



# Enhancing Coverage in Narrow Band-IoT Using Machine Learning

Marwa Chaffi, Faouzi Bader, Jacques Palicot

## ► To cite this version:

Marwa Chaffi, Faouzi Bader, Jacques Palicot. Enhancing Coverage in Narrow Band-IoT Using Machine Learning. IEEE Wireless Communications and Networking Conference (IEEE WCNC'2018), Apr 2018, Barcelona, Spain. 10.1109/wcnc.2018.8377263 . hal-01705715

**HAL Id: hal-01705715**

**<https://hal.science/hal-01705715>**

Submitted on 9 Feb 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Enhancing Coverage in Narrow Band-IoT Using Machine Learning

Marwa Chafii, Faouzi Bader, and Jacques Palicot  
IETR/CentraleSupélec, Campus de Rennes,  
Avenue de la Boulaie, 35510 Cesson-Sévigné, France  
Email: marwa.chafii@centralesupelec.fr

**Abstract**—Narrow Band-Internet of Thing (NB-IoT) is a recently proposed technology by 3GPP in Release-13. It provides low energy consumption and wide coverage in order to meet the requirements of its diverse applications that span social, industrial and environmental aspects. Increasing the number of repetitions of the transmission has been selected as a promising approach to enhance the coverage in NB-IoT up to 164 dB in terms of maximum coupling loss for uplink transmissions, which is a significant improvement compared with legacy LTE technologies, especially to serve users in deep coverage. However, a large number of repetitions reduces the system throughput and increases the energy consumption of the IoT devices, which reduces their battery lifetime and increases their maintenance cost. In this work, we propose a new method for enhancing the NB-IoT coverage based on machine learning algorithms. Instead of employing a random spectrum access procedure, dynamic spectrum access can reduce the number of required repetitions, increase the coverage, and reduce the energy consumption.

**Index Terms**—Narrow-band Internet of Things (NB-IoT), Coverage Enhancement (CE), Dynamic spectrum access, Reinforcement learning

## I. INTRODUCTION AND MOTIVATIONS

Internet of things (IoT) is becoming an increasingly growing topic due to its promise to change different aspects of our world. As recently predicted by Cisco, there will be 50 billions IoT connected devices by 2020, where each cell supports a massive number of devices (more than 50K connections per cell [1]). This rapid penetration of connected devices has the potential to impact the way we live, we work and how we interact with objects. The massive connected world advertised by IoT emerging companies requires a huge transfer of data, devices with long autonomy, as well as an extended coverage and indoor penetration.

In order to meet the requirements of IoT world, the 3GPP has designed the narrow band internet of things (NB-IoT) standard in its recent Release 13 [1]. The main design objectives of NB-IoT are increased coverage, long battery life (between 10 and 15 years), and low user equipment (UE) device complexity. To achieve these targets, several techniques have been adopted, including repetitions, power spectral density boosting, single-tone transmission, power saving mode, phase rotated modulations (e.g.  $\frac{\pi}{2}$ -BPSK and  $\frac{\pi}{4}$ -QPSK) to reduce peak-to-average power ratio in the uplink and so forth.

Other technologies have preceded the proposition of NB-IoT such as the LoRaWan standard [2] and the ultra narrow band protocol used by Sigfox [3], which operate in the unlicensed

bands (industrial, scientific and medical (ISM) radio bands). However, cellular IoT standards which operate in a licensed spectrum technologies have been worth the wait. In fact, NB-IoT can be deployed in any of the 2G/3G/4G spectrum (from 450 MHz to 3-5 GHz), since it achieves an excellent co-existence and compatibility performance with legacy cellular systems. NB-IoT needs only a small portion of the existing available cellular spectrum to operate without interfering with it. Hence, NB-IoT provides more reliability and more quality of service (QoS) as it operates in regulated spectrum. Moreover, NB-IoT uses existing cellular network infrastructure, which reduces the deployment costs.

However, since repeating transmission data and control signals has been selected as a major solution to enhance coverage of NB-IoT systems, this leads to reducing the system throughput and thereby a spectral efficiency loss. Another cost of providing deep indoor coverage is lowering the number of supported devices per sector. Hence the interest of studying new techniques of extending the coverage and reducing the number of retransmissions.

