Decentralized Adaptive Spectrum Learning in Wireless IoT Networks based on Channel Quality Information

Ahmed Abdelghany, B. Uguen, C. Moy, Dominique Lemur

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2022.3167016, IEEE Internet of Things Journal

Decentralized Adaptive Spectrum Learning in Wireless IoT Networks based on Channel Quality Information

Ahmed Abdelghany, Bernard Uguen, *Member, IEEE*, Christophe Moy, *Senior Member, IEEE*, Dominique Lemur

***Abstract*—In the Internet of Things (IoT) context such as Low Power Wide Area Network (LPWAN), it is essential to reduce the packet losses, e.g., to save energy. Decentralized artificial intelligence techniques have been proposed to combat radio collisions, but the approach here is extended to deal additionally with the channel propagation effects. In this article, a Quality of Channel Allocation (QoC-A) learning technique based on bandit algorithms is proposed in order to choose the transmission channel. This aims to reduce the effect of the propagation impairments between the radio channels while using the Effective Signal Power (ESP) as a quality metric. In addition, a Discounted QoC-A (DQoC-A) algorithm is proposed to adapt rapidly to any abrupt change in the channels' conditions. An experimental campaign on a real IoT device is carried out to demonstrate the low complexity and efficiency of these proposed decentralized algorithms. In the given results, QoC-A outperforms the classical UCB policy with a more accelerated learning process. On the other hand, the feasibility of using the DQoC-A in non-stationary scenarios is illustrated by its rapid convergence when abrupt changes in the channels' conditions occur. At the end of the process, these proposed learning techniques give 4.1 and 2.4 times fewer packet losses than the traditional ones with a random channel assignment scheme, in the stationary and non-stationary scenarios, respectively.**

***Index Terms*—IoT, LPWAN, LoRa, Packet Loss Rate, RSSI, Effective Signal Power, Measurement, Smart City, Spectrum Allocation, Machine Learning, Reinforcement Learning, UCB.**

I. Introduction

A. Background

INTERNET of Things (IoT) has been utilized in a wide range of domains, such as smart city, health monitoring, livestock, surveillance, etc [1], [2], [3]. For some IoT applications, the characteristics of large coverage, low power consumption and a low-cost deployment network are required, which could be fulfilled using Low Power Wide Area Network (LPWAN). For providing these capabilities, LoRaWAN (Long Range Wide Area Network) is considered as the leading technology that is based on a Chirp Spread-Spectrum (CSS) solution and operating in unlicensed bands, i.e. $868\,\mathrm{MHz}$ in Europe and $915\,\mathrm{MHz}$ in the USA. Hence, it can interconnect low-cost and mostly battery-powered devices over long ranges to the gateway [4].

The authors are with *Univ Rennes, CNRS, IETR - UMR 6164*, F-35000, Rennes, France (e-mail: {ahmed.abdelghany, bernard.uguen, christophe.moy, dominique.lemur}@univ-rennes1.fr).

As LoRaWAN protocol has no exclusive rights in these unlicensed bands [5], also called Industrial, Scientific and Medical (ISM) application bands, channel impairments may occur on the link between the end node and the gateway, reducing the reliability of communications in these networks [6]. Consequently, packet loss will be the weak point of these networks as it degrades the performance of the entire network in the long term. These packet losses may occur either due to collisions as depicted in [7], or propagation effects as manifested in [8], [9] and [10]. Serious outcomes are caused by this transmission failure that may affect various IoT applications, in particular those using the acknowledged messaging mode for the important sensor data. If the end node does not receive the Acknowledgment (ACK) packet, the end node will retransmit the data packet. However, extra energy consumption is required by this retransmission that impacts the battery life of the end node devices plus occupying an additional spectrum and raising the interfering level.

B. Related Works and Motivation

Recently, many works are proposing different solutions for reducing the packet loss rate. For example, in [11] the authors propose DyLoRa, a dynamic LoRa transmission control system to improve energy efficiency, as an alternative to the state-of-the-art LoRaWAN Adaptive Data Rate (ADR). Thus, the idea of DyLoRa is to adjust the transmission parameters, i.e. transmission power and Spreading Factor (SF), among different environments. Moreover, a Collision Avoidance Resource Allocation (CARA) algorithm is proposed in [12] with the objective of increasing the system capacity by mitigating packet losses from the collisions. However, these previous works are either centralized approaches (on the network side), hence, they do not consider the different interference conditions in the proximity of each end node which is located in a specific area. Or, they are imposing changes on the IoT protocol with extra packet transmissions and time synchronization between the end nodes, as depicted in Table I.

While in [13] and [14], the first implementation of a decentralized spectrum learning for IoT wireless networks is proposed in order to mitigate radio collisions without changing anything to LoRaWAN protocol. Moreover, they can be used in addition to conventional ADR techniques. They propose to use a learning algorithm on LoRa devices that can cope

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2022.3167016, IEEE Internet of Things Journal

with spectrum scarcity which could occur in unlicensed bands. Nevertheless, this learning algorithm does not exploit the channel quality at each frequency band on the end node side to accelerate the learning process. This decentralized approach could be extended, which was just to avoid collisions to the management of channel quality too. Hence, the channel quality is estimated from the channel parameters, i.e. SNR (Signal-to-Noise Ratio), Received Signal Strength Indicator (RSSI) and Effective Signal Power (ESP), which are considered as one of the main factors affecting the packet loss rate [8], [9]. Moreover, these channel parameters could be obtained on the end node side for implementing a decentralized approach, thanks to the reciprocity between the uplink and downlink Channel State Information (CSI).

C. Contributions

The main contributions of this article are threefold as follows:

- A Quality of Channel Allocation (QoC-A) algorithm is proposed by extending the classical Upper Confidence Bounds (UCB) technique to deal with the channel propagation effects besides collisions. This proposed decentralized machine learning algorithm takes into account the quality of the channel in addition to the traditional binary reward. Subsequently, it reduces the packet losses by learning a proper frequency allocation scheme that averts using a channel with poor propagation conditions and then accelerates the learning process.

- ESP is utilized instead of RSSI as the channel quality indicator, to overcome the RSSI limitation as discussed in [8].

- A Discounted QoC-A (DQoC-A) algorithm is proposed in order to tackle situations where the channels' conditions are non-stationary on a long-term basis. Here, this adaptive algorithm fulfills the necessity of introducing a specific mechanism for forgetting the outdated channels' states. Based on that, the CSI shape evolution which happens when an end node moves from one place to another is typically the kind of non-stationary behavior that is addressed by this algorithm. Another non-stationary scenario could occur when the end node exhibits a changing environment model, for example, if it is placed in a road where there are moving cars and objects. These moving obstacles may introduce a temporal evolution of the CSI shape between the morning and night period [15].

On the other hand, the proposed reinforcement learning algorithms could be embedded at the IoT end node side with low hardware extra cost in terms of processing power, memory footprint, etc. Hence, in a real network deployment within the university campus in Rennes, these algorithms are implemented inside an end node device and executed in different scenarios, i.e. stationary and non-stationary. Consequently, the whole results from the experiment are extensively evaluated among the proposed reinforcement learning algorithms against the UCB policy as well as the traditional scheme with random frequency allocation. Moreover, important remarks are given for adjusting the algorithms' configurations based on the properties of each of the potential IoT applications.

