Multiband Massive IoT: A Learning Approach to Infrastructure Deployment

Enes Krijestorac, *Student Member, IEEE*, Ghaith Hattab, *Member, IEEE*, Petar Popovski, *Fellow, IEEE*, and Danijela Cabric, *Fellow, IEEE*

Abstract—We consider a novel ultra-narrowband (UNB) low-power wide-area network (LPWAN) architecture design for uplink transmission of a massive number of Internet of Things (IoT) devices over multiple multiplexing bands. An IoT device can randomly choose any of the multiplexing bands to transmit its packet. Due to hardware constraints, a base station (BS) is able to listen to only one multiplexing band. Our main objective is to maximize the packet decoding probability (PDP) by optimizing the placement of the BSs and frequency assignment of BSs to multiplexing bands. We develop two online approaches that adapt to the environment based on the statistics of (un)successful packets at the BSs. The first approach is based on a predefined model of the environment, while the second approach is measurement-based model-free approach, which is applicable to any environment. The benefit of the model-based approach is a lower training complexity, at the risk of a poor fit in a model-incompatible environment. The simulation results show that our proposed approaches to band assignment and BS placement offer significant improvement in PDP over baseline random approaches and perform closely to the theoretical upper bound.

Index Terms—LPWA networks, UNB, IoT, channel assignment, BS placement

I. Introduction

THE Internet of Things (IoT) has the potential to change the technological landscape and bring great economical and societal benefits. The success of IoT on a large scale depends on several key enabling technologies, one of which is wireless communication. Low-power wide-area (LPWA) networks are a new paradigm of wireless networking that is expected to become one of the key drivers of massive IoT [1]. Compared to the legacy technologies, such as cellular and short-range wireless networks, the benefits offered by LPWA networks include wide-area connectivity for low-power and low-data-rate devices and low capital expenditure due to the use of the unlicensed spectrum. To enable long-range connectivity, LPWA networks primarily use sub-1GHz bands due to the favorable propagation conditions.

Enes Krijestorac and Danijela Cabric are with the Electrical and Computer Engineering Department, University of California, Los Angeles, CA 90095, USA. e-mail: enesk@g.ucla.edu, danijela@ee.ucla.edu.

Ghaith Hattab was with the Department of Electrical and Computer Engineering, University of California at Los Angeles, Los Angeles, CA 90095-1594 USA. He is now with Apple Inc., Cupertino, CA 95014 USA (e-mail: ghattab@ucla.edu).

Petar Popovski is with the Department of Electronic Systems, Aalborg University, 9220 Aalborg, Denmark (e-mail: petarp@es.aau.dk).

Ultra-narrowband (UNB) LPWA solutions apply the ultra-narrowband transmissions, which enable demodulation at a very low received power. Furthermore, UNB LPWA networks normally rely on simple ALOHA-like access protocols, where IoT devices avoid associating and synchronizing with any UNB base station (BS); in essence, IoT devices operate in a broadcast mode, transmitting their packets at arbitrary time and frequency. Normally, the BSs operate in a decentralized manner and packet decoding occurs at the BSs, as opposed to at a central server where the received symbols from all BSs in the network could be combined. Therefore, the packet is successfully transmitted if any BS decodes any of the packet transmissions.

Due to the extremely low bandwidth (hundreds of Hz), the frequency drift of the local oscillator in the commodity hardware becomes comparable to the transmission bandwidth, thus rendering slotted access infeasible. UNB networks solve this problem by allowing the devices to transmit in an unslotted manner, while the receiving BSs do the task of accurately syncing on to a signal in frequency. This is accomplished by performing fast Fourier transform (FFT) at a very fine sampling interval over the entire bandwidth of the multiplexing band, which is the portion of the spectrum across which UNB transmissions occur. Naturally, a UNB network would benefit from a wider multiplexing band. However, since the complexity of the FFT scales with the bandwidth of the multiplexing band, the multiplexing band has a feasibility limit on its bandwidth. One way to introduce more frequency diversity would be to use several multiplexing bands with each BS associated to one band and IoT devices transmitting freely across any band. This would keep the capital expenditure the same since the BSs would still use the same hardware, only tuned to different multiplexing bands, while the capacity of the network could potentially increase. Indeed, in [2], [3] it has been shown that the capacity does increase by applying this paradigm at no additional cost. However, the scope of [2], [3] does not consider optimization of the BS infrastructure, leaving room for further enhancements.

Optimizing the deployment of the BS infrastructure, namely the BS placement and frequency band assignment, is precisely the focus of this paper. The BS placement and channel assignment are jointly optimized, however, the approaches that we will propose naturally extend to optimizing channel assignment without placement and vice versa. We have developed an approach for joint optimization of BS placement and band

assignment, however, this approach can be applied to solve the problem of band assignment only, when the placement is predetermined. It can also be applied to solving for placement only, when there is a single band. Moreover, we assume that some BSs may be present in the environment already and the network is expanded by introducing additional BSs. In that case, the optimization will include channel assignment for the current BSs as well as both channel assignment and placement for the new BSs. The objective is to maximize the probability of decoding of a packet (PDP), irrespective of the source IoT device.

We cast the problem as an integer non-linear problem (INLP), for which we then find an approximate solution. In order to implement this solution, we are faced with learning the *average decoding probability (ADP)* at several locations in the network and the *joint decoding probability (JDP)* between several pairs of locations in the IoT network. JDP between a pair of locations is defined as the probability of successfully decoding a transmission from an IoT device at both locations in the pair. The model-based approach requires fewer measurements, however, it is based on stochastic models of device placement, channel, and interference, which the environment should match. The measurement-based approach directly estimates the required JDPs and ADPs. It involves no assumptions on device placement, channel, and interference at the cost of a more demanding training phase. Due to its training complexity, this approach is more suitable for optimizing only channel assignment when the placement is not optimized, due to, for example, infrastructure constraints that decide the BS placement. In both approaches, we do not assume to know the locations of IoT devices, rather, our optimization is based on the information collected by the BSs that are receiving the packets.

Our main contributions can be summarized as follows:

- We formulate joint BS placement and frequency assignment for maximizing PDP as an optimization problem and derive an approximate solution, which we show to require the estimation of ADP and JDP at several locations in the network.
- We propose a model-based approach for predicting ADP and the JDP at locations for which we do not have any measurements. First, we use stochastic geometry to derive models of ADP and JDP as a function of the location. We then use the derived models to develop an algorithm for prediction of ADP and JDP for any location in the network.
- We develop an algorithm for efficient direct learning of ADP and JDP for candidate locations for the BSs, which is the basis for the approach that we refer to as the measurement-based approach.

Organization: The remainder of the paper is organized as follows. In Sec. II, we summarize the related work and state the novelty compared to related work. In Sec. III, we describe the model of the UNB IoT network and the incumbent devices. In Sec. IV, we mathematically describe the problem of BS placement and BS band assignment and pose it as an INLP. We then show how the problem can be relaxed such that it can be solved by knowing ADPs and JDPs at a number of locations in the environment. In Sec. V, we describe the model-based approach and in Sec. VI, we describe its training procedure. In Sec. VII, we describe the measurement-based approach. In Sec. VIII, we run simulations to evaluate and compare our approaches, and in Sec. IX, we provide conclusions and directions for future work.

Notations: Scalars, vectors, and matrices are denoted by regular, bold lower-case, and bold upper-case letters, respectively. The (i, j)-th element of $\mathbf{A}$ is denoted by $[\mathbf{A}]_{i,j}$. A two-element tuple is denoted as (a, b). An uniform random distribution in a closed interval $[a, b]$ is denoted as $\mathcal{U}[a, b]$. The set of all vectors of size N with binary valued entries is denoted as $\mathbb{Z}_2^N$. The function $\mathbf{1}_A$ is an indicator function equal to 1, if the logic statement A is true, and 0 otherwise. $|\mathcal{A}|$ is the Lebesque measure of a continuous set $\mathcal{A}$ or the cardinality of a discrete set $\mathcal{A}$. $\text{vec}(\mathbf{A})$ flattens a matrix into a column vector.

