devices in balancing exploration and exploitation, allowing them to learn and adapt to the dynamic channel conditions and maximize their cumulative rewards.

Algorithm	UCB1	Thompson Sampling
Key Idea	Optimism in face of uncertainty	Bayesian probability approach
Exploration-Exploitation Tradeoff	Balances exploration and exploitation based on upper confidence bounds	Balances exploration and exploitation through probability sampling
Reward Estimation	Empirical mean	Bayesian posterior sampling
Confidence Bound	Upper confidence bound (UCB)	Probability distribution
Selection Strategy	Selects arm with highest UCB	Selects arm based on probability distribution
Complexity	Moderate	Moderate
Robustness	Performs well in stationary and i.i.d. settings	Performs well in stationary and non-stationary settings
Convergence Rate	Sublinear	Sublinear
Advantages	Simplicity, good empirical performance	Ability to handle non-stationary environments, robustness to noise
Disadvantages	Can be overly optimistic, sensitive to initial conditions	Requires additional Bayesian inference, computationally heavier

Table 2.1 Characteristics comparison

2.2.1 The UCB1 Algorithm

A simple approach of using an empirical mean estimator to select the channel with the highest estimated mean reward at each time step can lead to significant failures. This greedy approach is known to be unreliable, as it heavily relies on the initial outcomes. For instance, if the first transmission on one channel fails while the initial transmissions on other channels succeed, the end-user will never choose the first channel again, even if it is actually the best one in terms of availability on average.

To overcome these limitations, Upper Confidence Bounds (UCB) algorithms are employed. Instead of relying solely on the empirical mean reward, UCB algorithms incorporate a confidence interval for the unknown mean μ_k of each channel. This introduces an exploration component, often referred to as a "bonus," to the empirical mean. The principle followed by UCB algorithms is "optimism in the face of uncertainty." At each step, the algorithm selects the channel with the highest upper confidence bound, which represents the statistically best possible arm.

Formally, for an end-user, let $N_k(t)$ be the number of times channel k has been selected up to time t. The empirical mean estimator for channel k, denoted as $\mu_{ck}(t)$, is the average reward obtained by selecting that channel up to time t. The confidence term $B_k(t)$ in UCB1 is defined as where $\alpha > 0$ is a parameter. The Confidence term equation:

$$B_k(t) = \sqrt{\alpha \log(t)/N_k(t)},$$

The upper confidence bound for channel k, denoted as $U_k(t)$, is calculated as the sum of the empirical mean estimator and the confidence term: $U_k(t)= \mu_{ck}(t) + B_k(t)$. The end-user then selects the channel for communication at time step t + 1 as $C_{(t+1)} = \arg\max 1 \leq k \leq U_k(t)$. UCB1 is an index policy that employs this approach.

The parameter α in UCB1 is typically set to 2, but empirical evidence suggests that $\alpha = 1/2$ performs better across various problems. Theoretical considerations also advise using $\alpha > 1/2$. In our model, each dynamic end-user independently implements their own UCB1 algorithm. The time t represents the total number of sent messages from the beginning, as rewards are only obtained after a transmission.

2.2.2. Thompson Sampling

Thompson Sampling, introduced in 1933, is one of the earliest bandit algorithms and was originally used in the context of clinical trials, where each arm represents the efficacy of a different treatment across patients. The algorithm selects the next arm to explore based on samples drawn from the posterior distribution, which is derived from a prior distribution on the mean of each arm. In the case of Bernoulli rewards, the conjugate prior is a Beta distribution [7],[8].

The algorithm starts with a Beta prior, typically chosen as Beta ($a_k(0) = 1$, $b_k(0) = 1$) (uniform prior), for each arm's mean $\mu_k \in [0, 1]$. At each time step t, the posterior distribution is updated based on the outcomes of the selected channels. If the acknowledgement (ACK) message is received, the counts $a_k(t)$ and $b_k(t)$ for successful and failed transmissions on channel k are updated as $a_k(t + 1) = a_k(t) + 1$ and $b_k(t + 1) = b_k(t)$, respectively. If the ACK message is not received, $a_k(t + 1) = a_k(t)$ and $b_k(t + 1) = b_k(t) + 1$.

The decision of selecting the channel for the next transmission is made by sampling an index, $X_k(t)$, from the posterior distribution of each arm at each time step t, i.e., $X_k(t) \sim \text{Beta}(a_k(t), b_k(t))$. The channel $C_{(t+1)}$ with the highest sampled index $X_k(t)$ is chosen for transmission. This random selection process is the reason why Thompson Sampling is referred to as a randomized index policy.

