# Sub-sampling of channels with time and frequency sparsity access

Yves LOUET[*], Vincent SAVAUX[+], Apostolos KOUNTOURIS[#] and Christophe MOY[*]

[*] *IETR/CentraleSupélec, Cesson-Sévigné, France, {first name.last name}@centralesupelec.fr*

[+] *b<>com, Rennes, France, Vincent.savaux@b-com.com*

[#] *Orange Labs, Grenoble, France, apostolos.kountouris@orange.com*

***Abstract*—This paper shows that sub-sampling of signals can help in reducing the amount of data to be processed and stored when time and frequency sparsity is considered. The context is the one of the *Internet of Things* (IoT) for which a huge quantity of users (i.e. objects) communicate with very few time and frequency accesses. Taking advantage of the occupancy theory of probabilities, we propose a theoretical model for such communications and we show that the sampling frequency of signals can be significantly reduced under these assumptions.**

***Keywords*-Sub-Nyquist sampling, sparsity, Multi-channel access, Occupancy problem, Internet of things**

## I. INTRODUCTION

A huge quantity of objects will be soon connected all together to a network namely the Internet of Things (IoT) and will be able to collect and exchange data of any kind and anywhere. Experts estimate that the IoT will consist of almost 50 billion objects by 2020 [1]. As a result, the amount of data to be collected to a central base station (BS) and stored (depending upon applications) will increase exponentially and ways to mitigate this volume are more than expected.

From the channel access side the new context of IoT makes the data transmissions asynchronous (an object can communicate independently of any clock reference) and sparse (i.e. with a few time and frequency use). This new paradigm has two consequences: (i) collisions may occur between signals with an increase of the interference level (ii) the sparse use of time and frequency resources allows sub-sampling techniques to decrease the amount of data.

Signals collisions is a well-known issue in wired and wireless networks where different independent users share the same resources.We assume in this paper that a BS is only a scanning device and therefore the users (objects) do not get any feedback in case their packets collided or not. In [5] we developed a channel model for IoT communications including time asynchronism and how it influences the overall network capacity.

Sub-sampling techniques have gained interest with cognitive radio the past few years. The motivation is mainly to gain in consumption of analog to digital convertors especially when the considered bandwidth is not full [6]. The under use of channels (in frequency) makes possible the mitigation of the sampling frequency in between Shannon upper bound (twice the bandwidth) and Landau lower bound (depending upon the sparsity of the bandwidth) with perfect reconstruction. Many techniques have been proposed to under select the samples (uniform, non uniform, random) [7]. The purpose of this paper is not to propose a sub-sampling technique but to show that sub-sampling with almost perfect reconstruction is possible when considering both time and frequency access sparsity.

In this paper we analyze the time-frequency use of the band by means of the *occupancy problem* [8] exemplified by the situation where $m$ balls are thrown randomly to fill $n$ bins. This model has been previously applied in [5] to analyze the throughput capacity of the associated IoT network. We apply this *occupancy problem* to show that the sampling frequency can be decreased considering random access and sparsity both in time and frequency. Note that in current applications IoT communications are not channelized due to the low costs of devices. However the proposed study provides a first model from which one can draw some key conclusions.

The remaining of the paper is organized as follows: Section II presents the system model. In Section III some useful results using the *occupancy problem* are provided and Section IV covers the sub-sampling proposal of the paper. Finally, Section V concludes this paper.

## II. SYSTEM MODEL

The network setting considered in the following: $N_u$ users are randomly and homogeneously distributed in a cell in the center of which a BS is located. The users may operate in different ways depending on the nature of the deployed applications (like IoT objects for instance). Thus, users in the area $C_i$ may exchange data with others in area $C_i$ and/or transmit information to the BS independently of users in another area $C_j$. Moreover, it is assumed that they all share the same frequency band $\mathcal{B}$ of width $B$.