II. RELATED WORK

A. Our previous work

The theoretical benefits of multiband UNB LPWAN networks have been analyzed in [2], [3] using a stochastic geometry framework and simulation studies. Using multiple bands reduces IoT collisions and interference with incumbents. However, to fully exploit these gains when each BS is restricted to listen to one of the bands, it is necessary to optimize the BS-band selection policy. In [4], we proposed an algorithm for optimal BS-to-band assignment in a LPWA UNB multi-band network. This paper is an extension of that work with the focus expanding to joint BS placement and band assignment optimization. One of the two main proposed algorithms that we will discuss in this paper, the measurement-based approach, is based on our work in [4], but in this work it has been extended to be applicable for placement optimization. The second approach that will introduce in this paper, the model-based approach is entirely novel.

B. Existing prior literature

While the channel assignment and BS placement problems were previously researched in the context of various wireless technologies such as LPWANs, Wi-Fi networks [5], cellular networks [6]–[8], and cognitive unlicensed radio networks [9], to the best of our knowledge, no problem setting has considered the problem of optimal BS placement and band assignment in a multi-band UNB LPWAN. The distinctions arise due to two characteristics: (i) We consider a multi-band network in which IoT devices are able to transmit in any band but the receivers (BSs) are tuned to only one band over a long period of time; and (ii) we assume that there is no association between devices and BSs, i.e. any BS that receives a transmission is expected to decode it, which is the case for a limited number of network architectures. While LPWANs such as LoRa often satisfy characteristic (ii), the characteristic (i) is not met. Other technologies, such as Wi-Fi or cellular, in general neither meet (i) nor (ii). In the context of LPWANs, there are several works that consider BS
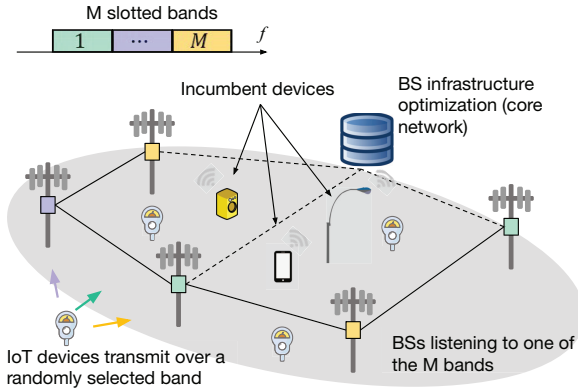
Fig. 1: The system model of the LPWAN IoT network.

III. MODEL OF THE IoT NETWORK

In this section, we describe the IoT network considered for the BS placement optimization and BS frequency assignment. We are interested in the planning of the operation of a LPWA UNB IoT network operating across multiple sub-1GHz unlicensed bands.

The interference is assumed to come not only from other devices of the same technology, as the IoT network is assumed to share the unlicensed band with other incumbent technologies such as LoRa, Zigbee or 802.11 WLANs.

Among the assumptions made in this section, only the assumptions described in Sec. III-A are used by the measurement-based approach while others are only necessary for model-based approach.

A. UNB network topology and transmission model

We denote the number of BSs present in the network as B. There are M multiplexing bands, each with bandwidth W. For ease of exposition, we assume a constant W across bands $m \in \{1, .., M\}$, even though our analysis is not dependent on this assumption. IoT devices use ALOHA random access protocol to access the bands and transmit signals at power P, occupying a bandwidth w with a random center frequency. Each packet is repeated R times, consecutively over time, yet randomly hopping from one frequency to another across any of the M bands, as shown in Fig. 1. The packet transmission duration is T. We label the transmissions of IoT devices with a tuple (n, r), where $n \in \{1, 2, .., \infty\}$ is the packet index and $r \in \{1, .., R\}$ corresponds to the repetition number of the packet. Therefore, the probality of a device being active is $\frac{NRT}{1\,h}$. In UNB LPWANs such as Sigfox, IoT devices are allowed to transmit at an average rate of up to N packets per hour [13], although they may transmit less in order to conserve energy. We assume that all devices are transmitting N packets per hour and that packets are transmitted as soon as they arrive into the queue, since there is no listen-before-talk mechanism. Hence, we model packet arrivals as a Poission process, with independent arrivals across devices.

We denote the time of a transmission (n, r) as $t(n, r)$. The transmission signals are extremely narrowband and the channel access is assumed to be unslotted, therefore we model the carrier frequency $f(n, r)$ as a uniform random distribution $\mathcal{U}\left[\frac{w}{2}, WM - \frac{w}{2}\right]$. Similarly, since the devices randomly transmit across any of the multiplexing bands, we can model the band selected for a particular packet transmission $\beta(n, r)$ as a discrete uniform random variable that takes on the values $m \in \{1, ..., M\}$. The probability distributions across each band is approximately $\Pr(\beta(n, r) = m) = \frac{1}{M}$, since $w \ll W$.

B. Interfering networks

We will assume a similar regular access pattern for incumbent devices with N_I packets per hour, R_I repetitions, power P_I, bandwidth w_I and transmission time T_I . The transmit probability of incumbent devices is then $\frac{N_I T_I}{1\,h}$. We also assume that transmissions can be modeled by a Poisson point process that operates independently across devices. This

placement optimization but not channel assignment. In [10], the authors consider the optimal placement of BSs for LoRa networks, however, their main objective is to minimize the capital-expenditure costs and maximize energy efficiency of the devices. Furthermore, it is assumed that the locations of the IoT devices are known which makes it possible to fairly reliably predict the PDP of a device given a particular BS placement, which is taken advantage of in their approach. We do not make the same assumptions since it is unlikely that all devices served by the BSs will be equipped with a GPS or other localization technologies due them often being energy limited. In [11], a similar objective is considered for the general category of LPWAN networks with known IoT device locations. The work in [12] considers placement when gateways perform interference cancellation with the objective of maximizing the packet delivery ratio, equivalent to PDP in this paper, however, it is assumed that the locations of the IoT devices are known. In summary, no approaches that were previously developed for placement in LPWANs would apply to the problem that we consider. Furthermore, to the best of our knowledge, no works in LPWANs consider channel assignment on the BS side because the BSs are normally assumed to operate on a single channel.

C. Overview of current technologies

Several commercial UNB LPWAN technologies are available on the market at the time of this writing, such as Sigfox [13], WAWIoT NB-Fi [14], and Telensa [15]. For a more detailed survey of various LPWAN technologies, including UNB, the reader is referred to [16]. Here, we briefly summarize the main characteristics cited in the technical specifications of these technologies that will be relevant to this thesis. First, the expected range of reliable communication is up to 10 miles. The uplink packet size payload is up to 12 bytes, depending on the type of data being transmitted, and the bandwidth of packet transmission ranges from 50Hz to 600Hz across technologies. Multiplexing band bandwidth ranges from 200kHz to 500kHz and the bands are located in the sub 1GHz unlicensed ISM bands for all existing commercial solutions.

assumptions is true only when the queues at each device are backlogged or packets arrivals at each device are sporadic. Otherwise, the analysis would be very difficult due to spatio-temporal correlation between incumbent transmissions [17].

To simplify our analysis, we further assume that the interference across multiplexing bands is uncorrelated. With this assumption, if a particular device is occupying several bands, it can only be approximately modeled by separate uncorrelated virtual devices.

The number of incumbent devices operating on each multiplexing band m may not be the same, therefore IoT devices on some bands may experience a higher incumbent interference than on others. The assumption of unequal incumbent interference across different frequencies is consistent with measurement studies performed in IoT networks [18], which show non-uniform interference power distribution in sub-1 GHz ISM bands.