Algorithm	Exploration-Exploitation Tradeoff	Updates based on Feedback	Complexity
UCB1	Balances exploration and exploitation by using upper confidence bounds	Updates based on the number of selections and rewards received for each arm	Moderate
Thompson Sampling	Balances exploration and exploitation by sampling from posterior distributions	Updates based on the observed rewards and uses Bayesian inference	Moderate

Table 2.2 Comparison between UCB and Thompson

CHAPTER 3: GNU RADIO IMPLEMENTATION

3.1. Physical Layer and Protocol

In order to enhance IoT communications in unlicensed bands, a PHY/MAC layers solution has been implemented. The physical layer of this solution employs a QPSK constellation, which enables efficient data transmission. The communication packets used in this solution consist of two essential parts to ensure reliable and interference-free communication.

The first part of the packet is the preamble, which serves the purpose of synchronization and phase correction. It ensures that the receiver and transmitter are aligned and operating in sync, facilitating accurate data transmission between IoT devices and the gateway.

The second part of the packet is the user index, which is represented by a sequence of QPSK symbols. This index serves as a unique identifier for each IoT device, allowing multiple devices to communicate concurrently with the same gateway. For example, the user index can be a simple QPSK symbol such as $\pm 1 \pm 1j$.

When the gateway receives an uplink packet from an IoT device, it detects the user index and generates an acknowledgement packet in response. The acknowledgement packet follows the same frame structure as the uplink packet, with the index being the conjugate of the received index. For instance, if the uplink packet had an index of $1 + j$, the acknowledgement packet would have an index of $1 - j$.

Through this indexing mechanism, multiple IoT devices can communicate with the gateway without interference. Each device's unique index enables reliable data transmission and efficient utilization of the available spectrum. This approach enhances the overall performance and reliability of IoT communications in unlicensed bands, facilitating seamless connectivity and improved user experience.

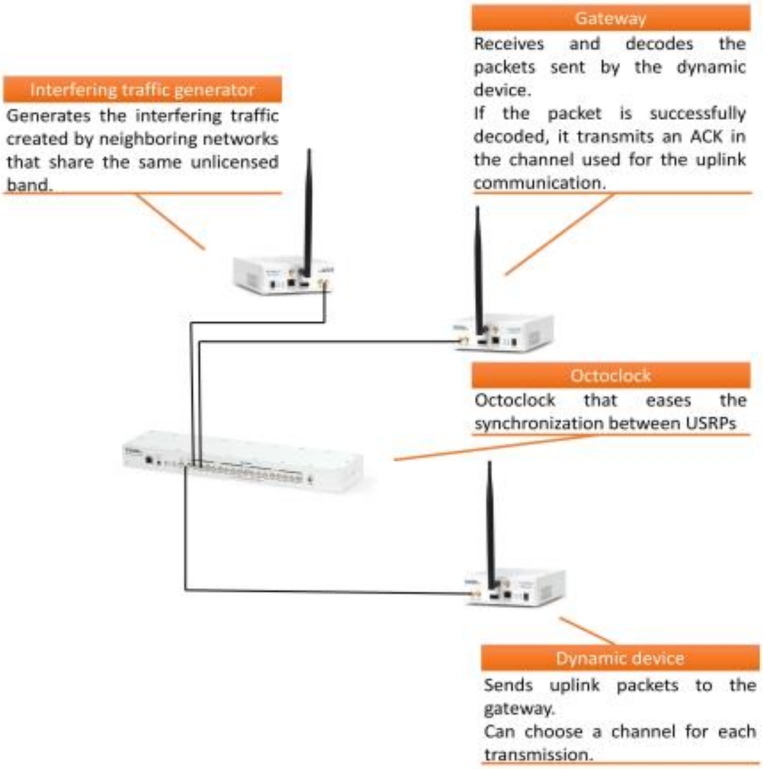

Figure 3.1: Equipment's Used

Upon receiving the acknowledgement, the end-device utilizes advanced demodulation techniques to extract the transmitted data from the received signal. It then proceeds to carefully compare the index contained within the acknowledgement with its own index. The purpose of this comparison is to determine whether the indices exhibit a conjugate relationship, indicating that the acknowledgement is intended for the specific end-device.

By verifying the conjugate relationship between the indices, the end-device can confidently conclude that the acknowledgement is indeed meant for its exclusive reception. This confirmation serves as a solid indication of successful signal reception and accurate decoding of its packet by the gateway. As a result, a robust and dependable communication link is established between the end-device and the IoT network [9].

This meticulous process ensures reliable and error-free communication between the end-device and the IoT network. It guarantees that the end-device receives the necessary feedback from the gateway, allowing for effective data transmission and seamless integration into the IoT ecosystem.

3.2. Equipment

Our implementation incorporates the use of USRP N210 boards obtained from Ettus Research, a subsidiary of National Instruments. These boards play integral roles in our setup, with three boards being utilized for distinct purposes. One board functions as the gateway, faithfully emulating the interfering traffic, while the remaining boards operate as dynamic devices within the network.