D. Organization

The remainder of this article is organized as follows. Section II presents the main factors of the packet losses. Section III provides sufficient detail about the system model of the proposed algorithm. The proposed algorithms are then illustrated in Section IV. While Section V shows the experimental architecture and network configuration used for the measurement campaign. In Section VI, the experimental results of the proposed algorithms are tested through the different scenarios. Additional remarks are given for adjusting the algorithms' configurations in Section VII. Finally, the work is concluded in Section VIII. A list of key acronyms used throughout this article is presented in Table II.

II. Key Factors of Packet Loss

Packet losses are the main drawback for IoT [16]. Hence, packet losses cause many retransmissions at the cost of a lower battery lifetime of the end nodes and may lead to an increase of the RF (Radio Frequency) contention level. Even worse, a total failure of the IoT service could happen, either because end nodes can not succeed in sending any data to the gateway or because all their energy is consumed much faster than expected due to the multiple repetitions of the transmission. Without loss of generality, LoRaWAN is used as an example in this article but any other IoT protocol could be utilized. Thus, the major sources of the packet loss are described in the following subsections, as shown in Figure 1.

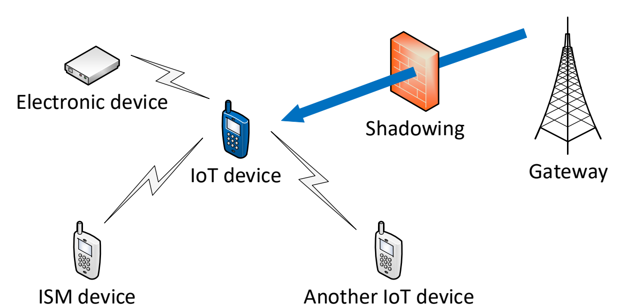


Fig. 1: An envisioned IoT network with the potential factors of the packet losses.

A. Packet collision

Channel contention occurs when multiple devices attempt to send data over the same channel simultaneously [7], [14]. Whereas the end nodes are uncoordinated, the packet transmission is initiated by the end node, not the network. Hence, the traditional LoRa device does not check if the channel is preempted by other devices before transmitting a packet, so

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2022.3167016, IEEE Internet of Things Journal

TABLE I: Main distinctive characteristics of different solutions for LoRa

Algorithm	Decentralized	No synchronization among the end nodes	Utilizing the channel quality information
DyLoRa [11]	✗	✓	✓
CARA [12]	✗	✗	✗
UCB [13]	✓	✓	✗
QoC-A & **DQoC-A**	✓	✓	✓

TABLE II: ACRONYMS

ACK	Acknowledgment
ADR	Adaptive Data Rate
AI	Artificial Intelligence
AS	Application Server
CARA	Collision Avoidance Resource Allocation
CSI	Channel State Information
CSS	Chirp Spread Spectrum
DQoC-A	Discounted Quality of Channel Allocation
DyLoRa	Dynamic LoRa
ESP	Effective Signal Power
IoT	Internet of Things
ISM	Industrial, Scientific and Medical
LPWAN	Low Power Wide Area Networks
LoRaWAN	Long Range Wide Area Network
LNS	LoRa Network Server
MAB	Multi-Armed Bandit
OSA	Opportunistic Spectrum Access
OTAA	Over-The-Air-Activation
PDR	Packet Delivery Rate
QoC-A	Quality of Channel Allocation
RF	Radio Frequency
RSSI	Received Signal Strength Indicator
SNR	Signal-to-Noise Ratio
SF	Spreading Factor
UCB	Upper Confidence Bounds

it may cause packet collision. Additionally, the IoT networks are superposed with no coordination between them. So, these collisions could also occur no matter whether the interfered signal is from other end nodes of the same network or the surrounding IoT networks, using the same IoT standard or not. Moreover, interference could occur from other radio signals present in the ISM bands which are not IoT signals, so they can be considered “jammers” by definition. Last but not least, electromagnetic radiations could disturb the end nodes due to the proximity of other electronic devices suffering from leakage radiations, for instance, this could happen in an industrial environment.

B. Channel factors

The end node can be located far away from the gateway, especially in LPWAN networks, which causes shadowing and signal attenuation across the transmission range [17]. Packet transmission success is closely related to the channel state, indeed, a channel facing bad propagation conditions has the same effect as caused by collisions. These propagation conditions are directly determined by the channel parameters which are SNR, RSSI and ESP. SNR is the ratio of signal power to the measured noise power. Consequently, the higher SNR is, the smaller the noise mixed in the signal, and the easier it is to separate the effective signal. RSSI is a relative value of signal power measured by the end node or the gateway at its receiving end. However, ESP is a more reliable parameter as it overcomes the RSSI limitation, especially at low SNR ($< 0\,\mathrm{dB}$) as depicted in [8] and [9].

Without loss of generality, ESP is computed as:

$$ESP_{dBm} = RSSI_{dBm} + SNR_{dB} - 10\log_{10}(1 + 10^{\frac{SNR_{dB}}{10}}). \quad (1)$$

In general, ESP value decays with the shadowing effect and increase of distance, in particular for the long transmission range, as in most IoT applications. Due to the low transmission power, i.e. maximum transmit power of the LoRa signal is $14\,\mathrm{dBm}$, and the large propagation attenuation, this value is usually negative in dBm. Furthermore, these channel parameter values depend on the specific environmental conditions around each end node, hence, each CSI is different from one location to another.

III. System Model

The proposed learning approach can help IoT devices to reduce packet losses due to both weak channel propagation conditions and also collisions. That for, the proposed approach is inspired from [13] and [14] where a solution to radio collisions is proposed, while imposing no change on the IoT protocols, as, for instance, LoRaWAN. Employing the acknowledged messaging mode is the only required condition to utilize the proposed solution. Based on the investigation in [8], the underlying hypothesis is that the channels' Packet Delivery Rate (PDR) is not equally balanced across the K different frequency bands an IoT device can use. In other words, some channels are less attenuated than others. These channel conditions are possibly be predicted online in time and space in a decentralized manner (on the end node side). As the end nodes may be quite far away from the gateways and suffer from different jamming and channel conditions, it is much more efficient to implement a spectrum allocation approach on the end node side than on a centralized unit. While taking into consideration that no extra processing can be afforded at the end node where every Watt is counted at transmission to save energy.