C. UNB and Interfering devices' topology

There are D UNB IoT devices across an area of interest $\mathcal{A}$ whose locations are assumed to be unknown, however, we assume that their locations are sampled from a uniform distribution. The area $\mathcal{A}$ contains all the IoT devices for which we aim to maximize the PDP by optimizing the BS deployment. Modeling the distribution of the locations of wireless devices as uniform is justified in many instances, including IoT networks. The number of incumbent devices on each multiplexing band m is $D_{I,m}$ and their locations are also sampled from a uniform distribution across $\mathcal{A}$.

IV. PROBLEM STATEMENT AND PROPOSED SOLUTIONS

In this section, we define the problem of optimal BS placement and the optimal assignment of BSs to multiplexing bands.

A. Problem statement

The main objective of the algorithm will be to maximize the probability of decoding a packet (PDP) for a typical device in the network. There are $B \geq 0$ currently deployed BSs and we assume that there is a need to expand the BS infrastructure by installing new BSs in $\mathcal{A}$. The network will be expanded by introducing additional $\Delta B \geq 0$ BSs to the network, and we need to find an optimal placement for these BSs and their multiplexing band assignment. When new BSs are installed, the band assignment of the currently deployed BSs may need to be updated. We assume ΔB is pre-determined by the resource availability of the network operator and not optimized. When $\Delta B = 0$, then the problem reduces to the optimal assignment of BSs to bands. When installing new BSs, we assume that there are $C \geq \Delta B$ candidate locations for new BSs. This would be determined by practical constraints such as connectivity to the core network or real-estate considerations. Among the C candidate locations, ΔB locations will be populated by new BSs.

Let us introduce an extended set of BS locations that includes the locations currently occupied by BSs and the candidate locations for the new BSs $b \in \{1, \ldots, B, \ldots B+C\}$, where the locations $b \in \{1, \ldots, B\}$ correspond to the locations currently occupied by the BSs.

We will be solving for the binary assignment variable $\mathbf{X} \in \mathbb{Z}_2^{(B+C) \times M}$, where

$$[\mathbf{X}]_{b,m} = \begin{cases} 1 & \text{A BS is placed at } b \text{ and assigned to band } m \\ 0 & \text{No BS at location } b \text{ and band } m \end{cases}.$$
(1)

Given an assignment $\mathbf{X}$, our objective is to maximize the average PDP of a packet n, defined as follows:

$$P_{\text{PDP}}(\mathbf{X}) = \Pr\left(\bigcup_r \bigcup_b [\mathbf{X}]_{b,\beta(n,r)} \text{SINR}_{b,\beta(n,r)}(n,r) > \tau \right)$$
(2)

We are interested in maximizing the probability of decoding of a packet irrespective of its source. The union over r in (2) captures that if any of the repetitions $r = 1, ..., R$ is decoded, the packet n is considered decoded. The union over b captures that a packet repetition is decoded if at any of the locations b there is a BS synchronized to band $\beta(n,r)$ (i.e. $[\mathbf{X}]_{b,\beta(n,r)} = 1$) that is able to decode it. A BS can decode a repetition at location b if the SINR of the packet repetition at the location b, $\text{SINR}_{b,\beta(n,r)}(n,r)$, exceeds the decoding threshold τ.

The optimization problem can be stated as:

$$\max_{\mathbf{X}} \quad P_{\text{PDP}}(\mathbf{X}) \tag{P.1.1}$$

$$\text{s.t.} \quad \begin{bmatrix} \mathbf{1}_B \\ \mathbf{0}_C \end{bmatrix} \leq \mathbf{X}\mathbf{1}_M \leq \begin{bmatrix} \mathbf{1}_B \\ \mathbf{1}_C \end{bmatrix}, \tag{P1.2}$$

$$\begin{bmatrix} \mathbf{0}_B^T & \mathbf{1}_C^T \end{bmatrix} \mathbf{X}\mathbf{1}_M = \Delta B, \tag{P1.3}$$

$$\mathbf{X} \in \mathbb{Z}_2^{(B+C) \times M} \tag{P1.4}$$

In the above, $\mathbf{1}_B$ is a vector of ones of size $B \times 1$ and $\mathbf{0}_C$ is a vector of zeros of size $C \times 1$. The constraint (P1.2) ensures that BSs $b \in \{1, \ldots, B\}$ that are already installed, remain at their location and are assigned to listen to a multiplexing band and none of the BSs are assigned to more than one band. The constraint (P1.3) ensures that ΔB BSs are placed to one of the C candidate locations and assigned to one of the M multiplexing bands. The constraint (P1.4) constrains the assignment variable $\mathbf{X}$ to have entries limited to $\{0, 1\}$.

In order to be able to solve this problem, we would have to know what the function $P_{\text{PDP}}(\mathbf{X})$ is in terms of the assignment variable $\mathbf{X}$. Learning or estimating this function would be non-trivial since the number of possible $\mathbf{X}$ is $\binom{C}{\Delta B} M^{B+\Delta B}$.

B. Suboptimal solution to P1

To relax the optimization problem (P1), we seek to optimize the probability of decoding of a repetition of a packet, $P_{\text{TDP}}(\mathbf{X}) = \Pr\left(\bigcup_b [\mathbf{X}]_{b,\beta(r)} \text{SINR}_{b,\beta(r)}(r) > \tau \right)$, which is a lower bound on $P_{\text{PDP}}(\mathbf{X})$. $P_{\text{TDP}}(\mathbf{X})$ is a lower bound on $P_{\text{PDP}}(\mathbf{X})$ since a packet is decoded if any of its R repetitions is decoded, therefore $P_{\text{PDP}}(\mathbf{X})$ is always larger than $P_{\text{TDP}}(\mathbf{X})$. As we will demonstrate in this section, working with the expression for $P_{\text{TDP}}(\mathbf{X})$ makes the analysis more tractable

and allows us to arrive at a simpler optimization problem. The relaxed problem (P1) is:

$$\max_{\mathbf{X}} \quad P_{\text{TDP}}(\mathbf{X})$$
$$\text{s.t.} \quad \text{(P1.2), (P1.3), (P1.4)} \tag{P2}$$

Note that (P2) is equivalent to (P1) when the number of repetitions is $R = 1$.

Proposition 1: For a given τ, a suboptimal solution to P2 can be obtained by solving the following convex integer optimization problem:

$$\max_{\mathbf{X}} \quad \sum_m \bigg(\sum_b [\mathbf{X}]_{b,m} \mathbb{E}_r\{A_{b,m}(r)\} -$$
$$\sum_{b<v} [\mathbf{X}]_{b,m}[\mathbf{X}]_{v,m} \mathbb{E}_r\{A_{b,m}(r)A_{v,m}(r)\} \bigg) \tag{P3}$$
$$\text{s.t.} \quad \text{(P1.2), (P1.3), (P1.4)}$$

where $A_{b,m}(r) = \mathbf{1}_{(\text{SINR}_{b,\beta(\mathbf{n},\mathbf{r})}(\mathbf{n},\mathbf{r})>\tau)}$. The objective function is a concave function and is also a lower bound on the objective function in (P3). Therefore, the objective function in (P3) is a lower bound on the objective function in (P1).

Proof: The proof is given in Appendix A. ∎

C. Practical solution of the proposed optimization problem

As shown by our results in the latter sections, approximating the problem (P1) by (P3) proves to be a suitable relaxation for finding the BS assignment and placement that will result in maximizing the PDP in a realistic environment.

While the solution given by (P3) can be used to maximize the lower bound on packet decoding probability, in order to implement it, we need to know $\mathbb{E}_r\{A_{b,m}\} \ \forall b$, which are the ADPs at each of the locations in the set $\{1, \ldots, B + C\}$. We also need the knowledge of $\mathbb{E}_r\{A_{b,m}(r)A_{v,m}(r)\} \ \forall b, v$, which are the JDPs between all pairs of locations in the set $\{1, \ldots, B + C\}$. Additionally, even with the knowledge of ADPs and JDPs the problem may be difficult to solve due to integer constraints on the optimization variables.