The role of the BS consists in scanning and sampling the band $\mathcal{B}$ containing the signals. The bandwidth $B$ to be scanned is regularly subdivided into $N_c$ channels $\{\mathcal{B}_0, \mathcal{B}_1, .., \mathcal{B}_{N_c-1}\}$ of width $B/N_c$. The considered random multi-channel multi-user transmission model can be formalized as follows:

- Each user can randomly access one of the channels with probability $\mathbb{P}_a = \frac{1}{N_c}$. In that way, many users may select the same channel i.e. a "reuse" of one or more channels may occur. In the following, we denote by $R$ the channel reuse factor, i.e. the number of users who simultaneously

share the same channel. For instance, $R = 0$ and $R = 2$ correspond to the cases where no user and two users share the same channel, respectively.
- An asynchronous slotted time traffic is considered in each channel at the BS, due to the independence between users. The probability of transmission of a packet (or duty cycle[1]) is denoted by $q \in [0, 1]$, and for simplicity, it is supposed that $q$ is the same for each user. Collisions may occur between packets when at least two users transmit on the same channel.

The proposed transmission model can be analyzed and characterized by means of combinatorial probability as presented hereafter.

## III. Occupancy Problem

The random access to $N_c$ channels by $N_u$ users is known as *occupancy problem* in probability. This theory provides useful tools to deal with the time-frequency use of the band $\mathcal{B}$ in the considered transmission model. In particular, two theorems will be used in this paper.

### A. Formulation of the Problem

**Theorem 1.** *Let $N_u$ users randomly accessing $N_c$ channels with probability $\mathbb{P}_a = \frac{1}{N_c}$. Then the probability that b channels are used (at least by one user) is*

$$\mathbb{P}(b) = \binom{N_c}{N_c - b} \sum_{\nu=0}^{b} (-1)^{\nu} \binom{b}{\nu} \left(1 - \frac{N_c - b + \nu}{N_c}\right)^{N_u}. \quad (1)$$

In the following, we denote by $\Theta_{N_u,b}$ the probability mass function corresponding to $\mathbb{P}(b)$ in (1).

*Proof:* see [8], Chapter 4. ∎

**Theorem 2.** *Under the same assumptions as in Theorem 1, the probability that $R$ users choose the same channel ($0 \leq R \leq N_u$) is:*

$$\mathbb{P}(R) = \binom{N_u}{R} \left(\frac{1}{N_c}\right)^{R} \left(1 - \frac{1}{N_c}\right)^{N_u - R}. \quad (2)$$

*Proof:* Let $\mathbb{P}(j)$ the probability of the event "the $j$-th channel is chosen", and $\mathbb{P}(X_j = R)$ the probability that $R$ users use the $j$-th channel. As $\mathbb{P}(j) = \frac{1}{N_c}$ and $\mathbb{P}(X_j = R)$ are equal for any $j$, then we obtain

$$\mathbb{P}(R) = \sum_{j=1}^{N_c} \mathbb{P}(j)\mathbb{P}(X_j = R) = \mathbb{P}(X_j = R).$$

The reuse factor $R$ can be defined as the sum $X_j = R = \sum_k x_{j,k}$, where $x_{j,k}$ is a random variable which counts the number of users in the channel $j$. Since users access the channel independently of each other, $x_{j,k}$ are Bernoulli trial outcomes with probability $1/N_c$ to access the channel $j$. Therefore, the sum $X_j = \sum_k x_{j,k}$ has the binomial distribution defined in (2). ∎

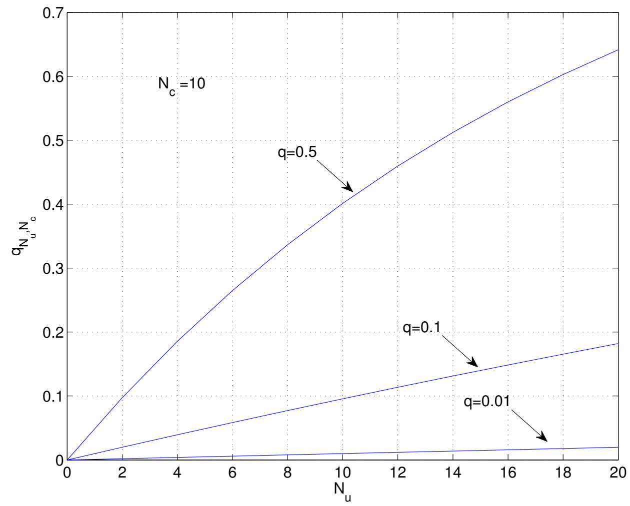


Fig. 1. Probability of occupancy $q_{N_u,N_c}$ versus $N_u$, using $N_c = 10$.