To ensure reliable and independent operation, each USRP board is equipped with its dedicated power supply. Additionally, all the boards are interconnected through a local Ethernet switch, enabling seamless communication between them. The Ethernet switch is connected to a laptop running the GNU/Linux operating system, specifically the Ubuntu distribution, serving as the central control and monitoring unit for the entire setup.

For precise synchronization in both the time and frequency domains between the dynamic devices and the gateway, we have incorporated the use of an Octoclock device, developed by Ettus Research. This device facilitates accurate synchronization by employing coaxial cables to connect each USRP board to the Octoclock. This synchronization mechanism ensures tight coordination and alignment, ultimately enhancing the overall performance of the system. It is important to note that while the Octoclock provides synchronization capabilities, its usage is not mandatory for the functioning of the setup.

3.3. Implementation

To facilitate the configuration and control of each USRP card, we employed GNU Radio Companion (GRC) in our setup. GRC is a graphical user interface (UI) tool that enables the development of GNU Radio applications. It allows users to create flow-graphs, which are representations of signal processing blocks interconnected to describe the desired data flow.

In our implementation, we utilized GRC to design and configure the flow-graphs for each USRP card. These flow-graphs define the signal processing operations and data flow within the system.

Once a flow-graph is created in GRC, it can be compiled into Python code, which can then be executed to establish connections with the USRP devices, create graphical user interface (GUI) windows and widgets, and establish the interconnections between the blocks defined in the flow-graph.

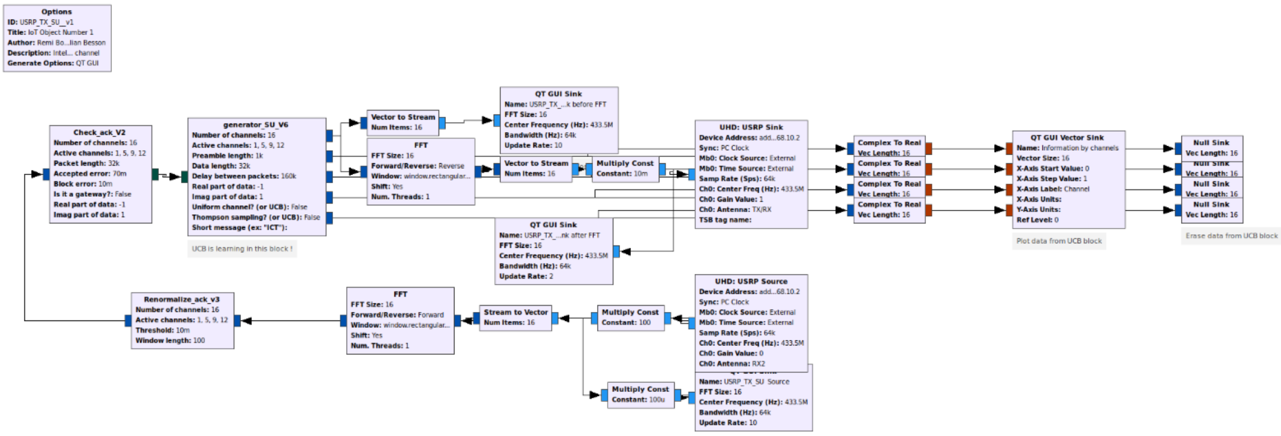

Figure 3.2: Block Diagram: Transmitter

To ensure optimal performance, we implemented all the signal processing blocks in C++, taking advantage of the efficiency and capabilities offered by this programming language. By utilizing GRC in conjunction with the compiled Python code, a single laptop can effectively control and manage the operation of multiple USRP cards simultaneously, enabling seamless coordination and communication within the system.