To be compatible with the constraint of low complexity of the end node hardware, a kind of Artificial Intelligence (AI) algorithm is considered [18]. This proposed approach is based on reinforcement learning algorithms which have been introduced by the machine learning community [19] and first proposed for cognitive radio communications more than 10 years ago [20]. It is also experimentally validated on real radio signals for cognitive radio and especially for Opportunistic Spectrum Access (OSA) in [21]. As asserted for OSA, Multi-Armed Bandit (MAB) problem can be utilized to model the IoT spectrum access issue [13]. Thus,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2022.3167016, IEEE Internet of Things Journal

reinforcement learning is based on a feedback loop that gives a success/failure measure of the action. In the IoT context, a binary reward, i.e. 1|0 for the presence|absence of the ACK packet which is sent by the gateway to the end node, is considered, as shown in Figure 2. If a message has been successfully transmitted from the end node to the gateway on the uplink channel, as well as the ACK message has been successfully transmitted from the gateway and received by the end node on the same channel in the downlink, a reward of "1" is given plus the channel parameters, i.e. SNR, RSSI and ESP, are estimated to optimize and accelerate the learning process as illustrated in the following sections. While a reward of "0" means that packet loss has happened either on uplink or downlink so that ACK has not been received by the end node. Maximizing the PDR is the main target, or equivalently, maximizing its cumulated reward. This proposed approach can adapt to any other IoT protocol, moreover, it has the following main advantages:

- Coordination between the end nodes is not necessary. Therefore, no extra retransmission, no data to be added into the uplink or downlink packet. While the content of the ACK packet isn't changed.
- The very low processing and memory overhead of both the implementation and execution of the proposed approach [13]. Consequently, it is possible to execute the proposed algorithm in the end node devices whose complexity is negligible in terms of processing, hardware, memory as well as energy consumption overhead.
- The strong mathematical proof of convergence of the Bandit algorithms [22], [23]. Thanks to the good matching between models and reality, these proofs are verified in real radio conditions [21].
- Rapid convergence of the proposed learning techniques in real experiments [24].
- Thanks to the reinforcement learning concept, any prior training isn't needed. This proposed algorithm can efficiently start learning from scratch.
- These proposed learning algorithms' results will always outperform the given state-of-the-art ones with random frequency allocation [24].

IV. Proposed Reinforcement Learning Techniques

Bandit algorithms are used at the end node side to handle IoT wireless spectrum issues. While LoRaWAN is the considered example, a simple ALOHA-based protocol is the utilized communications between the end nodes and the gateway. Whenever the end nodes decide, they can transmit their packets in one of the $K \geq 2$ channels which are predefined in frequency f_i. As stated in the previous section, the channel conditions in unlicensed ISM bands suffer in

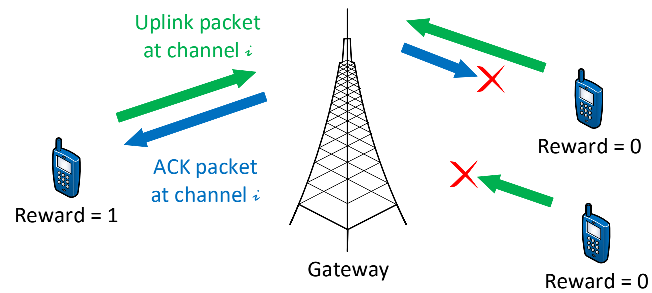


Fig. 2: An end node sends a packet with the acknowledged messaging mode, whereas the reward value depends on the presence|absence of the ACK packet.

particular from propagation conditions and interference which are different from one place to another. Even in one place, these channel conditions are very often unevenly distributed over the K different channels.

From the point of view of a single end node in the network, it has to choose one channel (or arm), denoted as $i \in \{1, ..., K\}$, every step index n when it sends a packet to the gateway, then, it starts to wait for a fixed delay (one second in LoRaWAN) in the same channel i for receiving an ACK packet from the gateway, as shown in Figure 2. But due to propagation issues, the message sent by the end node to the gateway, or the ACK sent by the gateway to the end node, could be lost. Hence, selecting the channel i at packet index n yields a random feedback, i.e. the binary reward $r_i(n)$ and the channel quality $g_i(n)$. Maximizing the transmission success rate of the end node or, equivalently, maximizing the cumulative reward is the main target. Sequence of rewards drawn from a given arm i is assumed to be $i.i.d.$, consequently, this problem is considered as a "stochastic" MAB [23]. A player (here, an end node) has to try all arms (here, channels), a sufficient number of times to get a robust estimate of their qualities, while not selecting the worst arms too much. This requires tackling the so-called exploration-exploitation dilemma to be able to progressively focus on the best arm i.e., the arm with the largest average reward, using the proposed algorithms in the following subsections.

A. *UCB Algorithm*

Exploiting the channel with the highest estimated mean by selecting it at each time could be a first naive solution, however, this "greedy" approach is known to fail dramatically [19]. Thus, the selection of arms is highly dependent on the first draws with this greedy policy. For example, if the first transmission in one channel succeeds and the first one on the other channel fails, the end node will never use the other channel again, even it is the best one and has the best channel conditions on average. While Upper Confidence Bounds (UCB) algorithm instead is adding an extra exploration term to the empirical mean, which can be viewed as a confidence bound [23]. At each end node, the number of times the channel i is selected up-to packet index $n \geq 1$ is calculated as:

$$T_i(n) = \sum_{m=1}^{n} \mathbb{1}_{A(m)=i} \; \forall i, \tag{2}$$

where $A(m) \in \{1, ..., K\}$ is a discrete action which corresponds to the chosen channel index at packet index m. Hence, the empirical mean estimation of the successful transmissions obtained in channel i by selecting it up to packet index n, is denoted by the mean reward as:

$$R_i(n) = \frac{1}{T_i(n)} \sum_{m=1}^{n} r_i(m) \mathbb{1}_{A(m)=i} \; \forall i. \tag{3}$$

Subsequently, the upper confidence bound $B_i(n)$ of the channel i is denoted by the sum of the empirical mean and a confidence term as:

$$B_i(n) = R_i(n) + \alpha \sqrt{\frac{\ln n}{T_i(n)}}, \tag{4}$$

where the exploration factor α is recommended by the theory to be ≥ 0.25 [25]. At this point, the maximum upper confidence bound $B_i(n)$ is chosen by the end node to decide the next channel for sending the packet, thus, the next action $A(n+1)$ is obtained by choosing the best channel as:

$$A(n+1) = \arg\max_i (B_i(n)). \tag{5}$$

For each end node, the packet index n corresponds to the total number of transmitted packets from the beginning. As time is not slotted, this packet index n isn't shared across different end nodes. Every end node implements its own UCB algorithm independently, which can be described as follows in Algorithm 1.

Data: $K, [\alpha]$
Result: $A(n+1)$
for $n = 1$ *to* ∞ **do**
    if $n < K$ **then**
        Initialize policy by trying each channel for at least one time.
        $A(n+1) = n+1$
    else
        $T_i(n) = \sum_{m=1}^{n} \mathbb{1}_{A(m)=i} \; \forall i$
        $R_i(n) = \frac{1}{T_i(n)} \sum_{m=1}^{n} r_i(m) \mathbb{1}_{A(m)=i} \; \forall i$
        $B_i(n) = R_i(n) + \alpha \sqrt{\frac{\ln n}{T_i(n)}}$
        $A(n+1) = \arg\max_i (B_i(n))$
    end
end

Algorithm 1: UCB policy as depicted in [23].