The *occupancy problem* provides the probabilities of the frequency use of the band $\mathcal{B}$: Theorems 1 gives the probability that $b$ among $N_c$ channels have been chosen at least by one user, whereas Theorems 2 gives the probability for a given channel to be chosen by $R$ users.

### B. Time-Frequency Use of the Band $\mathcal{B}$

It is assumed that users transmit packets independently of each others with the same duty cycle $q$. Therefore, the probability $q_R$ that a channel is occupied (in the time domain) when it is accessed by $R$ users can be expressed as $q_R = 1-(1-q)^R$ where $(1-q)^R$ is the probability that the channel is free. We denote by $q_{N_u,N_c}$ the overall probability of occupancy of the band $\mathcal{B}$ when $N_u$ users transmit in $N_c$ channels with random time and frequency channel access. The probability $q_{N_u,N_c}$ can be defined as the following weighted sum:

$$q_{N_u,N_c} = \sum_{R=0}^{N_u} \mathbb{P}(R) q_R. \quad (3)$$

Fig. 1 depicts the probability $q_{N_u,N_c}$ in (3), and gives the degree of time-frequency sparsity of the band $\mathcal{B}$. It can be observed that $q_{N_u,N_c} << 1$, in particular for low duty cycle values. These results allow us to deduce that the band $\mathcal{B}$ can be sampled at a sub-Nyquist frequency rate.

In the next section, we derive from (1) and (2) an upper bound of the achievable sampling rate[2] with respect to the sparsity of the channels.

## IV. Sub sampling analysis

### A. General Rule for defining the Sub-Nyquist Sampling $f_s$

In this section, we show that it is possible to take advantage of the time-frequency sparsity of the received signal in order

[1]Sometimes the term *duty cycle* is used to refer to the overlapping factor between two packets, as in [9].

[2]the minimum sampling rate that can be achieved for a specific channel occupation pattern without loss of information

to reduce the sampling frequency at the BTS, and therefore the corresponding amount of data to be tranfered to a storage server. Performing the sub Nyquist sampling (SNS) and the reconstruction is beyond the scope of this paper (see [6], [7] for details on uniform and non uniform sampling for instance). Nevertheless, in this work our aim is to provide a condition that guarantees a sampling frequency $f_s$ less than the Nyquist upper bound with probability $\mathbb{P}_L$.

In this paper, we rather assess an achievable sub-Nyquist sampling frequency according to the following principle:

- we denote by $f_s \in [0, f_{Ny}]$ the sampling frequency for the band $B$, with $f_{Ny} = 2B$. Alternatively, we can write $f_s = \nu_b f_{Ny}$ with $\nu_b \in [0, 1]$, or more precisely, $\nu_b \in \{0, \frac{1}{N_c}, .., \frac{b}{N_c}, .., 1\}$.
- According to: i) the time-frequency sparsity of the received signals at the BTS, ii) the random time-frequency channel access by the user nodes, it can be deduced that for any $f_s \in [0, f_{Ny}]$ corresponds the event $L$="$f_s$ is large enough to properly sample the band $B$" with probability $\mathbb{P}_{L,b}$ (the subscript $b$ indicates that this probability depends on $b$). For instance, whatever the sparsity of the signals is (depending on $N_c$, $N_u$, and $q$), we have $\mathbb{P}_{L,b} = 1$ if $f_s = f_{Ny}$ (it is sure that the multi-user signals are properly sampled), and $\mathbb{P}_{L,b} = 0$ if $f_s = 0$ (it is sure that the multi-user signals are not properly sampled).
- In order to define $f_s$, we use the following rule: set a given threshold $\gamma \in [0, 1]$ (actually $\gamma$ should be set close to 1) such that $f_s = \nu_b f_{Ny}$ where $b$ is chosen according to $\mathbb{P}_{L,b-1} < \gamma$ and $\mathbb{P}_{L,b} \geq \gamma$.