In this work, we propose a new method to enhance the coverage based on machine learning. Instead of getting access to the spectrum in a random way, dynamic spectrum access based on reinforcement learning algorithms can help increasing the coverage, decreasing the number of repetitions, and thereby reducing the energy consumption. This problem can be modelled as multi-armed bandit (MAB) framework, as has been proposed for cognitive radio in the literature [4]. It has also been reported that MAB works well in real channel conditions [5]. One of the most relevant methods to address MAB problems is upper confidence band (UCB) algorithms, which are efficient and converge quickly when the traffic is stationary and independent and identically distributed [6]. Using this framework, the NB-IoT end device will select the best channel following not only the criterion of availability but also the best coverage condition.

This paper is organized as follows: Section II explains the NB-IoT challenges in terms of coverage enhancement and energy consumption, while Section III shows how can dynamic spectrum access and its tools enable to follow the requirements of NB-IoT. Section IV provides simulation results about our proposed solution. Finally, Section V concludes this paper.

## II. NB-IoT CHALLENGES

Enhancing the coverage and reducing the energy consumption are among the most relevant key targets of NB-IoT. In this section, we explain the challenges of increasing the coverage and saving energy as well as the recently deployed solutions to fulfill these main requirements.

### A. Coverage enhancement

NB-IoT requires 20 dB of maximum coupling loss (MCL) higher than LTE, and then reaches up to 164 dB of MCL in order to serve end devices in deep coverage such as basements. Several modifications have been deployed on the different LTE protocol layers to achieve this significant gain. A major selected technique consists in increasing the number of retransmissions that reaches 128 repetitions for the uplink communications and 2048 for the downlink. These repetitions are combined at the receiver side in order to increase the signal-to-noise (SNR) ratio. Along with repeating the same transmission several times, other techniques have been used to extend the coverage such as cross-subframe channel estimation and frequency hopping. More details on these techniques can be found in [7].

Three coverage classes are allowed by a serving cell to an NB-IoT end device:

- CE level 0: normal coverage with  $MCL \approx 144$  dB and 15 kHz sub-carrier spacing.
- CE level 1: robust coverage with  $MCL \approx 154$  dB and 15 kHz sub-carrier spacing.
- CE level 2: extreme coverage with  $MCL \approx 164$  dB and 3.75 kHz sub-carrier spacing.

The choice of the coverage level depends on the channel conditions. The extreme coverage level corresponds to a low power received level, and a normal coverage level corresponds to a high power received level. Each selected coverage class determines the transmission parameters including the number of repetitions. Such a deployment allows the UE to be served in different coverage conditions characterized by different ranges of path loss. Depending on the coverage level, the serving cell indicates to the UE to repeat the transmission  $\{1, 2, 4, 8, 16, 32, 64, 128\}$  times, using the same transmission power on each repetition. Combining the different retransmissions allows a coverage extension.

### B. Energy consumption

In addition to the normal connected mode, there are mainly two energy efficient techniques that have been designed in NB-IoT in order to minimize the power consumption in end devices and increase their battery life:

- Idle mode extended discontinuous reception (I-eDRX): this mode allows a discontinuous reception for maximum of 3 hours, which saves UE battery but still allows it to be reachable by the network through paging messages or downlink control channels.
- Power saving mode (PSM): this energy saving mode allows unconnected state for up to 13 days, where UE enters to a deep sleep. Unlike in I-eDRX, UE is unreachable

while remains registered in the network. This mode save more energy than the idle mode.

As expected, the power consumption of these modes is substantially lower than the power consumed during transmission. Therefore, in a normal environment conditions (MCL of 154 dB), configuration of these power saving techniques allows a battery life of more than 10 years. However, in deep indoor coverage conditions, the targeted level of battery life (more than 10 years) cannot be achievable since the uplink repetitions get large. In addition, the energy consumption and the coverage enhancement mechanism imply high latency as the network waits a long period before being able to transmit its information data.

More techniques that allow both extending the coverage and reducing the number of required repetitions should be investigated to help prolong battery life. We show hereafter that dynamic spectrum access can help enhancing the coverage along with reducing the number of retransmissions and improving the latency.