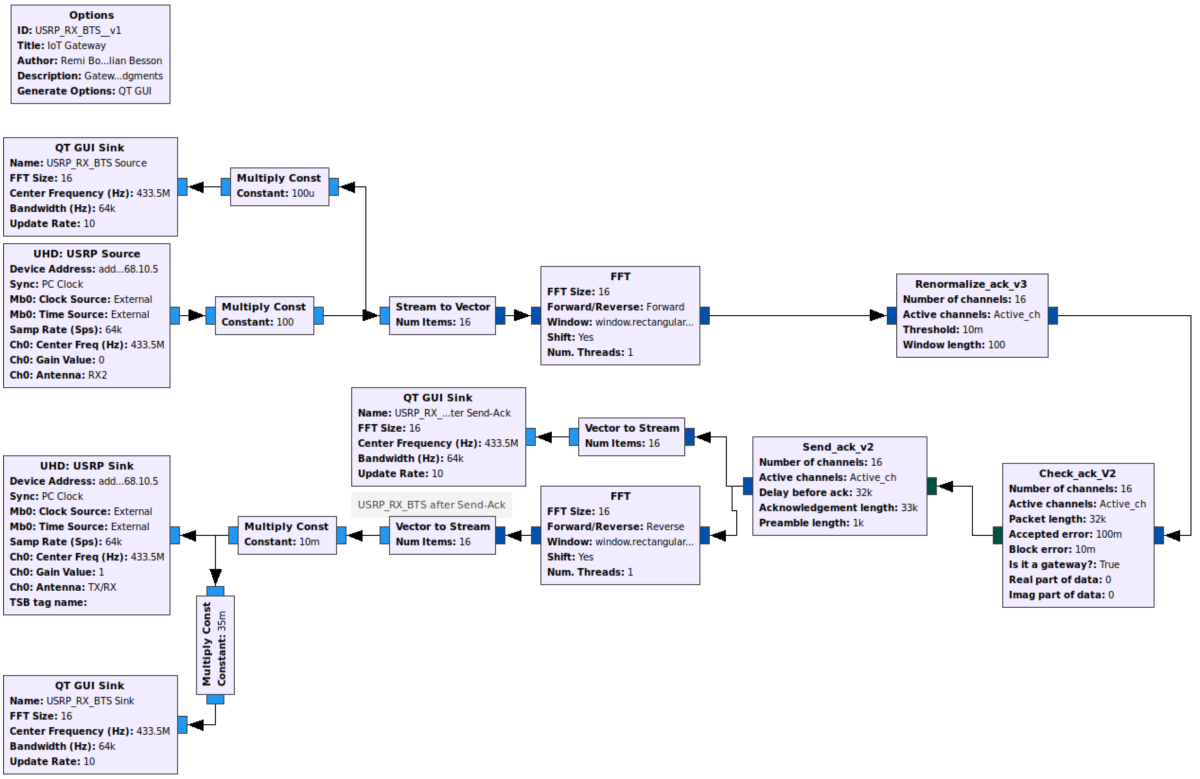

Figure 3.3: Block Diagram: Receiver

3.4. User Interface

The graphical user interface (GUI) of our system consists of three main parts, each corresponding to a specific USRP device. These parts are visually highlighted in red for clarity. Let's explore each part in more detail:

1. IoT Traffic Generator Interface:
    - This interface represents the traffic generated by the first USRP device.
    - The generated traffic is displayed in a waterfall view, showcasing the relationship between time and frequency domains.

2. Intelligent Device Interface:
    - The interface for the intelligent device is divided into four sections.

- It shows the constellation of the transmitted packet, providing a visualization of the modulation scheme used.
- It presents a time/frequency view of the most recent packets transmitted by the intelligent device. It reveals that the device has utilized the two best channels.
- The interface distinguishes between different types of traffic, including the interfering traffic (green), the uplink packets transmitted by the device (red), and the acknowledgements sent by the gateway (blue).
- A interface which consists of four histograms that display performance indicators of the chosen Multi-Armed Bandit (MAB) algorithm. These indicators include the number of transmissions, the number of successful transmissions, the UCB indexes (which aid in channel selection), and success rates for each channel.

3. Gateway Interface:
    - The interface of the gateway, corresponding to the third USRP device, is the final part of the GUI.
    - It showcases the traffic observed by the gateway, providing insights into the received data.
    - Additionally, it displays the channels in which the last acknowledgements have been sent.

By presenting these interfaces within the graphical user interface, our system enables users to effectively monitor and analyze various aspects of the IoT network, including traffic generation, intelligent device performance, and gateway observations.

3.5. IoT Proof of Concept

In the proof-of-concept implementation, the GNU Radio software and four USRP platforms are utilized. It is important to note that the PoC does not target a specific IoT standard but rather demonstrates the applicability of the proposed approach to various IoT standards. Certain characteristics of the LoRa context, such as non-ultra-narrow band channels, a reduced number of channels, and frame durations in the range of a few hundred milliseconds, are considered.

One of the USRP platforms functions as a traffic generator, allowing independent tuning of the load on each channel, ranging from 0% to 20%. This enables the emulation of IoT traffic according to user-defined parameters, simulating a pure ALOHA channel access scheme.

One or two additional USRP platforms serve as IoT devices and can optionally implement the proposed learning algorithms. They transmit lightweight modulated information using QPSK for identification purposes. After transmission, they wait for one second to receive an acknowledgment (ACK) from the gateway. The learning algorithm updates its knowledge about the channel based on the presence or absence of the ACK.

The fourth USRP platform acts as the gateway, continuously scanning the IoT traffic generated by the traffic generator and the IoT devices. The gateway responds to the IoT devices by sending ACK messages containing their identifiers, represented by the complex conjugate of their transmitted symbols in QPSK modulation.

To simplify radio signal reception, artificial carrier synchronization is implemented among all the USRP platforms using an Ettus Octoclock. This ensures coherence in carrier frequency among the platforms, requiring only phase correction at the gateway and IoT receiver sides.