In the following, we propose two ways to define $f_s$ first considering the frequency sparsity only and then both time and frequency sparsity.

### *B. Sampling frequency reduction according to the frequency sparsity*

We remind that the frequency sparsity of the signal can be defined as $\lambda(\mathcal{F})/f_{Nyq}$, where $\lambda(\mathcal{F})$ is called the Lebesgue measure. In our simple scenario, this frequency sparsity equals to $b/N_c$.

Considering frequency sparsity is equivalent to the case for which all the users transmit with a duty cycle $q = 1$. In this scenario, the probability $\mathbb{P}_{L,b}$ is defined as the cumulative distribution function $P_{L,b} = \sum_{k=1}^{b} \Theta_{N_u,k}$ where $\Theta_{N_u,k}$ is the corresponding probability mass function defined in (1). Note that since only frequency sparsity of the signals is considered, the sampling frequency $f_s$ can be seen as an upper bound of $f_s$ that can be achieved in practice. In order to analyze the achievable upper bound of sampling frequency $f_s$, we define $D$ as the density of users with respect to the number of channels, namely $D = N_u/N_c$. Fig. 2 depicts the obtained $f_s$ given as a ratio of $f_{Ny} = 2B$ versus density $D$, using three values of threshold $\gamma \in \{0.8, 0.9, 0.99\}$. The case $\gamma = 0.8$ relaxes the constraint on $f_s$, i.e. potential aliasing phenomenon is allowed, whereas $\gamma = 0.99$ corresponds to a case where almost no aliasing is expected. Fig. 2 shows the achieved results for $N_c = 10$.

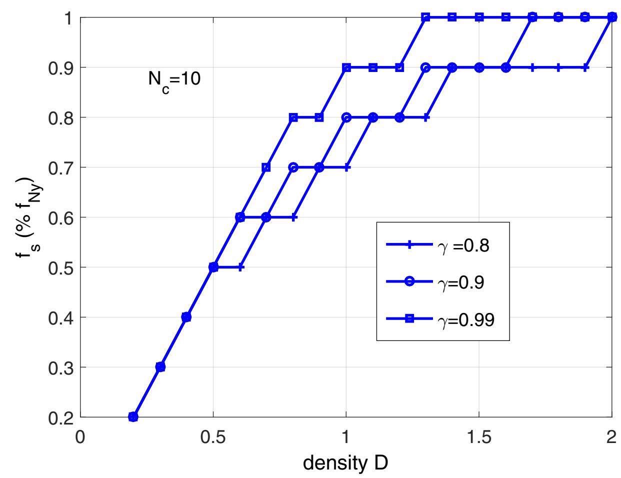


Fig. 2. Achieved $f_s$ (given as a percentage of $f_{Ny} = 2B$) for $N_c = 10$

It can be observed on Fig. 2 that $f_s$ values increases with the density, and the lower $\gamma$ the lower $f_s$. More generally, it is worth noting that a large reduction of the average sampling frequency can be achieved at $D \lessapprox 1.2$, e.g. up to 30% at $D = 1$. This interesting results could still be improved if both frequency and time sparsity is taken into account, as shown hereafter.

### *C. Sampling frequency reduction according to time and frequency sparsity*

The developments undertaken in this section require several new notations and definitions. The rigorous assessment of $f_s$ through the analysis of the sparsity of the signal in both time and frequency would necessitate the exhaustive search of all the patterns and their corresponding probabilities.

**Definition 1.** Pattern*: Let $N_u$ nodes randomly access $N_c$ channels in a given band $\mathcal{B}$. The term* pattern *points out the basic dispositions of the users in the channel, from which all other dispositions can be deduced by permutation. For instance, if $N_u = 4$ and $N_c = 3$, then the set $\Omega$ of the patterns is $\Omega = \{(4, 0, 0), (3, 1, 0), (2, 2, 0), (2, 1, 1)\}$, where the integers indicates the number of users per channel. In a more general way we denote by $\Omega = \{\omega_i\}$, with $1 \leq i \leq S$ and $S = card(\Omega)$, and the probability of the event $\omega_i$ is denoted by $\mathbb{P}_{\omega_i}$.*