## III. DYNAMIC SPECTRUM ACCESS CAN HELP

Instead of a random access based on slotted ALOHA which selects randomly the channel where to transmit in order to establish connection with the cell, we propose in this work to use a dynamic spectrum access in order to learn the channel which is more likely to be available and in good coverage conditions.

The spectrum learning process can be modelled as a multi-armed bandit (MAB) framework as proposed in [4], [8]. Depending on the location of UE (outdoor, indoor, basements) and the channel conditions (high or low SNR), the quality of the physical channels changes. Therefore, choosing the channel with the best quality (i.e. coverage level), potentially leads to reliable transmissions, less costly in terms of energy consumption.

### A. Multi-armed bandit framework

The MAB problem is a reinforcement learning game where a player have to decide which machine  $k$  to play (among  $K$  machines i.e  $k \in \{1, 2, \dots, K\}$ ) at each discrete time slot  $t = 0, 1, 2, \dots$ , based on informations of their reward. The player plays the machine that has the maximum reward. The rewards associated to each machine  $k$  are independent and identically distributed (i.i.d.) and follows a fixed and unknown distribution law  $d_k$ . In general, the reward distributions  $\{d_1, d_2, \dots, d_K\}$  differ from one machine to another, and the player does not have any knowledge about these distributions.

In our case, the player is the NB-IoT end device, and the machines are the spectrum channels used for cell connection. In the following we define some concepts related to the MAB framework.

**Reward** Let  $r_t(k)$  be the reward of the data transmission for a channel  $k$  at instant  $t$ . The reward in our scenario takes two values 0 or 1. The reward is equal to 1 if the selected channel is vacant (acknowledgement is received) and in a good coverage.

Note that the quality of the channel is only known when an acknowledgement is received.

**Exploitation and exploration dilemma.** It refers to a trade-off between the exploitation of the channel with highest mean reward, and exploring the other channels in order to get more information about their payoffs.

**Regret.** It means the loss represented by the difference between the expected reward associated to the suboptimal channel learned by the end device, and the ideal reward associated to the optimal channel. Since the user does not have any knowledge about the distribution of the reward, he cannot avoid a loss when selecting a channel.

Denote  $\pi$  the learning channels policy. Let  $\mu_k = \mathbb{E}[d_k]$  the stationary mean reward of the  $k^{\text{th}}$  channel, where  $\mathbb{E}[\cdot]$  denotes the expectation function. The regret of a policy  $\pi$  is defined as

$$R_t^\pi = t\mu^* - \sum_{l=0}^{t-1} r_l, \quad (1)$$

where  $\mu^*$  stands for the expected value of the reward of the optimal channel.

Based on (1), we define the expected cumulative regret as

$$\mathbb{E}[R_t^\pi] = \sum_{k=1}^K (\mu^* - \mu_k) \mathbb{E}[T_k(t)], \quad (2)$$

where  $T_k(t)$  being the total number of times channel  $k$  has been selected from instant 0 to instant  $t - 1$ .

The MAB problem can be solved using reinforcement learning algorithms such as UCB approaches. In the following section, we briefly define the use of UCB in dynamic spectrum access.

#### B. Upper confidence bound algorithm

The policy that we seek should help the NB-IoT device to make a decision on which channel to transmit. We choose to build a policy based on the UCB algorithms since this approach requires few processing resources and memory, and guarantees asymptotically optimal performance. The upper confidence bound index  $B_k(t)$ , is computed at each instant  $t$  and for each channel  $k$ , and gives an estimation of the expected reward of a channel  $k$ . The UCB index is expressed as:

$$B_k(t) = \bar{X}_k(t) + A_k(t), \quad (3)$$

$$\text{such that } \bar{X}_k(t) = \frac{1}{T_k(t)} \sum_{l=0}^{t-1} r_l(k) \mathbb{1}_{\{a_l=k\}} \quad (4)$$

$$A_k(t) = \sqrt{\frac{\alpha \ln t}{T_k(t)}}, \quad (5)$$

where  $\bar{X}_k$  is the sample mean of the channel  $k$  reward, and  $A_k$  is an upper confidence bias.  $\mathbb{1}$  is the indicator function and  $a_t$  is the selected channel using the policy  $\pi$  at the  $t^{\text{th}}$  transmission. Therefore, we have  $\mathbb{1}_{\{a_l=k\}} = 1$  only if the channel  $k$  has been chosen at instant  $l$ . The factor  $\alpha$  in (5) is an exploration coefficient for channel availability and coverage. If  $\alpha$  gets larger, the UCB algorithm will explore more channels for better

Table I: Allowed LTE PRB indices for cell connection in NB-IoT in-band deployment.