B. QoC-A Algorithm

As mentioned in the previous sections, there is a strong dependency between the PDR and the channel quality. Hence, a Quality of Channel Allocation (QoC-A) policy is proposed to find the channel which is optimal in terms of both quality and probability of receiving the ACK packet [26]. As first stated in Algorithm 2, all the channels are explored at least once to acquire the initial binary reward and their channel qualities. At the final step, Algorithm 2 returns the channel index $A(n+1)$ which has to be used in the next packet index. Again, $T_i(n)$ defines the number of times channel i has been utilized up to packet index n. After $n \geq K$ steps, the term $B_i(n)$, i.e. corresponding to the score of the $i-$ th channel at packet n, are updated as [26]:

$$B_i(n) = R_i(n) + Q_i(n) + \alpha \sqrt{\frac{\ln n}{T_i(n)}}, \tag{6}$$

where $R_i(n)$ denotes the exploitation term, or equivalently the empirical mean of the states (ACK packet received or not) of the $i-$ th channel at packet index n. While the bias term $\alpha\sqrt{\frac{\ln n}{T_i(n)}}$ forces to explore the other channels, if the scheme leads to a channel whose average rewards degrades. Furthermore, the quality term $Q_i(n)$ defines the quality information of the channel i which can be calculated as:

$$Q_i(n) = \beta(\frac{G_i(n)}{G_{max}(n)} - 1)\frac{\ln n}{T_i(n)} \tag{7}$$

with

$$G_i(n) = \frac{1}{T_i(n)} \sum_{k=1}^{T_i(n)} g_i(k) \tag{8}$$

and

$$G_{max}(n) = \max_i G_i(n), \tag{9}$$

where $G_i(n)$ is the empirical mean of quality observations $g_i(n)$ collected from channel i and $G_{max}(n)$ is the maximum expected quality value within the set of channels. Moreover, the new parameter β forces the algorithm to give some weight to the quality term $Q_i(n)$ in the score computation, whereas the parameter α forces the exploration of other channels to check their probability of receiving the ACK packet. Such a formulation tends to select a channel with the highest quality and probability to be acknowledged. For instance, if $G_i(n) = G_{max}(n)$, the term $Q_i(n)$ will be equal to zero. While if the average quality of channel i is less than the average quality of the best channel ($G_i(n) < G_{max}(n)$), the value of $Q_i(n)$ will be negative and then the score term $B_i(n)$ will decrease. Accordingly, channel i will be less likely to be selected.

An important contribution of this article compared to the previous works in the cognitive radio field [26], is that the quality observation $g_i(n)$ is the channel quality, i.e. ESP in linear scale, of the ACK packet. On the contrary, the channel parameters are obtained from a sensing phase before sending the packet in the OSA context, as the "listen-before-talk" model in [27], which consumes power. It also worth mentioning that ESP is the used parameter that is obtained as a function of SNR, thus, it has an enlarged range when the SNR is very low, unlike the RSSI which has a limitation as detailed in Section II. Accordingly, this proposed algorithm is simple to implement and to use in practice, even on embedded microprocessors with limited computation and memory capabilities as detailed in the following sections.

Data: $K, [\alpha, \beta]$
Result: $A(n+1)$
for $n = 1$ *to* ∞ **do**
  if $n < K$ **then**
    Initialize policy by trying each channel for at least one time.
    $A(n+1) = n+1$
  else
    $T_i(n) = \sum_{m=1}^{n} \mathbb{1}_{A(m)=i} \ \forall i$
    $R_i(n) = \frac{1}{T_i(n)} \sum_{m=1}^{n} r_i(m) \mathbb{1}_{A(m)=i} \ \forall i$
    $G_i(n) = \frac{1}{T_i(n)} \sum_{k=1}^{T_i(n)} g_i(k)$
    $G_{max}(n) = \max_i G_i(n)$
    $Q_i(n) = \beta(\frac{G_i(n)}{G_{max}(n)} - 1) \frac{\ln n}{T_i(n)}$
    $B_i(n) = R_i(n) + Q_i(n) + \alpha \sqrt{\frac{\ln n}{T_i(n)}}$
    $A(n+1) = \arg\max_i (B_i(n))$
  end
end

Algorithm 2: QoC-A policy.

C. DQoC-A Algorithm

In many IoT applications, the end node could be in a non-stationary scenario, for instance, when it moves across different locations whose CSI shapes are different. Based on that, the probability of receiving the ACK packet and the quality of each channel are likely to experience changes in time, which exhibits the limitation of the aforementioned MAB algorithms. Although the convergence of these algorithms is very fast, however, the stationarity of the environment is required. Resetting the learning algorithm from time to time can be a simple solution, however, it may fail to determine when the propagation conditions change as the acquired reward distribution evolves with time occasionally.

In this article, a Discounted QoC-A (DQoC-A) algorithm is proposed to learn on the same channel conditions, i.e. quality and the probability of receiving the ACK packet, as the previous QoC-A but in non-stationary scenarios [28]. Since the confidence interval of the standard QoC-A policy becomes tighter when time goes up, it is not appropriate for the non-stationary environment as stated before. While the motivation for the DQoC-A policy is to find an optimal channel in the case of changing environments, with less exploration. Therefore, discount factors (λ and λ_g) are considered for the DQoC-A to guaranty the adaptiveness of DQoC-A policy in a non-stationary environment, as stated in Algorithm 3. The idea behind the inclusion of these discount factors (λ and λ_g) is to give more weight to recent observations compared to the ones acquired in the past. A remarkable contribution of this article compared to the previous work in [28], is that two different discount factors (λ and λ_g) are considered rather than using one. Thus, λ and λ_g should have different values as they are used to average two different distributions, which are the binomial distribution of the binary reward $r_i(n)$ and the log-normal distribution of the channel quality $g_i(n)$, respectively.

Based on that, this proposed DQoC-A policy learns a channel that is optimal in terms of the probability of receiving the ACK packet and quality in a gradual manner. As with QoC-A policy, an end node employing DQoC-A policy first starts to explore all channels at least once initially. After $n \geq K$ steps, it updates the scoring term $B_i(n)$, however, each term in the equation is adapted to take into account the non-stationary hypothesis as:

$$B_i(n) = R_i(n) + Q_i(n) + \alpha \sqrt{\frac{\ln W(n)}{N_i(n)}}, \tag{10}$$

where $N_i(n)$ is the discounted number of times channel i has been used up to packet index n, and $W(n)$ is the total discounted time. Contrary to QoC-A policy, the empirical mean of rewards $R_i(n)$ and channel quality $G_i(n)$ are estimated by taking into account the discount factors ($0 < \lambda < 1$ and $0 < \lambda_g < 1$), as shown in Algorithm 3. While the coefficients α and β, are the same as in the QoC-A policy, to weight exploration for the probability of receiving the ACK packet and channel quality, respectively. Furthermore, a relative comparison between the three policies is depicted in Table III.