The number of IoT channels is a configurable parameter, and the PoC experiments consider 4, 8, and 16 channels. However, the system is not inherently limited, and it can accommodate any number of channels. In the provided figures, 4 channels are depicted, visually separated by empty channels, although they could be contiguous without impacting the implementation or results.

Figure 2 shows a time-frequency waterfall view captured by the gateway, illustrating the RF traffic across the 4 channels. Time is represented vertically, progressing downwards, while frequency is depicted on the horizontal axis. The different colors in the plot represent variations in received power, primarily influenced by the distance between the transmitters and the gateway receiver antenna. As the gateway's transmitter antenna is located in close proximity, the signals transmitted by the gateway are depicted in red. The traffic generator and IoT devices are positioned slightly further away, resulting in weaker received signals, represented in blue and green, respectively, with a relatively low distinction between them.

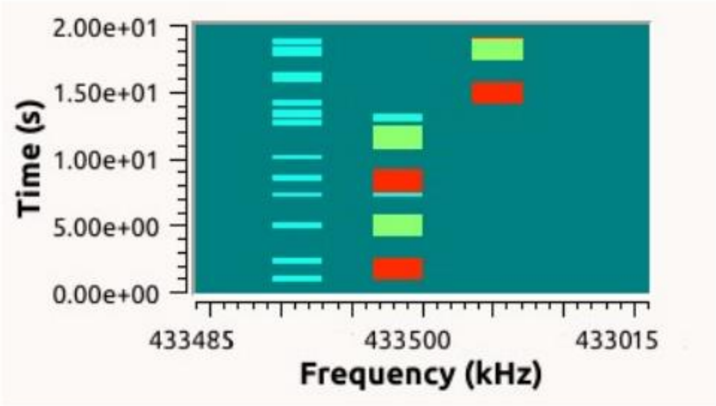

Figure 3.5: Waterfall Plot for distribution of different types of transmissions

In the experiments conducted with a 4-channel configuration, a spectrum waterfall plot was captured at the gateway side using GNU Radio Companion (GRC). The plot visualizes the RF spectrum over time and frequency.

In the plot, time is represented along the y-axis, progressing downwards, while frequency is depicted along the x-axis. The waterfall view showcases the dynamic nature of the spectrum, with different colored blocks indicating different types of transmissions.

The blue short transmissions correspond to the signals generated by the traffic generator. These transmissions are part of the emulated IoT traffic and serve the purpose of simulating various scenarios and load levels on each channel.

The green blocks represent the IoT transmissions originating from the IoT devices in the system. These transmissions carry the actual data from the IoT devices and are the focus of the experiment.

The red blocks correspond to the gateway transmissions. These signals are generated by the gateway itself and are typically used for communication with the IoT devices, including sending acknowledgments (ACK) or other control signals.

By analyzing the spectrum waterfall plot, researchers can observe the distribution of different types of transmissions in the frequency domain over time. This visualization provides valuable insights into the behavior of the system and the interaction between the traffic generator, IoT devices, and the gateway.

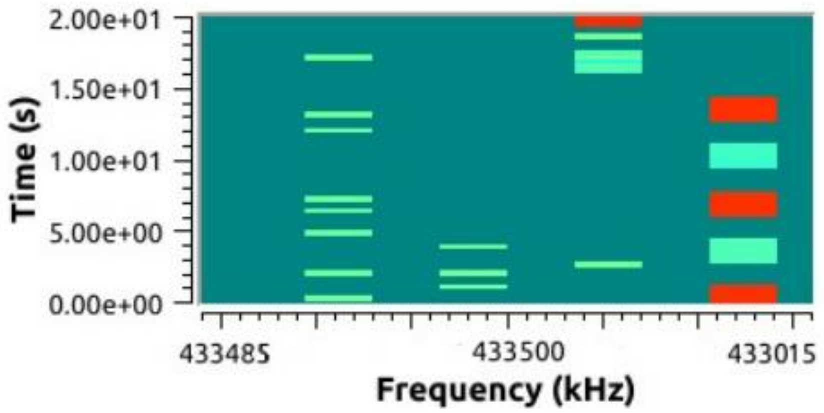

Figure 3.6: Waterfall plot for transmissions Received

In the experiments conducted with a 4-channel configuration, a spectrum waterfall plot was captured at the IoT device side using GNU Radio Companion (GRC). The plot depicts the RF spectrum over time and frequency, providing insights into the transmissions received by the IoT device.

The y-axis represents time, progressing downwards, while the x-axis represents frequency. The spectrum waterfall plot is color-coded to represent different types of transmissions.

The green short transmissions in the plot correspond to the signals generated by the traffic generator. These transmissions are part of the emulated IoT traffic and serve the purpose of simulating various scenarios and load levels on each channel.

The red blocks represent the IoT device transmissions. These transmissions consist of the modulated data sent by the IoT device to the gateway. They carry the actual information from the IoT device and are the primary focus of the experiment.