LTE bandwidth	3 MHz	5 MHz	10 MHz	15 MHz	20 MHz
LTE PRB indices for NB-IoT	2, 12	2, 7, 17, 22	4, 9, 14, 19, 30, 35, 40, 45	2, 7, 12, 17, 22, 27, 32, 42, 47, 52, 57, 62, 67, 72	4, 9, 14, 19, 24, 29, 34, 39, 44, 55, 60, 65, 70, 75, 80, 85, 90, 95

vacancy and coverage. Otherwise, when  $\alpha$  takes lower values, the exploitation is privileged.

The selected channel  $a_t$  resulting from UCB algorithm is the one with the highest UCB index, i.e.

$$a_t = \arg \max_k (B_k(t)). \quad (6)$$

#### IV. SIMULATION RESULTS

##### A. NB-IoT Scenario

NB-IoT standard designed by 3GPP is a slotted protocol [9]. We assume that the different communications using the studied spectrum are slotted i.e. all devices share synchronized time. We also assume that the end device knows in advance the finite number of dedicated radio frequency (RF) channels.

The NB-IoT supports the following deployment modes:

- In-band mode where NB-IoT is deployed within the LTE bandwidth, and occupies one or multiple physical resource blocks (180 kHz).
- Guard-band mode where NB-IoT operates within the guard-band of an LTE carrier
- Stand-alone mode where NB-IoT can either occupies one or more GSM carrier (200 kHz), or it can be deployed in an adjacent band to LTE.

It is worth mentioning that in the in-band deployment, the channels supported by NB-IoT are well-defined to avoid interfering with resources used by the LTE system such as synchronization, broadcast and control channels. The allowed physical resource blocks (PRB) are provided in Table I. We assume in our scenario that the NB-IoT is deployed in in-band mode associated to LTE system bandwidth of 15 MHz. 14 physical channels are then allowed for cell connection as given by Table I.

When the NB-IoT end device turns on, it tries to connect to the cell using the channel that maximizes its reward:

- If the device receives a feedback from the base station, it will update the reward (1 if good coverage level, 0 otherwise) and proceeds to the data transmission.
- Otherwise, the reward of the selected channel is updated to 0, and the device waits till the beginning of the next slot to select a new channel.

We define  $\mu_{\text{vac}}^k$  the expected mean reward associated to the vacancy of the PRBs  $k$ , and  $\mu_{\text{cov}}^k$  the expected mean reward

Table II: NB-IoT channel parameters for an LTE Bandwidth of 15 MHz when in-band mode is deployed.

PRB index	Channel index	$\mu_{\text{vac}}^k$ [%]	$\mu_{\text{cov}}^k$ [%]	$\mu_{\text{vac,cov}}^k = \mu_{\text{vac}}^k \cdot \mu_{\text{cov}}^k$
2	1	10	50	0.5
7	2	20	90	0.18
12	3	25	55	0.137
17	4	30	80	0.24
22	5	35	75	0.262
27	6	40	60	0.24
32	7	45	40	0.18
42	8	50	70	0.35
47	9	55	65	0.357
52	10	60	35	0.21
57	11	65	50	0.325
62	12	70	40	0.28
67	13	80	85	0.68
72	14	90	20	0.18

related to the coverage level of the PRBs  $k$ . Table II shows the values used in our learning policy, which are defined in a random way. Without loss of generality, we assume  $\mu_{\text{vac}}^1 \leq \mu_{\text{vac}}^2 \dots \mu_{\text{vac}}^{14}$ . Note that  $\mu_{\text{vac}}^1 = 10\%$  means that channel 1 is available 10% of the time, and  $\mu_{\text{cov}}^1 = 50\%$  means that channel 1 has a probability of 50% to be in a good coverage when it is free.  $\mu_{\text{vac,cov}}^1 = 0.05$  is associated to the probability that channel 1 is both available *and* in a good coverage.