Data: $K, [\alpha, \beta, \lambda, \lambda_g]$
Result: $A(n+1)$
for $n = 1$ *to* ∞ **do**
  if $n < K$ **then**
    Initialize policy by trying each channel for at least one time.
    $A(n+1) = n+1$
  else
    $N_i(n) = \sum_{m=1}^{n} \lambda^{n-m} \mathbb{1}_{A(m)=i} \ \forall i$
    $W(n) = \sum_{i=1}^{K} N_i(n)$
    $R_i(n) = \frac{1}{N_i(n)} \sum_{m=1}^{n} \lambda^{n-m} r_i(m) \mathbb{1}_{A(m)=i} \ \forall i$
    $Ng_i(n) = \sum_{m=1}^{n} \lambda_g^{n-m} \mathbb{1}_{A(m)=i} \ \forall i$
    $G_i(n) = \frac{1}{Ng_i(n)} \sum_{m=1}^{n} \lambda_g^{n-m} g_i(m) \mathbb{1}_{A(m)=i} \ \forall i$
    $G_{max}(n) = \max_i G_i(n)$
    $Q_i(n) = \beta(\frac{G_i(n)}{G_{max}(n)} - 1) \frac{\ln W(n)}{N_i(n)}$
    $B_i(n) = R_i(n) + Q_i(n) + \alpha \sqrt{\frac{\ln W(n)}{N_i(n)}}$
    $A(n+1) = \arg\max_i (B_i(n))$
  end
end

Algorithm 3: DQoC-A policy.

V. Experiment Setup

Comparing the proposed algorithms against the state-of-the-art method with random frequency allocation in real conditions of operation, is the main target of the experiment. This is done by setting the LoRaWAN configuration as

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2022.3167016, IEEE Internet of Things Journal

TABLE III: Comparison of the implemented policies

Algorithm	Node mobility condition	Complexity	Preconfigured parameter	Online parameter
UCB	Stationary	$O(K)$	$K, [\alpha]$	Binary reward
QoC-A	Stationary	$O(K)$	$K, [\alpha, \beta]$	Binary reward + ESP
DQoC-A	Stationary | Non-stationary	$O(K)$	$K, [\alpha, \beta, \lambda, \lambda_g]$	Binary reward + ESP

presented in Table IV, but it could be done with any other IoT standard, as soon as it uses acknowledged messaging. A Tektelic KONA Macro Gateway is used whose antenna is fixed on the roof of the university building [29], as shown in Figure 3a and 3b. While the end node is implemented using a Pycom card, i.e. programmed in the MicroPython language, composed of an Expansion Board and a LoPy 4 module which can support LoRa wireless connectivity [30], as shown in Figure 3c.

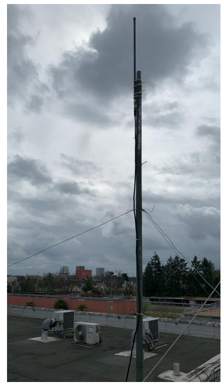
TABLE IV: LoRaWAN configuration

LoRaWAN Parameter	Value
Modulation technique	LoRa (based on CSS)
SF	7
Coding rate	4/5
Bandwidth W	$125\,\mathrm{kHz}$
Transmission power	$14\,\mathrm{dBm}$
Number of channels K	8
Center frequency f_i	$\{867.1, 867.3, 867.5, 867.7, 867.9, 868.1, 868.3, 868.5\}$ MHz

First, the network is joined with an Over-The-Air-Activation (OTAA). Then, the proposed algorithms are executed sequentially at each step index n inside the Pycom node. An uplink packet is transmitted at the channel i which is chosen by the algorithm, then the node waits about 10 seconds before sending the next packet to respect the duty cycle. By default, the gateway attempts to send one acknowledgment at the same channel i and center frequency f_i as the message transmitted. Subsequently, the node writes the information of the last received downlink packet (packet number, ESP, etc.) to the payload of the next uplink packet. In this experiment, there is no retransmission attempt if the node does not receive the acknowledgment packet. For analyzing purposes only in this experiment, a desktop computer that runs a Python program is used as an Application Server (AS) which is connected with the LoRa Network Server (LNS) through the internet. This computer receives data from the LNS, as well as LoRa metadata with all parameters of the LoRaWAN transmission (f_i, SF, W, RSSI, SNR, etc.). Those data are analyzed as detailed in the following sections. Moreover, they are provided to the research community on this online repository [31].

VI. Measurement Results

The scope of this section is to compare the aforementioned learning policies against the state-of-the-art ones with random frequency allocation that is considered as the reference. For emulating the state-of-the-art frequency allocation, a uniform Round-Robin algorithm is executed that simply transmits the packets sequentially at each frequency band f_i across the packet index n. Furthermore, the experiments are performed in two different environments as detailed in the following subsections.



(a) Position of the gateway monopole antenna fixed on the roof of the IETR lab.

(b) The gateway connected to the antenna cable.

(c) The packaged Pycom device is fixed on a rod and connected to a battery inside a building.

Fig. 3: Views from the end node and gateway sites.

A. Scenario 1: Stationary IoT node

The first experiment is carried out by placing the end node in a fixed position without moving it to have stationary channel conditions. Hence, the CSI shapes are almost identical across the whole experiment duration of ≈ 8.8 hours and 800 steps (each algorithm transmits 800 packets), as shown by the received power level in Figure 4. Each CSI shape is acquired by averaging the ESP values at each frequency band f_i independently every 200 steps. Moreover, the frequency selectivity of this CSI is clear, particularly at $867.3\,\mathrm{MHz}$ with a deep fade of more than $10\,\mathrm{dB}$ depth. These unequal channel qualities across the frequency bands f_i could be exploited using the proposed algorithm QoC-A.

As depicted in Figure 5, the performance of UCB ($\alpha = 0.6$), and QoC-A ($\alpha = 0.6$ and $\beta = 0.2$), i.e. with ESP in the linear scale as the quality observation $g_i(n)$ that is shown in Algorithm 2, are compared against the uniform frequency allocation. As promised in the previous sections, all the learning algorithms outperform the uniform frequency allocation. On the other hand, the final average reward obtained using the QoC-A policy outperforms the UCB policy, as manifested in Figure 5a. Accordingly as shown in Figure 5b, the cumulative regret obtained for all the methods preserve the same performance rank over the whole packet index n with final total lost packets of 32, 39 and 132, while using the proposed algorithm QoC-A, the classical UCB, and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2022.3167016, IEEE Internet of Things Journal

the uniform frequency allocation, respectively.

As shown in Table V, QoC-A policy outperforms all the other algorithms by a total number of successfully transmitted packets of 768 over 800, which is actually 4.1 times fewer packet losses than the uniform frequency allocation. This indicates that the proposed policy QoC-A is exploiting properly the channel quality to converge faster. Hence, it does not need to lose some time to acquire this knowledge before learning actually ends.