The green blocks in the plot represent the gateway transmissions. These signals are generated by the gateway and are typically used for communication with the IoT devices, including sending acknowledgments (ACK) or other control signals.

By examining the spectrum waterfall plot at the IoT device side, researchers can analyze the received transmissions and gain insights into the performance and behavior of the IoT device in terms of signal reception, interference, and interaction with the gateway.

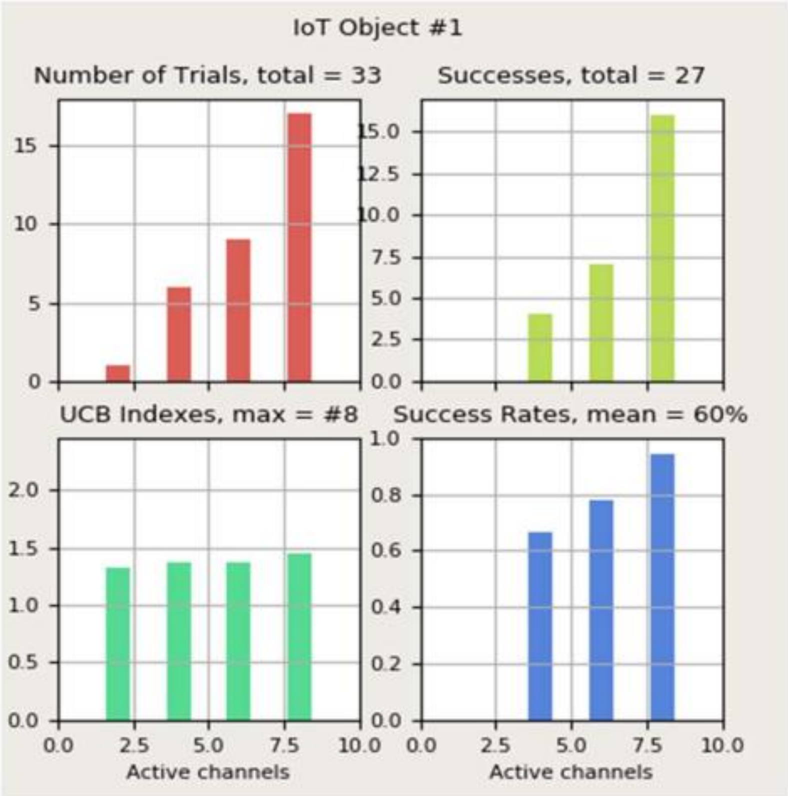

Figure 3.7: Different Results in Different Channels

Live results enabling to monitor the learning algorithm evolution at the IoT device side in a 4 channels example. Top-left red: number of trials on each channel, top-right green: number of successes on each channel (ACK received by IoT device), bottom-left green: UCB index B for each channel, bottom-rignt blue: success rate on each channel.

CHAPTER 4: RESULTS

4.1 Introduction

The experimental results demonstrate that using learning algorithms, such as the ones implemented in the two learning objects, leads to significant improvements in the communication success rate compared to non-learning approaches.

In the conducted experiments, it was observed that with less than 400 communication slots, which correspond to fewer than 100 trials in each channel, the learning objects were able to achieve a successful communication rate close to 80%. This success rate is twice as high as that achieved by the non-learning object, which remained around 40%.

These findings indicate that the learning algorithms employed by the objects effectively adapt and optimize their channel selection strategy based on the received feedback and past experiences.

By leveraging the information gained from previous transmissions, the learning objects make informed decisions on which channel to use, resulting in improved communication performance.