While general integer optimization problems are known to be NP-hard, the optimization problem (P3) stands out in that its objective function is convex and its constraints are linear. It can be shown that there exists an algorithm that can solve a convex integer problem in polynomial time given a fixed number of variables [19, p. 574]. However, the worst-case computational complexity asymptotically grows exponentially with the number of variables. In our case, the number of variables is $M \times (B+C)$. Therefore, even though this problem is solvable it may not be tractable. In our solution we use a branch-and-bound algorithm to solve this problem and place a T_{sol} threshold on the solver CPU time as a stopping criterion. T_{sol} was selected such that in the majority of cases the optimal solution was found on our machine. To guarantee tractability, relaxations of the problem (P3) could be pursued to lower computational complexity, however that is beyond of the scope of this paper.

Therefore, the main challenge that we will focus on is to accurately and efficiently estimate the ADPs $\mathbb{E}_r\{A_{b,m}\} \ \forall b$ and JDPs $\mathbb{E}_r\{A_{b,m}(r)A_{v,m}(r)\} \ \forall b, v$. Due to this critical aspect

of our problem, we will propose two alternative algorithms for BS placement and band assignment:

1) *Measurement-based (MEAS) placement and assignment approach*: In this approach, we directly measure $\mathbb{E}_r\{A_{b,m}\} \ \forall b$ and $\mathbb{E}_r\{A_{b,m}(r)A_{v,m}(r)\} \ \forall b, v$ through an optimized training procedure during the operation of the network. However, the complexity of this training procedure scales with the number of candidate locations since the number of JDP parameters to be estimated is $\binom{B+C}{2}$. For this reason, we have developed an alternative, more practical approach.

2) *Model-based (MOD) placement and assignment approach*: In this approach, we will use models of ADP and JDP to predict $\mathbb{E}_r\{A_{b,m}\} \ \forall b$ and $\mathbb{E}_r\{A_{b,m}(r)A_{v,m}(r)\} \ \forall b, v$. This approach necessitates development of suitable models of ADP and JDP and estimating the parameters of these models to make the correct predictions on ADP and JDP.

V. MODEL-BASED PLACEMENT AND ASSIGNMENT APPROACH

In this section, we describe the proposed model-based approach. In Sec. V-C, we will develop a model of ADP as a function of a set of parameters $\psi_{\text{ADP},m}$ dependent on band m, denoted as $\Pr(\text{SINR}_{b,m} \geq \tau \mid \psi_{\text{ADP},m})$ (Theorem 2). In Sec. V-D, we will develop a model of JDP as a function of a set of parameters, $\psi_{\text{JDP},m}$, and the separation between receiving locations b and v, $d_{b,v}$, denoted as $\Pr(\text{SINR}_{b,m} \geq \tau, \text{SINR}_{v,m} \geq \tau \mid d_{b,v}, \psi_{\text{JDP},m})$ (Theorem 4). Then, the models of ADP and JDP can be applied in a real environment by estimating the parameters $\psi_{\text{ADP},m}$ and $\psi_{\text{JDP},m}$ and using the estimated parameters to predict ADPs and JDPs as:

$$\mathbb{E}_r\{A_{b,m}(r)\} = \Pr(\text{SINR}_{b,m} \geq \tau \mid \psi_{\text{ADP},m}) \ \forall b, \tag{4}$$

$$\mathbb{E}_r\{A_{b,m}(r)A_{v,m}(r)\} =$$
$$\Pr(\text{SINR}_{b,m} \geq \tau, \text{SINR}_{v,m} \geq \tau \mid d_{b,v}, \psi_{\text{JDP},m}) \ \forall b, v \tag{5}$$

Hence, (P3) can be solved using the predicted ADPs and JDPs.

A. Stochastic geometry preliminaries

In this section, we introduce some fundamental concepts from stochastic geometry that will assist us in development of the model of the ADP, $\Pr(\text{SINR}_{b,m} \geq \tau \mid \psi_{\text{ADP},m})$, and the model of the JDP, $\Pr(\text{SINR}_{b,m} \geq \tau, \text{SINR}_{v,m} \geq \tau \mid d_{b,v}, \psi_{\text{JDP},m})$.

Stochastic geometry is linked to spatial point processes. A point process (PP) is a countable random collection of points that reside in some measure space, usually the Euclidean space $\mathbb{R}^d$ [20]. Formally, a PP is a countable random set $\Phi = \{x_1, x_2, \ldots\} \in \mathbb{R}^d$ consisting of random variables $x_i \in \mathbb{R}^d$ as its elements.

A PP model that is often used to model the topologies of wireless networks is the Homogeneous Poisson Point Process (HPPP). A HPPP with density λ is a PP in $\mathbb{R}^d$ such that for

every compact set $\mathcal{A}$, the number of points has a Poisson distribution with mean $\lambda|\mathcal{A}|$, where $|\cdot|$ is the Lebesgue measure in d dimensions, and if two compact sets $\mathcal{A}_1$ and $\mathcal{A}_2$ are disjoint, then the number of points in each is independent of one other.

One important result for the analysis of HPPP that we will utilize is the probability generating functional (PFGL) of a HPPP. Let $\mathcal{V}$ be the family of measurable functions, then for a $v(x) \in \mathcal{V}$ and a HPPP with density λ [20, p. 150]:

$$\mathbb{E}_\Phi \left(\prod_{x \in \Phi} v(x) \right) = \exp\left(-\int_{\mathbb{R}^d} [1 - v(x)]\lambda dx \right). \quad (6)$$

With these basic concepts covered, we will explain how stochastic geometry can be used to model our problem.

B. Stochastic modeling of the IoT network

In this section, we describe a way of modeling the activity of IoT and incumbent devices as random processes. In our modeling, we preserve the assumptions about the IoT network and the environment, stated in Sec. III. By applying theorems from stochastic geometry on these random processes, we will develop models of the ADP and JDP in the later sub-sections.

1) Modeling the topology of active devices as a HPPP: While we assume that the number and the position of the nodes in the network is fixed, given that the IoT and incumbent devices access the medium using the ALOHA protocol, the set of devices that are transmitting at some time $\tilde{t}$ and frequency $\tilde{f}$ in $|\mathcal{A}|$ is random. We will argue that that the distribution of both IoT and incumbent devices of active devices at some time $\tilde{t}$ and frequency $\tilde{f}$ can be modeled as a HPPP and can then be analyzed using the tools from stochastic geometry. Furthermore, it is important to emphasize that the analysis of networks using stochastic geometry is in most cases applied to finding the average of some random quantity over many realizations of wireless networks sampled from some spatial point process. In our case, we are interested in a deterministic topology of devices but as we will argue in this section due to the random access nature of IoT and incumbent devices, their activity from one time slot to the next can be approximately modeled as a HPPP.

Let us focus on the IoT devices first, since our model for the incumbent network is similar in all aspects except that incumbents are operating over a single multiplexing band. Since we assume an ALOHA-like protocol and Markovian packet arrivals, the probability of an IoT device being active at some time $\tilde{t}$ and frequency $\tilde{f}$ is p, where

$$p = 2\frac{NRT}{1\ h} \times 2\frac{w}{MW}. \quad (7)$$

Let the set of active IoT devices at time $\tilde{t}$ and frequency $\tilde{f}$ be Φ. Then the size of Φ, $|\Phi|$, is given by a Binomial distributions, $\Pr(|\Phi|) = \binom{D}{|\Phi|}p^{|\Phi|}(1-p)^{(D-|\Phi|)}$. Since D is large and $p \ll 1$, we can approximate the Binomial distribution model of $|\Phi|$ by a Poisson distribution with a parameter Dp, $\Pr(|\Phi|) = \frac{(Dp)^{|\Phi|}\exp(-Dp)}{|\Phi|!}$. According to a rule of thumb, this approximation is good if D is large and p is small, and Dp is not large [21]. Furthermore, the devices

in Φ are uniformly distributed across $\mathcal{A}$, therefore the set of active devices Φ can be approximated to be a realization of an HPPP with density $\lambda = \frac{Dp}{|\mathcal{A}|}$.