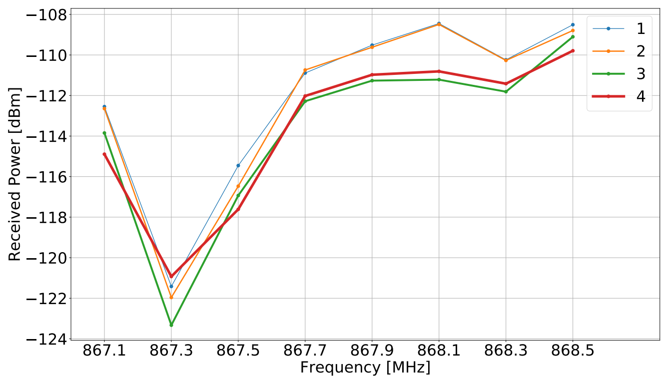


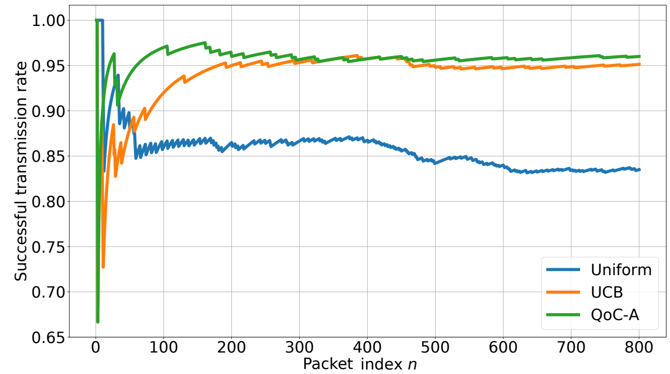
Fig. 4: Stability of the CSI shape throughout the 800 steps (one CSI every 200 steps) in the stationary scenario.

TABLE V: Number of successfully transmitted and lost packets in the stationary scenario

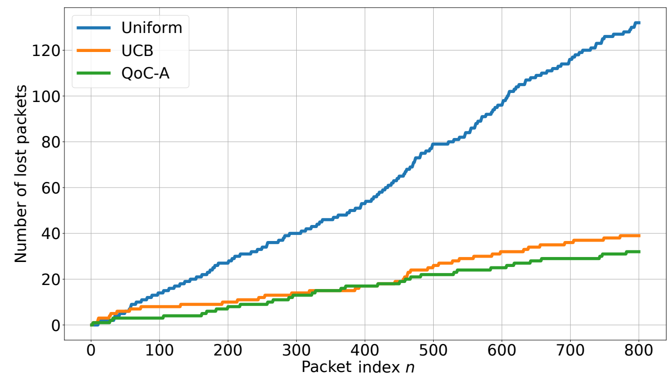
Algorithm	Succeed	Lost	Success rate
Uniform	668	132	83.5%
UCB	761	39	95.1%
QoC-A	768	32	96.0%

B. Scenario 2: Non-Stationary IoT node

For analyzing the behavior of MAB learning policies in a non-stationary scenario with an abrupt changing environment, the end node is moved during the experiment across three different positions to have unconstant channel conditions. As shown in Figure 6, the CSI shapes are different across the whole experiment duration of $\approx$ 6.6 hours and 600 steps (each algorithm transmits 600 packets). Again, each CSI shape is estimated by averaging the ESP values every 200 steps at each frequency band f_i independently. Although all the locations are indoor, nevertheless, a progressive reduction is observed in the received power across locations 1 to 3. This reduction is reasonable as the end node goes deeper inside the building from location 1, i.e. behind the window as shown in Figure 3c, to location 3. Moreover, the CSI shape which is represented in location 1 shows fades of $\approx$5 dB, especially at 867.1 MHz and 868.5 MHz, while the best channel quality is obtained at 867.9 MHz. The opposite behavior is obvious in the CSI shapes of locations 2 and 3 whose best channel conditions are almost around 868.1 MHz, 868.3 MHz and 868.5 MHz. While they have degradation in the channel quality at 867.9 MHz, contrary to the CSI of



(a) Average reward over 800 steps.



(b) Cumulative regret over 800 steps.

Fig. 5: Result of the algorithms against the uniform frequency allocation (non-learning) in the stationary scenario.

location 1. These abrupt changes in the channel qualities from one location to another could be learned using the proposed algorithm DQoC-A.

In this non-stationary environment, identifying the abrupt change in the reward distribution with reduced delay should be the ability of an optimal policy. As shown in Figure 7, the performance of UCB ($\alpha = 0.6$), QoC-A ($\alpha = 0.6$ and $\beta = 0.2$) and DQoC-A ($\alpha = 0.6$, $\beta = 0.2$, $\lambda = 0.98$ and $\lambda_g = 0.90$) are compared against the naive frequency allocation. For both the QoC-A and DQoC-A algorithms, the linear scale of the ESP value is utilized as the quality observation $g_i(n)$. Moreover, the discount factor λ is equal to 0.98 in the DQoC-A policy, while λ_g is set to 0.90 to acquire more rapidly the most recent ESP values, which follows a log-normal distribution from the channel shadowing when the end node moves and subsequently to converge faster. Motions of the end node are represented by the two breakpoints at step index $n = 200$ and $n = 400$ in which an abrupt change in the reward distribution is indicated. Before the first breakpoint (at $n = 200$), the evolution of the cumulative regret is almost flat using any algorithm. This is a reasonable behavior as most of the frequency bands f_i at location 1 have relatively high ESP values and low packet losses. While after the first breakpoint (at $n = 200$), DQoC-A achieves significantly lower regret

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2022.3167016, IEEE Internet of Things Journal

and higher average reward than QoC-A and UCB policy in this non-stationary environment.

As shown in Table VI, DQoC-A policy outperforms all the other algorithms by a total number of successfully transmitted packets of 520 over 600, which is 2.4 times fewer packet losses than the uniform frequency allocation. This is plausible due to the inclusion of the discount factors (λ and λ_g) in the DQoC-A calculation, so it wastes significantly less time than QoC-A to identify an abrupt change (at $n = 200$ or $n = 400$) in the channel conditions. While the UCB and QoC-A are prevented to adapt quickly to changes as the old rewards have a higher influence on them. In other words, all the learning algorithms converge to the optimal mean reward in the long run, but DQoC-A policy benefits from less dependency on past observations.

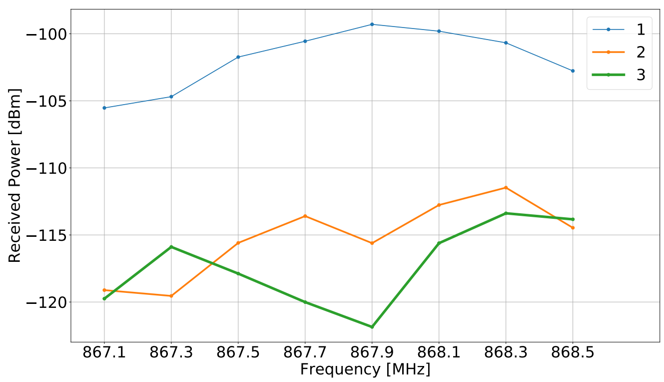


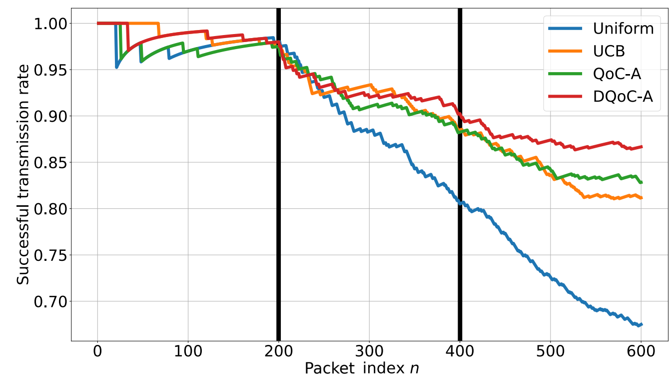
Fig. 6: Evolution of the CSI shape throughout the 600 steps (one CSI every 200 steps) in the non-stationary scenario.