#### B. Best channel selection and cumulative regret

In addition to the cumulative regret defined in Section III-A, the percentage of optimal channel selection is a relevant metric in the analysis of reinforcement learning policies. Since the device should transmit in the optimal channel i.e. that has the highest mean reward, the more is the percentage of optimal channel selection, the better is the policy. Both of these relevant metrics are compared in Fig. 3, Fig. 4, and Fig. 2 for the following proposed UCB scenarios:

- UCB(vac) means that the policy consider only the vacancy distributions  $\mu_{\text{vac}}^k$  and assumes that the different channels have the same coverage properties. The best channel in this case is  $14 = \arg \max_k (\mu_{\text{vac}}^k)$ . This scenario is extremely unlikely to happen, since the coverage level of each RF channel is independent from the probability of its availability.
- UCB(vac,cov) takes into account both the vacancy and the coverage criteria. The policy is then based on the distribution  $\mu_{\text{vac,cov}}^k$ . The best channel in this case is different from UCB(vac) case, and it corresponds to  $13 = \arg \max_k (\mu_{\text{vac,cov}}^k)$ .
- UCB(vac, $\overline{\text{cov}}$ ): In this scenario, the channels do not have the same coverage quality but follow the distributions  $\mu_{\text{cov}}^k$ . However, the end device does not take the coverage criterion in the calculation of the UCB index <sup>1</sup>, its calculation is based only on the vacancy distributions  $\mu_{\text{vac}}^k$ . The

best channel here is 13 since the channels have different coverage and vacancy properties.

The exploration coefficient  $\alpha$  is set to 1.5 for all the studied scenarios.

Fig. 1 and Fig. 2 show the impact of the coverage constraint on the UCB policies. The major finding resulting from comparing the UCB(vac, $\overline{\text{cov}}$ ) policy where the coverage criterion is not considered, and the UCB(vac,cov) policy where the choice of the RF channel is constrained with its coverage quality, is that the former policy does not converge to the optimal channel unlike the latter. This result was expected from the behaviour and the target of each policy. During the first iterations, both algorithms start exploring the physical channels with different objectives: UCB(vac, $\overline{\text{cov}}$ ) calculates its UCB index based on the mean reward of vacancy  $\mu_{\text{vac}}^k$ , while UCB(vac,cov) takes into account both availability and coverage i.e.  $\mu_{\text{vac,cov}}^k$  in its UCB index calculation. In the long run, UCB(vac, $\overline{\text{cov}}$ ) tends to choose the channel number 14, which is the most available one, while channel number 13 is the optimal one since it allows both a good availability and a good coverage level. The channel with the highest probability to be free is not necessarily the one with the best coverage. Therefore, the two compared policies behave in a symmetric manner, the best channel selection percentage decreases with time for UCB(vac, $\overline{\text{cov}}$ ) and tends to 0% while it converges to 100% for UCB(vac,cov) (see Fig. 1). The resulting cumulative regret increases linearly with time for UCB(vac, $\overline{\text{cov}}$ ) and seems insignificant for UCB(vac,cov) compared with the first policy (see Fig. 2). Note that in some other scenarios, the most available channel might be also the one with the best coverage level. Assuming this scenario, even if it is highly unlikely to occur, the policy UCB(vac, $\overline{\text{cov}}$ ) would not behave better than the proposed policy UCB(vac,cov).

We compare in Fig. 3 and Fig. 4 the previously analysed UCB(vac,cov) policy, with UCB(vac) scenario. It can be noticed that UCB(vac,cov) still gives the best performance. UCB(vac,cov) converges faster than UCB(vac) and produces the lowest cumulative regret. As an example, if we consider the transmission i.e iteration number  $t = 10^3$ , the best channel selection associated to UCB(vac,cov) reaches 56%, while it is equal to 43% for UCB(vac) which is less than half the percentage achieved by UCB(vac,cov) (16%). Regarding the cumulative regret, Fig. 4 reflects in fact the same performance behaviour. In order to give a numerical insight, let us consider another interval of the data transmission:  $t = 10^5$ . The cumulative regret corresponding to UCB(vac,cov) is only 492, while it reaches 615 for UCB(vac).

Note that for the different scenarios, their behaviour during the first iterations is not regular. This is due to the random selection of the channels during the first round since no knowledge about the vacancy nor the coverage is available during the beginning of the iterations.