Using a similar reasoning, we can approximately model the set of active incumbent nodes Φ_I at time $\tilde{t}$ and frequency $\tilde{f}$ as a HPPP with density $\lambda_{I,m} = \frac{D_{I,m}p_{I,m}}{|\mathcal{A}|}$, where the density of the incumbent devices depends on the multiplexing band m.

Note that the activity of incumbent and IoT devices could be modelled slightly more accurately using a homogenous binomial point process (HBPP), where the number of points in a closed set is modeled using a Binomial distribution. However, we choose to use the HPPP over the HBPP since it allows us to arrive at some general expressions for models of ADP and JDP.

2) Modelling the SINR of transmissions: Let the SINR of some transmission (n, r) from IoT device i to BS b over multiplexing band m be $\text{SINR}_{i,b,m}(n, r)$. A transmission (n, r) is considered to be successfully decoded if $\text{SINR}_{i,b,m}(n, r)$ at any of BS b listening to the band $\beta(n, r)$ exceeds a threshold τ. In practice, τ can be the minimum SINR required to achieve a certain bit error rate (BER) performance. The value of τ depends on specific coding, modulation, and detection schemes being employed. For example, in the Sigfox technical documentation it is stated that a good detection probability can be achieved if the SINR exceeds 8 dB [22]. The $\text{SINR}_{i,b,m}$ from an IoT device i is modeled as

$$\text{SINR}_{i,b,m} = \frac{h_i p_{i,b}^{-\alpha}}{\hat{P}_N + \sum_{j \in \Phi} f_j p_{j,b}^{-\alpha} + \sum_{j' \in \Phi_{I,m}} \hat{P}_I f_{j'} p_{j',b}^{-\alpha}} \quad (8)$$

where $p_{i,b}$ is the distance of the device i to BS b, $p_{j,b}$ is the distance of an interfering IoT device j, and $p_{j',b}$ is the distance of an incumbent device j'. α is the path-loss exponent and h_i, f_j and $g_{j'}$ are fading gains, modeled as Exp(1). $\hat{P}_N$ and $\hat{P}_I$ are noise power and incumbent device power expressed as a fraction of the IoT device transmission power. In order for the path loss function $p^{-\alpha}$ to be integrable ($\int_0^\infty p \times p^{-\alpha} dp < \infty$), we assume that $\alpha > 2$.

C. Modeling the ADP

We first derive the expression for the probability of decoding a transmission from a device i at a receiver b, on band m, denoted as $\Pr(\text{SINR}_{i,b,m} \geq \tau \mid p_{i,b})$. Then we use $\Pr(\text{SINR}_{i,b,m} \geq \tau \mid p_{i,b})$ to derive a model of the ADP, $\Pr(\text{SINR}_{b,m} \geq \tau)$.

Theorem 1: The complementary cumulative distribution (CCDF) of $\text{SINR}_{i,b,m}$ when noise power is negligible ($\hat{P}_N \to 0$) is:

$$\Pr(\text{SINR}_{i,b,m} \geq \tau \mid p_{i,b}) = \exp\left(-\epsilon_m \tau^{\frac{2}{\alpha}} p_{i,b}^2 \right), \quad (9)$$

where $\epsilon_m = \pi \left(\lambda + \hat{P}_I^{\frac{2}{\alpha}} \lambda_{I,m} \right) \frac{2\pi/\alpha}{\sin(2\pi/\alpha)}$. The term ϵ_m is proportional to the interference in the network, since it depends on the density of incumbent and IoT users. Note that as α increases, the term $\hat{P}_I^{\frac{2}{\alpha}} \lambda_{I,m}$ approaches $\lambda_{I,m}$, in cases when either $\hat{P}_I > 1$ (the incumbent power is stronger than IoT device power) or when $\hat{P}_I \leq 1$, which indicates that as α
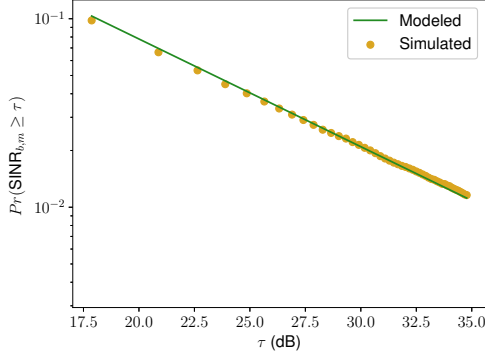
Fig. 2: Simulated and modeled ADP as a function of the decoding threshold τ. The modeled results are obtained using (10). We have simulated a disk shaped UNB IoT network with the following parameters: 300km radius, $D = D_I = 15E6$, $M = 1$, $N = 3$, $R = 3$, $w = 600$Hz, $w_I = 200$kHz, $W = 200$kHz, $T = (208/w)$s, and $T_I = (2080/w_I)$s.

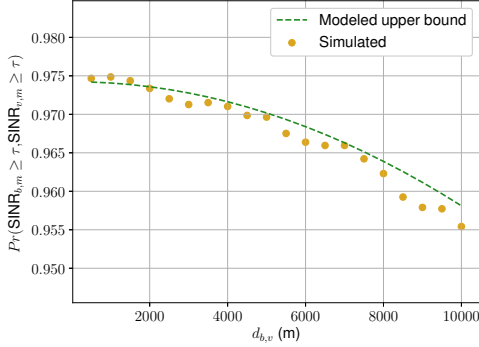
Fig. 3: Simulated and modeled JDP as a function of the separation between BS, $d_{b,v}$. The modeled results are obtained using the derived expression in 12. The simulated results are obtained by measuring the average rate of transmissions decoded by both BS b and v. The graph is obtained in the same simulation environment as the one described in Fig. 2. Additionally, we set $\tau = 0$ dB.

increases, the power of incumbent devices relative to IoT devices becomes less important, and only the density matters. This is sensible since when α is large, the signal power from distant devices decays significantly and only the activity of devices close to the BS is important. However, if an incumbent device close to the BS is active during a transmission its transmit power is not relevant as it will likely have a higher power than the IoT transmission anyway due its proximity to the BS.

Proof: The proof is shown in Appendix B. ∎

We will now evaluate the ADP across all transmitters for a finite network $\mathcal{A}$ around the receiver b.

Theorem 2: The ADP irrespective of the source is

$$\Pr\left(\text{SINR}_{b,m} \geq \tau\right) = \int_{\mathcal{A}} \exp\left(-\psi_m p_{i,b}^2\right) \frac{p_i}{|\mathcal{A}|} dp_i \theta_i, \quad (10)$$

where $|\mathcal{A}|$ and (p_i, θ_i) are the coordinates of the device i with respect to the origin. Let (p_b, θ_b) be the polar coordinates
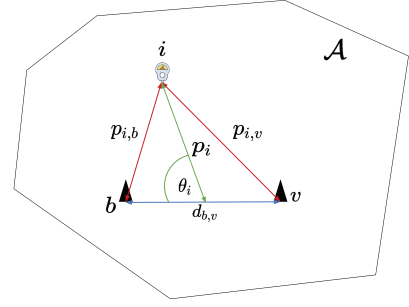

Fig. 4: Illustration of the example used to model the joint decoding probability between two BSs.

of the receiving location, then $p_{i,b}^2 = p_i^2 + p_b^2 - 2p_i p_b \cos(\theta_i - \theta_b)$. Moreover, $\psi_m = \epsilon_m \tau^{\frac{2}{\alpha}}$, and hence $\psi_{\text{ADP},m} = \{\psi_m, \mathcal{A}\}$. The unit of ψ_m is m^{-2}.

Proof: The expression is obtained by taking the expectation over p_i and θ_i in the expression in (9). The joint PDF of p_i and θ_i is $\frac{p_i dp_i \theta_i}{|\mathcal{A}|}$. ∎

To the best of our knowledge, the integral in (10) cannot be simplified to a closed-form expression, hence when evaluating the ADP we will use numerical integration.