TABLE VI: Number of successfully transmitted and lost packets in the non-stationary scenario

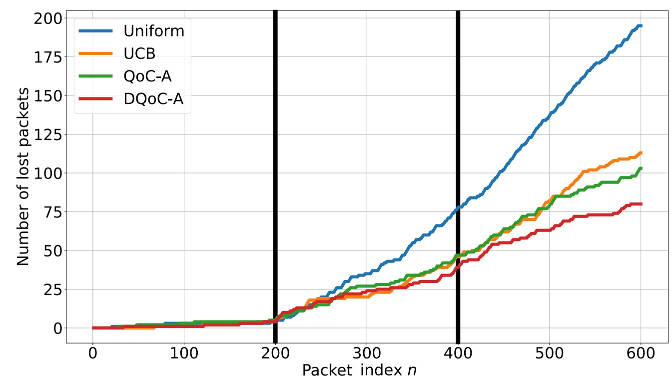
Algorithm	Succeed	Lost	Success rate
Uniform	405	195	67.5%
UCB	487	113	81.2%
QoC-A	497	103	82.8%
DQoC-A	520	80	86.7%

VII. Additional Remarks on the Algorithms' Configurations

The aforementioned results confirm the feasibility of utilizing the proposed policies to achieve lower packet losses in different scenarios. However, the parameters of all the proposed learning algorithms should be adjusted based on the environment's conditions. Thus, the exploration factor α is chosen depending on the channel conditions in terms of channel qualities or presence of interference signals [32]. Furthermore, the influence of the optimal choice of α is higher as the number K of channels increases, and less as it decreases. In the case of a small number of arms, choosing the value of the exploration factor α does not influence the policy performance much. In the opposite case, if the number of arms to explore is large, the choice of this factor will



(a) Average reward over 600 steps.



(b) Cumulative regret over 600 steps.

Fig. 7: Result of the algorithms against the uniform frequency allocation (non-learning) in the non-stationary scenario.

have a greater influence. If, for example, the chosen alpha parameter is deviated from the optimal value for a large number of arms to be explored, significant degradation in performances may be observed. Hence, the algorithm will then have either a tendency either to over-explore or to over-exploit.

On the other hand, the discount factors (λ and λ_g) values should be adjusted based on the requirements of the potential application. Thus, they will regulate properly the amount of dependency on the old observations to track an abrupt change in the reward distribution. For example, if an IoT node sends one packet every a fixed short time, therefore, the discount factors in the DQoC-A algorithm should be close to one to benefit also from the older observations as the environment is stationary over such a short horizon of time, and vice versa.

VIII. Conclusion

This article introduces a decentralized learning technique to mitigate the channel impairments in the IoT signal propagation. Thus, the proposed QoC-A policy could learn a proper frequency allocation based on the channel quality, i.e. ESP, as an extra observation to avert using the channels whose quality is poor. Besides, another DQoC-A policy is proposed to adapt rapidly to any change in the channels' conditions, as a result of the IoT node motion for example. To demonstrate the low complexity of these proposed algorithms and the feasibility of implementing them on the IoT device side, at a very low cost of implementation and no protocol overhead, a real experiment campaign is carried out in the city of Rennes. Consequently, the results show that QoC-A has a more optimized and accelerated learning process than the classical UCB policy, and then it has lower packet losses at the end of the process. On the other hand, DQoC-A policy converges faster than QoC-A policy in the non-stationary scenario, thanks to the discount factor in DQoC-A algorithm which decreases the dependency on the old observations gradually.

For future work, these proposed reinforcement learning algorithms could be implemented to decrease the packet losses in the potential IoT applications. Thus, the configurations of the implemented policy should be adjusted based on the given conditions. It would also be interesting to conduct long-term evaluations of the behavior of the proposed algorithms with new practical measurements. Using a greater number of nodes in more conducted experiments, the feasibility of using these learning algorithms in highly interfering scenarios could be investigated. Over and above, we are currently conducting measurements to calculate the amount of energy consumed by an IoT device with and without executing the proposed algorithms.

ACKNOWLEDGMENT

The authors would like to thank Jean François Legendre from Gwagenn company for borrowing from him the gateway [33].

REFERENCES

[1] A. Abdelghany, B. Uguen and D. Lemur, "A Pre-processing Algorithm Utilizing a Paired CRLB for TDoA Based IoT Positioning," *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, 2020, pp. 1-5.

[2] M. A. Al-Jarrah, M. A. Yaseen, A. Al-Dweik, O. A. Dobre and E. Alsusa, "Decision Fusion for IoT-Based Wireless Sensor Networks," in *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 1313-1326, Feb. 2020.

[3] D. Ciuonzo, P. S. Rossi and P. K. Varshney, "Distributed Detection in Wireless Sensor Networks Under Multiplicative Fading via Generalized Score Tests," in *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 9059-9071, June 2021.

[4] R. Oliveira, L. Guardalben and S. Sargento, "Long Range Communications in Urban and Rural Environments," 2017 *IEEE Symposium on Computers and Communications (ISCC)*, pp. 810-817, 2017.

[5] LoRa Alliance website. [Online]. Available: https://www.lora-alliance.org/.

[6] A. Augustin, J. Yi, T. Clausen and WM. Townsley, "A Study of LoRa: Long Range & Low Power Networks for the Internet of Things," *Sensors 2016*, vol. 16, no. 9, p. 1466.

[7] A. Rahmadhani and F. Kuipers, "When LoRaWAN Frames Collide," in *Proceedings of the 12th International Workshop on Wireless Network Testbeds*, Experimental Evaluation & Characterization. Association for Computing Machinery, New York, NY, USA, pp. 89–97, 2018.

[8] A. Abdelghany, B. Uguen, C. Moy and D. Lemur, "On Superior Reliability of Effective Signal Power versus RSSI in LoRaWAN," *28th International Conference on Telecommunications (ICT)*, London, United Kingdom, pp. 1-5, 2021.

[9] A. Abdelghany, B. Uguen, C. Moy and D. Lemur, "Modelling of the Packet Delivery Rate in an Actual LoRaWAN Network," *Electronics Letters*, vol. 57, no. 11, pp. 460-462, 2021.

[10] Q. Liu, Y. Mu, J. Zhao, J. Feng and B. Wang, "Characterizing Packet Loss in City-Scale LoRaWAN Deployment: Analysis and Implications," *2020 IFIP Networking Conference (Networking)*, Paris, France, pp. 704-712, 2020.

[11] Y. Li, J. Yang and J. Wang, "DyLoRa: Towards Energy Efficient Dynamic LoRa Transmission Control," *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*, Toronto, ON, Canada, 2020, pp. 2312-2320.

[12] N. Chinchilla-Romero, J. Navarro-Ortiz, P. Muñoz and P. Ameigeiras, "Collision Avoidance Resource Allocation for LoRaWAN," *Sensors 2021*, vol. 21, no. 4, p. 1218.

[13] C. Moy, "IoTligent: First World-Wide Implementation of Decentralized Spectrum Learning for IoT Wireless Networks," *2019 URSI Asia-Pacific Radio Science Conference (AP-RASC)*, 2019, pp. 1-4.