*Furthermore, we define $x_{i,j}$ the elements of the subsets $\omega_i$ such as $\omega_i = (x_{i,1}, .., x_{i,j}, .., x_{i,N_c})$.*

To show the difficulty of assessing $f_s$ by using the exhaustive patterns $\omega_i$, we define the variables:

- $\pi_b$ corresponding to the probability that $b$ channels are occupied in the same time.
- $\xi_b$ the event "$b$ channel are used in the same time", and by $\mathbb{P}(\xi_b|k)$ the conditional probability of $\xi_b$ given that $k$ channels are accessed.

Since time and the frequency accesses to the channels are two independent events, the probability $\pi_b$ can be written as

$$\pi_b = \sum_{k=1}^{N_c} \Theta_{N_{ut},k} \mathbb{P}(\xi_k|k) = \sum_{k=b}^{N_c} \Theta_{N_{ut},k} \mathbb{P}(\xi_k|k), \quad (4)$$

where $\mathbb{P}(\xi_k|k)$ can be expressed as

$$\mathbb{P}(\xi_b|k) = \sum_i \mathbb{I}_{x,i}^{(k)} \mathbb{P}_{\omega_i} \frac{1}{L_{\omega_i}} \sum_{l_i=1}^{L_{\omega_i}} \underbrace{\prod_{j=1}^{b} q_{x_{i,j,l_i}}}_{\text{Proba. of presence}}$$
$$\times \underbrace{\prod_{j=b+1}^{k} (1 - q_{x_{i,j,l_i}})}_{\text{Proba. of absence}}, \quad (5)$$

where $\mathbb{I}_{x,i}^{(k)}$ is the indicator function of the subsets $\omega_i$ containing exactly $k$ non-null elements $x_{i,j}$, $L_{\omega_i}$ is the number of possible permutations between non-null elements $x_{i,j}$ of $\omega_i$ ($l_i = 1$ corresponds to the pattern), and the two products in (5) are the probability of simultaneous presence and absence of signals. It appears that the assessment of the probability $\mathbb{P}_{\omega_i}$ is very difficult. Therefore we propose to derive an upper bound of $\mathbb{P}(\xi_k|k)$ thanks to some simplifications.

The major simplification consists of considering a single *virtual pattern* for each value $b$ whose definition is given by:

**Definition 2.** Virtual pattern*: Let $N_u$ nodes randomly access $N_c$ channel in a given band $\mathcal{B}$. The term* virtual pattern *points out the configurations of the users in the channel as*

$$\tilde{\Omega} = \{(\tilde{x}_1, 0, ..), .., (\tilde{x}_b, \tilde{x}_b, .., \tilde{x}_b, 0, ..), .., (\tilde{x}_{N_c}, \tilde{x}_{N_c}, .., \tilde{x}_{N_c})\},$$

*where $\tilde{x}_b = \frac{N_u}{b}$. It is worth mentioning that as $\tilde{x}_b$ may be not an integer, it has no physical sense, but it allows us to simplify (5) in a mathematical sense. Furthermore, the subsets $\tilde{\omega}_b$ are invariant by permutation of their non-null elements $\tilde{x}_b$.*

The use of the virtual pattern allows to remove the probability $\mathbb{I}_{x,k}\mathbb{P}_{\omega_i}$ in (5) since there is a single pattern $\tilde{\omega}_k$ for any value $k$. The conditional probability $\tilde{\mathbb{P}}(\xi_k|k)$ can then be rewritten as $\mathbb{P}(\xi_b|k) = q_{\tilde{x}_k}^b (1 - q_{\tilde{x}_k})^{k-b}$ and finally, the probability that $b$ channels are occupied in the same time can be estimated as

$$\tilde{\pi}_b = \sum_{k=b}^{N_c} \Theta_{N_{ut},k} q_{\tilde{x}_k}^b (1 - q_{\tilde{x}_k})^{k-b}. \quad (6)$$

Note that $\tilde{\pi}_b$ in (6) corresponds to $\mathbb{P}_{L,b}$ and therefore the achievable $f_s$ can be assessed by using $\tilde{\pi}_b$.