In Fig. 2, we show the simulated ADP as a function of the decoding threshold τ for a disk-shaped $\mathcal{A}$ centered around the receiver b. The ADP modeled using (10) closely matches the simulated results which supports our model of the IoT network as a HPPP and the model in (10).

D. Modeling the JDP

Here, we derive the expression for the JDP, denoted as $\Pr(\text{SINR}_{i,b,m} \geq \tau, \text{SINR}_{i,v,m} \geq \tau)$. First, we will derive the expression for the upper bound on the JDP for a given transmitter i, denoted as $\Pr(\text{SINR}_{i,b,m} \geq \tau, \text{SINR}_{i,v,m} \geq \tau \mid i)$. We assume that the BSs b and v are located at the origin of the polar coordinate system as shown in Fig. 4 with coordinates $\left(\frac{d_{b,v}}{2}, \frac{\pi}{2}\right)$ and $\left(\frac{d_{b,v}}{2}, -\frac{\pi}{2}\right)$, respectively.

Theorem 3: The upper bound on conditional JDP for a given transmitter i is:

$$\Pr(\text{SINR}_{i,b,m} \geq \tau, \text{SINR}_{i,v,m} \geq \tau \mid i) \leq$$
$$\exp\left(-\frac{1}{2}\epsilon_m \tau^{\frac{2}{\alpha}} d_{b,v}^2\right) \times \exp\left(\left(-\frac{2}{\alpha} - 1\right)\epsilon_m \tau^{\frac{2}{\alpha}} p_i^2\right) \quad (11)$$

where p_i is the distance of a device i to the origin of the coordinate system. Note that the validity of this model depends on the approximation that the midpoint between the two BSs is located at the origin, therefore it can only serve as an approximation for BS pairs that are located near the edge.

Proof: The proof is shown in Appendix C. ∎

Theorem 4: The JDP irrespective of the source is:

$$\Pr\left(\text{SINR}_{b,m} \geq \tau, \text{SINR}_{v,m} \geq \tau\right) = \Psi_m \exp\left(-\frac{1}{2}\psi_m d_{b,v}^2\right) \quad (12)$$

where $\Psi_m = \int_{\mathcal{A}} \exp\left(\left(-\frac{2}{\alpha} - 1\right)\epsilon_m \tau^{\frac{2}{\alpha}} p_i^2\right)$. Then, $\psi_{\text{JDP},m} = \{\psi_m, \Psi_m\}$. This theorem allows us to model JDP between two BSs as a function of their separation, $d_{b,v}$.

Algorithm 1: Model-based placement and frequency assignment algorithm

---

Input: $\{\psi_m, \Psi_m\} \; \forall m, \; (p_b, \theta_b) \; \forall b, \; \mathcal{A}$
Output: Assignment $\mathbf{X}$

1 **for** $b \in \{1, \dots, B + C\}$ **do**
2   % Estimate the ADP
3    $\mathbb{E}_r\{A_{b,m}(r)\} \leftarrow \int_{\mathcal{A}} \exp\left(-\psi_m p_{i,b}^2\right) \frac{p_i}{|\mathcal{A}|} dp_i \theta_i \; \forall m$
4   **for** $v \in \{1, \dots, b\}$ **do**
5       $d_{b,v} \leftarrow \sqrt{p_b^2 + p_v^2 - 2p_b p_v \cos(\theta_b - \theta_v)}$
6   % Estimate the JDP
7       $\mathbb{E}_r\{A_{b,m}(r) A_{v,m}(r)\} \leftarrow$
         $\Psi_m \exp\left(-\frac{1}{2}\psi_m d_{b,v}^2\right) \; \forall m$

8 Solve for $\mathbf{X}$ in (P3)

Proof: The proof is obtained by taking the expectation with respect to p_i of the expression in (11). ∎

In Fig. 3, we show the simulated JDP for two BSs, b and v, whose separation $d_{b,v}$ is varied. The JDP modeled using Theorem 4 is an upper bound to the simulated results.

E. Summary of the proposed algorithm

In Algorithm 1, we summarize the MOD approach for assignment and placement assuming that $\psi_{\text{JDP},m}$, $\psi_{\text{ADP},m}$ and positions of the BSs are known. The algorithm is based on predicting ADP and JDP using Theorem 2 and Theorem 4, respectively. The predicted ADPs and JDPs are used in (P3) to solve for optimal placement and frequency assignment of BSs.

In Sec. VI, we describe the method for estimation of parameters of $\psi_{\text{JDP},m}$ and $\psi_{\text{ADP},m}$ during network operation.

VI. LEARNING APPROACH FOR MODEL-BASED ASSIGNMENT

A. Parameter estimation

As shown in Algorithm 1, to perform the assignment, our approach requires the information about Ψ_m and ψ_m for all m, locations of current BSs and candidate locations and the contour of the area $\mathcal{A}$. It is assumed that the contour of the area $\mathcal{A}$ is known by the network operator network, without assuming that the locations of the IoT devices are known or their number. We will now discuss how to estimate parameters ψ_m and Ψ_m for all m during the operation of the network.

We will use estimates of JDP between BSs in $\mathcal{A}$ to estimate ψ_m and Ψ_m for all m. Furthermore, we will denote the estimates JDP between two locations b and v as $\tilde{\mathbb{E}}_r\{A_{b,m}(r) A_{v,m}(r)\}$. Let us denote the set of JDP estimates on each band as $\mathcal{S}_m$, where $|\mathcal{S}_m| = S_m$. To estimate ψ_m and Ψ_m for all m, we will use the least squares criterion between the measured JDP and the fitted model. Parameter estimation is then a non-linear least squares (NLSQ) optimzation problem:

$$\min_{\psi_m, \Psi_m} \sum_{m \in \{1, \dots, M\}} \sum_{(b,v) \in \mathcal{S}_m} |\tilde{\mathbb{E}}_r\{A_{b,m}(r) A_{v,m}(r)\} - \Psi_m \exp\left(-\psi_m d_{b,v}^2\right)|^2 \quad \text{(P4)}$$
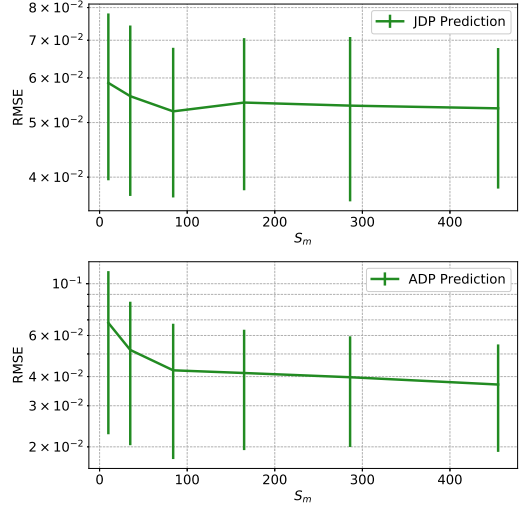

Fig. 5: Root mean square error (RMSE) in prediction of ADP and JDP at a random location in $\mathcal{A}$ as a function of S_m used to estimate the ADP and JDP model parameters. The graph is obtained in the same simulation environment as the one described in Fig. 2

which can be solved using the Levenberg–Marquardt algorithm [23] for all m.

The root mean square error (RMSE) of ADP and JDP prediction as a function of S_m is shown in Fig. 5. Based on the simulation results, we observe that the accuracy of prediction stops significantly decreasing past $S_m = 10$ and saturates past $S_m = 20$.

B. Training procedure for estimation of $\psi_{\text{ADP},m}$ and $\psi_{\text{JDP},m}$

In this section, we describe the training procedure for collection of measurements of JDP. The training episode is executed to keep up with changes in the network or the environment, for example, if the density of IoT devices changes or if the IoT network expands.