#### C. Coverage extension and reducing repetitions

In the previous section, we have strongly supported that integrating the coverage level constraint into the UCB policy improves the best channel selection percentage and decreases

<sup>1</sup>This is what we mean by the notation  $\overline{\text{cov}}$ .

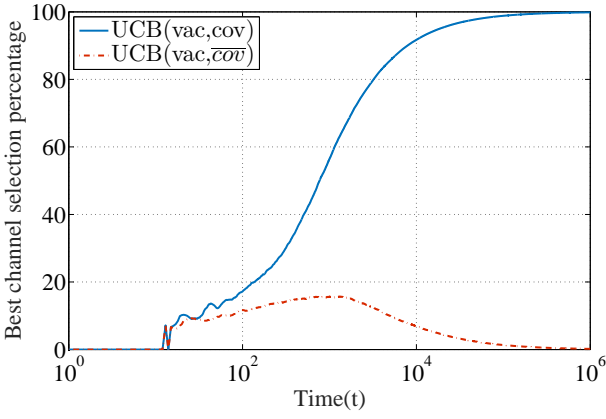


Figure 1: Impact of considering coverage on best channel selection percentage.

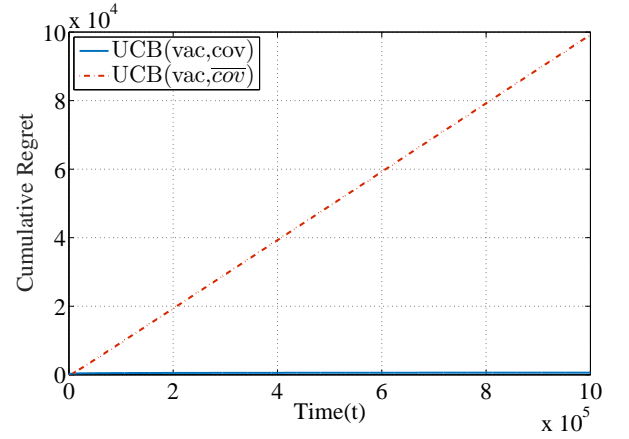


Figure 2: Impact of considering coverage on cumulative regret.

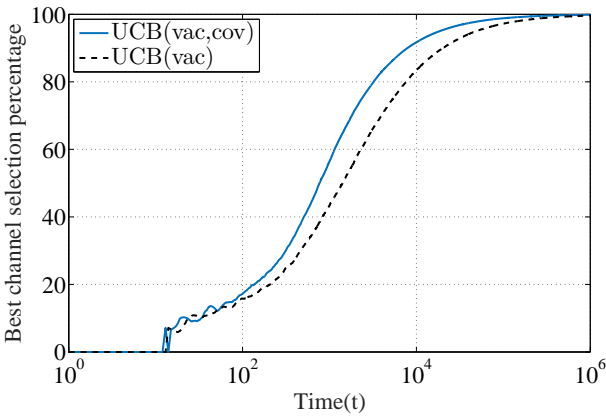


Figure 3: Best channel selection percentage for different scenarios.

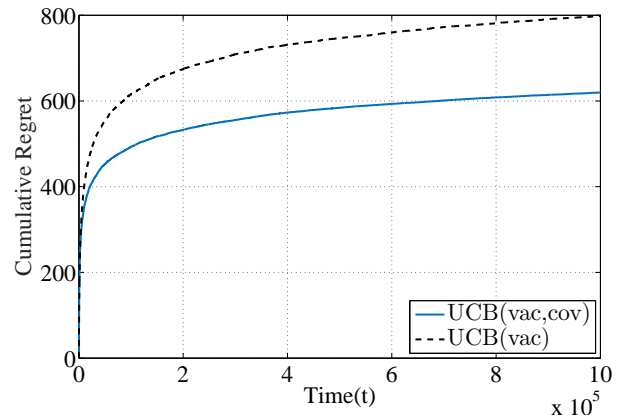


Figure 4: Cumulative regret for different scenarios.

the cumulative regret. Regarding the studied scenarios that are associated to the distributed values reported in Table II, using UCB(vac, cov) would rather choose a RF channel (channel number 13 in the provided example) with a good coverage that allows reaching easily the destination, than a RF channel (channel number 14) with an extreme coverage that makes reaching the receiver pretty hard. Moreover, UCB(vac, cov) allows avoiding transmission repetitions by selecting a good quality channel that requires no or few number of repetitions.