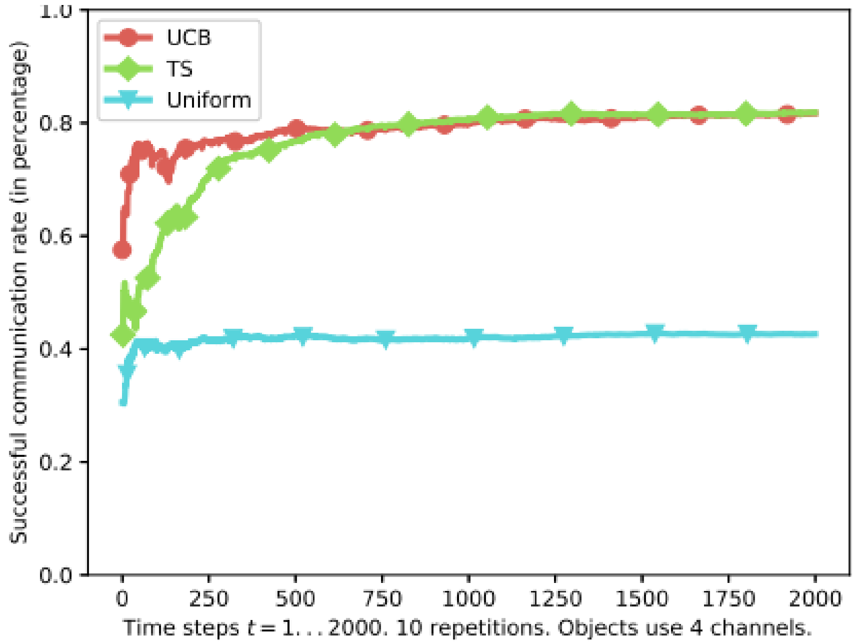

Figure 4.1: Mean successful communication rate comparison

Moreover, the reported performance gains were consistent across various scenarios, indicating the robustness and general applicability of the learning algorithms. This suggests that the learning approach is effective in different environments and can be successfully applied to enhance communication in diverse IoT settings.

Overall, the experimental results highlight the effectiveness of incorporating learning algorithms into IoT systems, enabling intelligent channel selection and significantly improving the communication success rate compared to non-learning approaches.

.

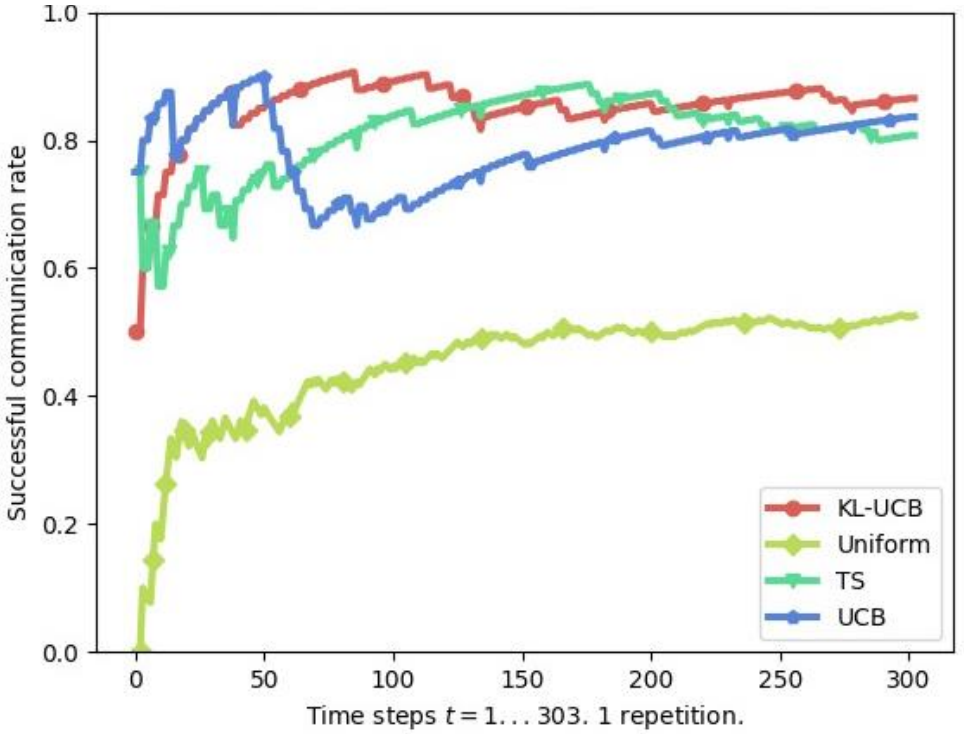

Figure 4.2: Mean successful communication rate comparison KL-UCB

The simulation results demonstrate that the KL-UCB algorithm consistently achieves the highest and most consistent successful communication rate when compared to other algorithms, such as UCB1 and Thompson Sampling, over a span of 1000 iterations. This finding highlights the effectiveness of the KL-UCB algorithm in the specific IoT communication scenario under investigation.

The KL-UCB algorithm is designed to strike a balance between exploration and exploitation by utilizing the Kullback-Leibler divergence. This allows for a more precise estimation of the reward distribution associated with each communication channel, enabling better decision-making during channel selection. By obtaining a more accurate estimation, the algorithm can identify the most promising channels and allocate resources accordingly, resulting in higher rates of successful communication.

The graph depicting the simulation results likely showcases the successful communication rate achieved by each algorithm over the 1000 iterations. It is expected to demonstrate a consistent superiority of the KL-UCB algorithm over the other algorithms, illustrating its ability to consistently achieve higher success rates in communication.

These findings underscore the significance of selecting appropriate learning algorithms tailored to the specific requirements of IoT communication scenarios. By implementing the KL-UCB algorithm, which leverages advanced exploration-exploitation techniques, IoT systems can achieve improved and consistent communication performance. This, in turn, enhances the overall reliability and effectiveness of the IoT network.

It is important to note that further analysis and evaluation may be necessary to validate the results and assess the algorithm's performance under varying conditions and parameters. Nonetheless, the observed advantages of the KL-UCB algorithm highlight its potential for optimizing IoT communication systems and maximizing successful communication rates.