[14] 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*, vol. 75, pp. 711–727, 2020.

[15] S. Rauh, T. Lauterbach, H. Lieske, J. Robert and A. Heuberger, "Multipath Characteristics of Indoor-to-Outdoor Radio Channels in the 868-MHz Band," *Smart SysTech 2016; European Conference on Smart Objects, Systems and Technologies*, 2016, pp. 1-11.

[16] Q. Guo, F. Yang and J. Wei, "Experimental Evaluation of the Packet Reception Performance of LoRa," *Sensors 2021*, vol. 21, no. 4, p. 1071.

[17] R. El Chall, S. Lahoud and M. El Helou, "LoRaWAN Network: Radio Propagation Models and Performance Evaluation in Various Environments in Lebanon," in *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2366-2378, April 2019.

[18] M. Min et al., "Learning-Based Privacy-Aware Offloading for Healthcare IoT With Energy Harvesting," in *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4307-4316, June 2019.

[19] T. L. Lai and H. Robbins, "Asymptotically Efficient Adaptive Allocation Rules," *Advances in Applied Mathematics*, vol. 6, issue 1, 1985, pp. 4-22, ISSN 0196-8858, doi:10.1016/0196-8858(85)90002-8.

[20] W. Jouini, D. Ernst, C. Moy and J. Palicot, "Upper Confidence Bound Based Decision Making Strategies and Dynamic Spectrum Access," *2010 IEEE International Conference on Communications*, 2010, pp. 1-5.

[21] C. Moy, "Reinforcement Learning Real Experiments for Opportunistic Spectrum Access," *WSR'14*, Mar 2014, Karlsruhe, Germany. 10 p. [Online]. Available: https://hal-supelec.archives-ouvertes.fr/hal-00994975.

[22] R. Bonnefoi, L. Besson, C. Moy, E. Kaufmann and J. Palicot, "Multi-Armed Bandit Learning in IoT Networks: Learning helps even in non-stationary settings," *Cognitive Radio Oriented Wireless Networks and Communications 2017, CrownCom'17*, September 2017, Lisbon, Portugal, doi:10.1007/978-3-319-76207-4_15.

[23] P. Auer, N. Cesa-Bianchi and P. Fischer, "Finite-time Analysis of the Multiarmed Bandit Problem," *Machine Learning*, vol. 47, no. 2, pp. 235–256, 2002, doi:10.1023/A:1013689704352.

[24] L. Besson, R. Bonnefoi and C. Moy, "GNU Radio Implementation of MALIN: "Multi-Armed bandits Learning for Internet-of-things Networks", *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, 2019, pp. 1-6.

[25] S. Bubeck and N. Cesa-Bianchi, "Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems," *Foundations and Trends in Machine Learning*, vol. 5, no. 1, pp. 1–122, 2012, doi:10.1561/2200000024.

[26] N. Modi, P. Mary and C. Moy, "QoS Driven Channel Selection Algorithm for Cognitive Radio Network: Multi-User Multi-Armed Bandit Approach," in *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 1, pp. 49-66, March 2017.

[27] N. Modi, P. Mary and C. Moy, "QoS Driven Channel Selection Algorithm for Opportunistic Spectrum Access," *2015 IEEE Globecom Workshops (GC Wkshps)*, 2015, pp. 1-6.

[28] N. Modi, "Machine Learning and Statistical Decision Making for Green Radio," PhD dissertation, CentraleSupélec, 2017. [Online]. Available: https://tel.archives-ouvertes.fr/tel-01668536.

[29] Tektelic KONA Macro IoT Gateway. [Online]. Available: https://www.tektelic.com/uploads/Brochures/Kona%20Macro.pdf.

[30] Pycom documentation. [Online]. Available: https://GitHub.com/PyCom/PyCom-libraries.

[31] Measurement Data. [Online]. Available: https://gitlab.com/ahmednagy/lorawan-beaulieu-measurement-2021.git.

[32] L. Melián-Gutierrez, N. Modi, C. Moy, F. Bader, I. Perez-Álvarez and S. Zazo, "Hybrid UCB-HMM: A Machine Learning Strategy for Cognitive Radio in HF Band," in *IEEE Transactions on Cognitive Communications and Networking*, vol. 1, no. 3, pp. 347-358, Sept. 2015.

[33] Gwagenn company website. [Online]. Available: http://www.gwagenn.com/en/home/.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2022.3167016, IEEE Internet of Things Journal

Ahmed Abdelghany received the B.Sc. degree in communications engineering from the German University in Cairo (GUC), New Cairo City, Egypt, in 2014, and the M.Sc. degree from the University of Erlangen-Nuremberg in Bavaria, Germany, in 2017. In December 2021, he accomplished his Ph.D. degree with the Institut d'Electronique et des Technologies du numéRique (IETR/UR1), University of Rennes 1, Rennes, France. He is currently with the Propagation and Localization Group, in the IETR lab. His research areas of interest include wireless communication, signal processing, IoT localization, spectrum allocation and machine learning.

Bernard Uguen (Member, IEEE) received the Diploma degree in electrical engineering and the Ph.D. degree in electronics from the INSA Rennes, in 1992. From 1993 to 2006, he was an Associate Professor with the INSA Rennes. Since September 2006, he has been a Full Professor with the University of Rennes 1. Since 2009, he has been heading the Propagation and Localization Group, Institut d'Electronique et des Technologies du numéRique (IETR/UR1). He has authored or coauthored over 200 technical articles in international journals and conferences. His interests include the physics of propagation to statistical signal processing and data science. He has been working with the UWB technology, indoor positioning, and wireless BAN for wearables devices fields. He is currently involved in channel modeling for 5G mmW systems and LPWAN systems.

Christophe Moy (Senior Member, IEEE) received the Eng., M.Sc., and Ph.D. degrees in electronics from the INSA of Rennes, France, in 1995 and 1999, respectively. He is a Professor at the University of Rennes 1, where he has been teaching electronics in ISTIC syllabus since 2017. His research has been done since 2005 in IETR entity of CNRS (UMR 6164) where he focuses on software radio, cognitive radio (learning for dynamic spectrum access - DSA) and green radio. He has been particularly focusing on DSA for Opportunistic Spectrum Access (OSA) and Internet of Things (IoT). He was involved in many European and French collaborative projects. He was with Mitsubishi Electric ITE-TCL research lab, from 1999 to 2005, where he initiated studies on software radio systems and concepts. He was a Professor with Supelec from 2005 to 2014, and with CentraleSupélec from 2015 to 2017.

Dominique Lemur received the Ph.D. degree in electronics from the Université de Rennes 1, Rennes Cedex, France, in 1996. Since 1998, he has been an Associate Professor with the Université de Rennes 1, where he is a Member of the Institut d'Electronique et des Technologies du numéRique. He has designed numerous radio frequency systems, transmitting or active receiving antennas, and compacts receiving HF arrays. His current research interests include radio waves propagation and radiolocation in the HF and very/ultra high frequency bands, especially ionospheric radio applications for direction of arrival/sound source localization and improvement in data bit rates with high-frequency single input multiple output and MIMO modems.