When using a random selection procedure, the channel to transmit is chosen in a random way, ignoring its probability of availability or its coverage properties. In the literature, several researchers have been defending the interest of dynamic spectrum access using the vacancy criterion of the channels. Here, we support the relevance of involving the quality of the channels in terms of coverage level in UCB policies, especially for applications that need a long battery lifetime such as NB-IoT. To be convinced of this, let us consider the best and worse case scenarios for the proposed policy UCB(vac, cov):

- Best case scenario for implementing UCB proposed poli-

cies: the selected channel by a random selection procedure is the one with the highest required number of repetitions 128 (in an extreme coverage), while the optimal channel in UCB policy requires only 1 (or 0) repetition. This scenario would reduce the latency and save significant amount of energy by avoiding large number of retransmissions.

- Worse case scenario: the randomly selected channel matches the one with the best quality of service. This means that our policy for this typical transmission time, would not save a priori more energy than legacy selection schemes.

Hence, our policy promises in the general case a considerable improvement of the quality of the transmission, and meets then the challenges required by NB-IoT applications.

## V. CONCLUSION

The emergence of new IoT applications such as NB-IoT requires to fulfill and meet several challenges. Reducing the energy consumption is one of the major features of NB-IoT end devices. This target is inherently correlated to enhancing the

coverage, since the fundamental adopted solution is increasing the number of retransmissions, and thereby consuming more energy. In order to enhance the coverage and to reduce the energy consumption, this paper proposes a new solution based on dynamic spectrum using machine learning algorithms. The random selection procedure is replaced by a more efficient selection method that chooses the channels with the highest probability to be available, and with the best coverage and the lowest number of required repetitions. The next step of our work targets the investigation of a multi-player scenario, where several NB-IoT end devices need to access dynamically the network.

#### ACKNOWLEDGMENT

This work has received a French state support granted to the *Enhanced PHY for cellular Low power communication IoT (EPHYL)* project and managed by the National Research Agency under reference Nb. ANR-16-CE25-0002-03.

#### REFERENCES

- [1] 3GPP TR 45.820 v13.1.0. Cellular system support for ultra low complexity and low throughput internet of things. *[Online]*, November 2015.
- [2] N Sornin, M Luis, T Eirich, T Kramp, and O Hersent. LoRa Alliance LoRaWAN specification. *LoRaWAN Specification, Release v1. 0*, 2015.
- [3] Mehdi Anteur, Vincent Deslandes, Nathalie Thomas, and Andre-Luc Beylot. Ultra narrow band technique for low power wide area communications. In *Global Communications Conference (GLOBECOM), 2015 IEEE*, pages 1–6. IEEE, 2015.
- [4] W. Jouini, D. Ernst, C. Moy, and J. Palicot. Upper confidence bound based decision making strategies and dynamic spectrum access. In *2010 IEEE International Conference on Communications*, pages 1–5, May 2010.
- [5] L. Melian-Gutierrez, N. Modi, C. Moy, I. Perez-Alvarez, F. Bader, and S. Zazo. Upper Confidence Bound learning approach for real HF measurements. In *2015 IEEE International Conference on Communication Workshop (ICCW)*, pages 381–386, June 2015.
- [6] Peter Auer. Using confidence bounds for exploitation-exploration trade-offs. *Journal of Machine Learning Research*, 3(Nov):397–422, 2002.
- [7] Vidit et al. Saxena. Reducing the Modem Complexity and Achieving Deep Coverage in LTE for Machine-Type Communications. In *Global Communications Conference (GLOBECOM), 2016 IEEE*, pages 1–7. IEEE, 2016.
- [8] Wassim Jouini, Christophe Moy, and Jacques Palicot. Decision making for cognitive radio equipment: analysis of the first 10 years of exploration. *Eurasip journal on wireless communications and networking*, 2012(1):26, 2012.
- [9] Rapeepat Ratasuk, Benny Vejlgaard, Nitin Mangalvedhe, and Amitava Ghosh. NB-IoT system for M2M communication. In *Wireless Communications and Networking Conference (WCNC), 2016 IEEE*, pages 1–5. IEEE, 2016.