Fig. 3 shows the sampling frequency $f_s$ (given as a percentage of $f_{Ny} = 2B$) versus $N_u$, assessed with (6). Compared to results in Fig. 2, it can be observed in Fig. 3 that the achievable $f_s$ is greatly reduced when the time sparsity (highlighted by the parameter $q$) is considered in addition to the frequency selectivity. In fact, for low duty cycle $q = 0.01$, a reduction by half of the sampling frequency can be done for $D = N_u/N_c = 2$. This shows that sub-Nyquist sampling can be carried out, even if numerous nodes are located in the cells.

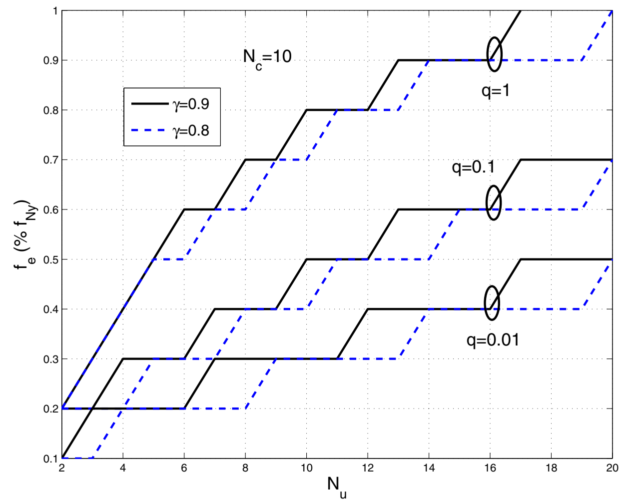


Fig. 3. Achieved $f_s$ (given as a percentage of $f_{Ny} = 2B$) versus number of users $N_u$ for $N_c = 10$.

## V. Conclusion

In this paper we modeled the multi-channel multi-user access network for IoT where a number of independent users (objects) access the spectrum band randomly both in time and frequency with sparsity. Thanks to this formulation the achievable reduction of the sampling rate have been assessed. From a practical standpoint it is very helpful to quantify the sampling rate reduction if spectrum samples need to be transfered and/or stored and processed in some cloud-based IoT network infrastructure.

## References


[1] J. A. Stankovic, "Research directions for the internet of things," *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 3–9, February 2014.

[2] P. Nardelli, M. Kaynia, P. Cardieri, and M. Latva-aho, "Optimal Transmission Capacity of Ad Hoc Networks with Packet Retransmissions," *IEEE Transactions on Wireless Communications*, vol. 11, no. 8, pp. 1536 – 1276, August 2012.

[3] G. Thomas, "Capacity of the Wireless Packet Collision Channel Without Feedback," *IEEE Transactions on Information Theory*, vol. 46, no. 3, pp. 1141 – 1444, May 2000.

[4] M.-T. Do, C. Goursaud, and J.-M. Gorce, "Interference Modelling and Analysis of Random FDMA scheme in Ultra Narrowband Networks," in *proc. of AICT'14*, Paris, France, July 2014.

[5] V. Savaux, A. Kountouris, Y. Louet, and C. Moy, "Throughput capacity analysis of a random multi-user multi-channel network modeled with occupancy problem," in *11th EAI International Conference on Cognitive Radio Oriented Wireless Networks*, May 2016.

[6] M. Mishali and E. Y. C., "Sub-Nyquist Sampling ," *IEEE Signal Processing Magazine*, vol. 28, no. 6, pp. 98 – 124, November 2011.

[7] R. Venkataramani and Y. Bresler, "Optimal Sub-Nyquist Nonuniform Sampling and Reconstruction for Multiband Signals," *IEEE Transactions on Signal Processing*, vol. 49, no. 10, pp. 2301 – 2313, October 2001.

[8] W. Feller, *An Introduction to Probability Theory and its Applications*, 3rd ed. John Wiley & Sons, 1968, vol. 1.

[9] M. Z. Win, P. C. Pinto, and L. A. Shepp, "AMathematical Theory of Network Interference and Its Applications," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 205 – 230, February 2009.