The required $\mathcal{S}_m$ estimates will be collected by the currently installed BSs and, if necessary, measurements collected by $\hat{B}$ temporary installed BSs that can assist with training. Installation of temporary BSs can be assisted with the help of UAVs that can carry and place the BSs to a designated location. During the operation of the network, installed BSs and temporary BSs will record transmissions that they receive, as well as their timestamps and the band that they were received on. Using the core network, BSs forward their recordings to a central processor, which then estimates the JDPs, $\tilde{\mathbb{E}}_r\{A_{b,m}(r) A_{v,m}(r)\}$. We assume that the central processor can identify all the successful transmissions that occurred from the captures. Hence, the network can only have information about the successful transmissions, but not directly about the unsuccessful ones. In order to find the expected JDP, the infrastructure should know how many packets have been sent and how many repetitions have been made. This can be known by, for example, IoT devices adding packet and repetition numbering to their transmissions.

The number of different S_m estimates of JDP that can be collected for a given number of temporary and installed BSs is determined by their assignment to bands. For example, if there are B_m BSs assigned to operate on a band m, then we can collect $\binom{B_m}{2}$ estimates of JDP on that band. However, if $S_m > \binom{B_m}{2}$ estimates are needed for parameter estimation then it will be necessary to move BSs across bands to collect the necessary measurements. While the assignment of temporary BSs to bands can be changed without repercussions, the reassignment of installed BSs to bands will impact the PDP and therefore cannot be changed arbitrarily, since we assume that the temporary BSs do not contribute to the PDP. For example, we cannot assign all of the installed BSs to a single band during training as leaving the other bands to be unoccupied would severely harm the PDP.

Let $\{1, \ldots, B, \ldots, \hat{B} + B\}$ be the set of locations of installed and temporary BSs combined. Then, let us use the assignment variable $\mathbf{X} \in \mathbb{Z}_2^{(\hat{B}+B) \times M}$ for the assignment of both temporary and installed BSs to bands, where $[\mathbf{X}]_{b,m}$ is equal to 1 if BS b is assigned to band m and 0 otherwise. Furthermore, let the set of viable assignments during training be $\mathcal{X}_{\text{train}}$, where $\mathcal{X}_{\text{train}} = \{\mathbf{X}_i \mid i = 1, \ldots, |\mathcal{X}_{\text{train}}|\}$. The set $\mathcal{X}_{\text{train}}$ is determined by the constraints on the assignment of installed BSs. The constraints may be heuristic such as ensuring that a certain number of BSs remains on a particular band or ensuring to select assignments whose predicted TDP exceeds a certain threshold. The predicted TDP for a particular assignment $\mathbf{X}$ can be obtained by using estimates of ADP and JDP from earlier training episodes and evaluating the objective function in (P3).

Let us now define the training procedure. There will be several phases, during each of which a different assignment $\mathbf{X}_i$ will be applied. Since our objective is to complete the training as fast as possible, we need to select the least number of assignments $\mathbf{X}_i$ such that S_m distinct JDP estimates are obtained on each band m. The set of assignments selected for training is $\mathcal{L}$.

We observe that the problem of selecting $\mathcal{L}$ is a variant of the NP hard *set cover problem*. In the set cover problem, given a collection of subsets $\mathcal{T}$ of a ground set $\mathcal{U}$, the goal is to cover $\mathcal{U}$ with the smallest number of subsets. An extension of the set cover problem is the *partition set cover problem*, also an NP problem. In the partition set cover problem, $\mathcal{U}$ is divided into M partitions, $\mathcal{U} = \{\mathcal{U}_1, \ldots, \mathcal{U}_M\}$, and the goal is to cover S_m elements from each $\mathcal{U}_m$ with the smallest number of subsets from $\mathcal{T}$. Let us demonstrate that the problem of selecting $\mathcal{L}$ is the partition set cover problem. Let $\mathcal{U}_m = \{\tilde{\mathbb{E}}_r\{A_{b,m}(r)A_{v,m}(r)\} \mid b = 1, \ldots, \hat{B} + B, \ v = 1, \ldots, b\}$ be the set of distinct JDP estimates that can be collected on band m. Furthermore, let $\mathcal{T}_i$ be the set of JDP estimates that can be collected given the assignment $\mathbf{X}_i \in \mathcal{X}_{\text{train}}$, and $\mathcal{T} = \{\mathcal{T}_1, \ldots, \mathcal{T}_{|\mathcal{X}_{\text{train}}|}\}$. We assume that $\cup_{i=1,\ldots,|\mathcal{X}_{\text{train}}|} \mathcal{T}_i$ and $\cup_{m=1,\ldots,M} \mathcal{U}_m$ are identical. If not, $\mathcal{U}$ can be revised to remove the JDPs $\mathcal{U} \setminus \cup_{i=1,\ldots,|\mathcal{X}_{\text{train}}|} \mathcal{T}_i$. Selecting the $\mathcal{L}$ assignments is equivalent to selecting $|\mathcal{L}|$ sets $\mathcal{T}_i$ that will cover S_m elements from each partition $\mathcal{U}_m$, therefore this is a partition set cover problem. For the sake of simplicity, in this paper, we will use

---

Algorithm 2: Greedy algorithm for the selection of $\mathcal{L}$ BS assignments $\mathbf{X}_i$ for training.

Input: $\mathcal{U}, \mathcal{T}, S_1, \ldots, S_M$
Output: $\mathcal{L}$

1   $\mathcal{L} = \emptyset$
2   **while** *any* $s_m > 0$ **do**
3   Find $l \in (\{1, \ldots, |\mathcal{X}_{\text{train}}|\} \setminus \mathcal{L})$ that maximizes $\min(\sum_m s_m, \sum_m |\mathcal{U}_m \cap \mathcal{T}_l|)$
4     $S_m \leftarrow S_m - |\mathcal{U}_m \cap (\cup_{l \in \mathcal{L}} \mathcal{T}_l)| \ \forall m$
5   Add l to L
6     $\mathcal{T}_i \leftarrow \mathcal{T}_i \setminus \mathcal{T}_l$ for $i = 1, \ldots, |\mathcal{X}_{\text{train}}|$

---

a polynomial time greedy approach to solve the partition set cover problem developed in [24]. Better solutions can also be pursued, such as the one in [25]. We describe the greedy approach for planning of the training procedure in Algorithm 2.

The algorithm returns the set of assignments $\mathcal{L}$ that will result in S_m estimates from $\mathcal{U}_m$ being collected for each m. Assignments are added to $\mathcal{L}$ with priority given to the assignments that can learn most new estimates in $\mathcal{U}$ given the currently selected assignments in $\mathcal{L}$.

VII. MEASUREMENT-BASED PLACEMENT AND ASSIGNMENT APPROACH

In this section, we will discuss an alternative approach to MOD, in which ADPs, $\mathbb{E}_r\{A_{b,m}(r)\} \ \forall b$, and JDPs $\mathbb{E}_r\{A_{b,m}(r)A_{v,m}(r)\} \ \forall b, v, m$ are estimated directly using measurements from BSs in the field instead of predicting ADPs and JDPs via modeling. As in the MOD approach, the estimated ADPs and JDPs are plugged into (P3) to solve for optimal placement and band assignment of BSs. We refer to this approach as MEAS for convenience.

A. Training procedure

The training involves the direct estimation of the necessary ADPs, $\hat{\mathbb{E}}_r\{A_{b,m}(r)\} \ \forall b, m$, and the necessary JDPs $\hat{\mathbb{E}}_r\{A_{b,m}(r)A_{v,m}(r)\} \ \forall b, v, m$. Similar to the training phase of the MOD approach, during the operation of the network, installed BSs and temporary BSs will record transmissions that they receive, as well as their timestamps and the band that they were received on. BSs forward their recordings to a central processors, which then estimates the ADPs, $\hat{\mathbb{E}}_r\{A_{b,m}(r)\} \ \forall b, m$, and JDPs $\hat{\mathbb{E}}_r\{A_{b,m}(r)A_{v,m}(r)\} \ \forall b, v, m$. Evidently, the assignment of installed and temporary BSs will determine which ADPs and JDPs can be measured. Hence, the training may have to take place over several stages with a different BS band assignment applied in each.