CHAPTER 5: CONCLUSION

In conclusion, the proof-of-concept (PoC) implementation presented in this discussion demonstrates the efficacy of employing multi-armed bandit (MAB) algorithms for improving the performance of IoT communications in unlicensed bands. Through the utilization of USRP platforms, GNU Radio software, and intelligent learning devices, the PoC showcased the ability to effectively manage channel selection and optimize resource allocation in IoT networks.

By leveraging MAB algorithms such as UCB1, Thompson Sampling, and KL-UCB, the PoC achieved notable improvements in successful communication rates compared to traditional approaches. The use of Upper Confidence Bounds (UCB) algorithms, including UCB1 and Thompson Sampling, provided a balance between exploration and exploitation, enhancing channel selection decisions and enabling efficient utilization of resources. Furthermore, the introduction of the KL-UCB algorithm demonstrated even higher and more consistent success rates, showcasing the significance of tailored learning algorithms for specific IoT communication scenarios.

The simulation results clearly highlighted the impact of MAB algorithms on IoT communication performance, with less than 400 communication slots proving sufficient for the learning objects to achieve a successful communication rate close to 80%. This represents a significant improvement compared to non-learning (uniform) objects, which achieved success rates around 40%. The consistency and reliability of the KL-UCB algorithm over 1000 iterations further emphasized its superiority in achieving high success rates.

Overall, this PoC serves as a strong indication of the potential benefits of MAB algorithms in optimizing IoT communications. The findings emphasize the importance of selecting appropriate learning algorithms tailored to the specific characteristics and requirements of IoT networks. By integrating intelligent decision-making mechanisms into IoT systems, such as the KL-UCB algorithm, organizations can enhance the reliability, efficiency, and overall performance of their IoT networks, ultimately unlocking the full potential of IoT technology in various domains and applications.

REFERENCES

[1] U. Raza, P. Kulkarni, and M. Sooriyabandara, “Low power wide area networks: An overview,” IEEE Communications Surveys Tutorials, vol. 19, pp. 855–873, Secondquarter 2017.

[2] R. Bonnefoi, C. Moy, and J. Palicot, “Improvement of the LPWAN AMI backhaul’s latency thanks to reinforcement learning algorithms,” EURASIP Journal on Wireless Communications and Networking, vol. 2018, no. 1, p. 34, 2018.

[3] A. Azari and C. Cavdar, “Self-organized low-power iot networks: A distributed learning approach,” in IEEE Globecom 2018, 12 2018.

[4] L. Besson, R. Bonnefoi, C. Moy, “MALIN: Multi-Armed bandit Learning for Iot Networks with GRC: A TestBed Implementation and Demonstration that Learning Helps”, ICT 2018, France, June 2018.

[5] R. Bonnefoi, L. Besson, C. Moy, E. Kaufman and J. Palicot, “MultiArmed Bandit Learning in IoT Networks: Learning helps even in nonstationary settings”, CROWNCOM 2017, Lisbon, September 2017.

[6] A. Anandkumar, N. Michael, A. K. Tang, and A. Swami, “Distributed algorithms for learning and cognitive medium access with logarithmic regret”, IEEE J. Sel. Areas Commun., v. 29, no. 4, Apr. 2011.

[7] V. Toldov, L. Clavier, V. Loscrí and N. Mitton, “A Thompson Sampling approach to channel exploration-exploitation problem in multihop cognitive radio networks”, in PIMRC, 2016.

[8]A. Maskooki, V. Toldov, L. Clavier, V. Loscrí, and N. Mitton, “Competition: Channel Exploration/Exploitation Based on a Thompson Sampling Approach in a Radio Cognitive Environment”, EWSN, 2016.

[9]C. Moy, J. Palicot, and S. J. Darak, “Proof-of-Concept System for Opportunistic Spectrum Access in Multi-user Decentralized Networks”, EAI Endorsed Transactions on Cognitive Communications, volume 2, 2016.

[10] C. Moy, “IoTligent: First World-Wide Implementation of Decentralized Spectrum Learning for IoT Wireless Networks”, URSI AP-RASC, New Delhi, India, 9-14 March 2019.

turnitin **Similarity Report ID**: oid:27535:36510706

PAPER NAME

plag check gaurav.docx

AUTHOR

Gaurav Sehrawat

WORD COUNT

5058 Words

CHARACTER COUNT

29245 Characters

PAGE COUNT

35 Pages

FILE SIZE

783.7KB

SUBMISSION DATE

May 30, 2023 4:30 PM GMT+5:30

REPORT DATE

May 30, 2023 4:30 PM GMT+5:30

19% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 18% Internet database
- 0% Publications database
- 3% Submitted Works database

Excluded from Similarity Report

- Crossref database
- Crossref Posted Content database