The planning of the training phases is similar problem to the one explained in Sec. VI-B for the MOD approach. The main difference is that $\mathcal{U}$ includes all ADPs and all JDPs. The assumption remains that a certain set of assignments $\mathcal{X}_{\text{train}} = \{\mathbf{X}_i \mid i = 1, \ldots, |\mathcal{X}_{\text{train}}|\}$ is possible during training and for each assignment i, a subset of measurements from $\mathcal{U}$, $\mathcal{T}_i$, can

TABLE I: Parameter values used in simulation

Parameter	Value	Parameter	Value
Noise power	-146 dBm	w_I	200 kHz
P_I	14 dBm	N_I	3
P	14 dBm	R_I	1
R	3	Packet size incum.	200 B
N	3	T_I	$200\ B/w_I$ s
Packet size IoT	20 B	B	6
w	600 Hz	τ	10 dB
T	$20\ B/w$ s	T_{train}	10 min
W	200 kHz	M	3
Λ_{IoT}	50 km^{-2}	Λ_I	50 km^{-2}

be collected. After, applying a similar analysis as in Sec. VI-B, the planning of the training phases can be posed as a set cover problem instead of a partition set cover problem. Likewise, the planning can be solved using the greedy approach described in Algorithm 2. Set cover solutions that are more optimal than a greedy approach can be applied and are abundant in literature. For simplicity, we implement the greedy solution since evaluating different set cover solution algorithms would be beyond the scope of this paper.

B. Comparison to model-based BS placement and band assignment

The training process of the MEAS approach is similar to the training process of the MOD approach, however, there are some notable differences. First, the MEAS approach requires all locations in $\{1, \ldots, B + C\}$ to be covered by either previously installed BSs or temporary BSs, since ADPs and JDPs are estimated through direct measurements rather than through modelling. On the other hand, the MOD approach only requires the ADP and JDP to be estimated for a subset of BS locations, and then uses modeling to predict ADP and JDP on the remainder of the locations. This means that fewer temporary BSs need to be installed at candidate locations for optimal infrastructure planning. For these reasons, the MEAS approach is more appropriate when $\Delta B = 0$, i.e., when only performing frequency assignment, or when the number of candidate locations C is small. Moreover, even if $\Delta B = 0$, if B is large, the MEAS approach will suffer from an extended training time, since $\binom{B}{2}$ JDP values need to be estimated, whereas MOD approach generally requires less than $\binom{B}{2}$ JDP measurements, as we discussed in VI-B and VI-A. When B is large or when C is large and $\Delta B > 0$, MEAS approach is not practical, however it still serves as a useful benchmark in our simulation results, since it is an upper bound for the MOD approach in terms of the PDP.

The MOD approach has a lower training complexity and is more practical than the MEAS approach, however, the incumbent network and the IoT network are expected to match the system model explained in Sec. III and additional channel modelling assumptions introduced in Sec. V. On the other hand, the MEAS approach only has to meet the modelling assumptions in Sec. III-A.

VIII. SIMULATION RESULTS

A. Simulation environment

We simulate a disk area of radius 10 km with BSs, UNB devices and incumbent devices randomly distributed over the area. In each Monte Carlo (MC) simulation, the BS locations are sampled from a uniform distribution. Similarly, in each MC simulation, the UNB devices are sampled from a homogeneous Poisson Point process (HPPP) with a density Λ_{IoT}, therefore, the expected number of devices across all Monte Carlo simulations is $\Lambda_{IoT}|\mathcal{A}|$, where $|\mathcal{A}|$ is the area size of the simulation area. The locations of incumbent devices are generated from a HPPP with density Λ_I. Furthermore, the placement of current BSs and candidate locations is also randomized. The distribution of incumbent devices across multiplexing bands is non-uniform. Let β_j be the operating band of the incumbent device j. The probability that an incumbent is assigned to operate in a multiplexing band m during an MC iteration is a non-uniform discrete PMF $Pr(\beta_j = m)\ \forall m$. The PMF $Pr(\beta_j = m)\ \forall m$ distribution is randomly generated in each MC iteration. When planning the training procedure of MOD and MEAS, we assume that $\mathcal{X}_{\text{train}}$ contains all assignments such that each band is covered by at least $\lfloor \frac{B}{M} \rfloor$ BSs. Each MC iteration has a training stage for the MOD or MEAS approach. The training stage lasts T_{train} minutes and is split equally across each training phase. After the training stage ends and the BS band assignment and BS placement are completed, the network runs for 1 h to collect the performance metrics. The results are averaged over 300 MC iterations.

Unless otherwise stated, we use the simulation parameters given in Table I, where the UNB network emulates the Sigfox network with US specifications [13]. The temporal traffic generation assumes that each device sends three unique packets per hour. For interfering incumbents, we consider specifications similar to LoRa IoT devices [26] and assume they have a similar temporal traffic generation characteristics as the UNB network. We note that the noise is not ignored, and is equal to the thermal noise power at room temperature across bandwidth of w Hz.

B. BS band assignment results

The following algorithms are evaluated only for BS assignment to bands in terms of the packet decoding probability (PDP):

1) Random assignment: The BSs are randomly assigned to bands with the constraint that each band is covered by at least $\lfloor \frac{B}{M} \rfloor$ BSs.
2) Maximum BS separation heuristic: The BSs are placed by solving the optimization problem

$$\max_{\mathbf{X}} \quad \sum_m \left(\sum_b \sum_{v<b} [\mathbf{X}]_{b,m}[\mathbf{X}]_{v,m} d_{b,v} \right)$$

$$\text{s.t.} \quad \text{(P1.2), (P1.3), (P1.4)},$$

which ensures that BSs that are placed on the same band have maximum separation and that new BSs are added such that separation from the current BSs is maximized. This heuristics requires only knowledge of the coordinates of the current BSs as well as the candidate locations for BSs. Maximizing the separation ensures that the BSs are not decoding packets from common transmissions, which may increase the network
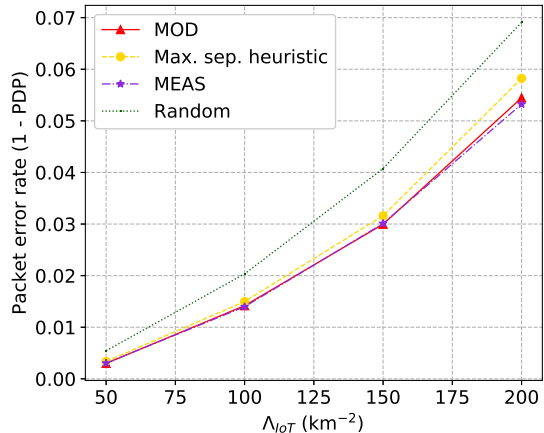
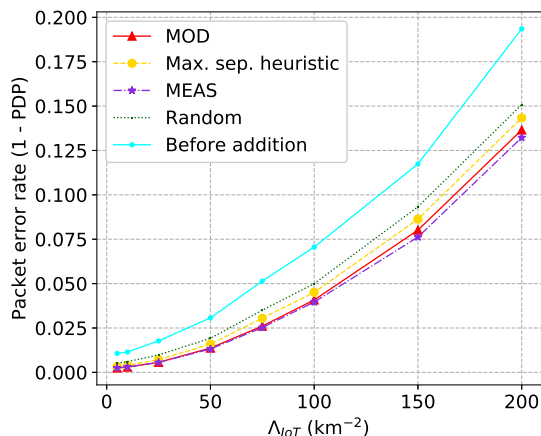Fig. 6: The PDP of the assignment algorithms with respect to the density of the IoT devices in the network.
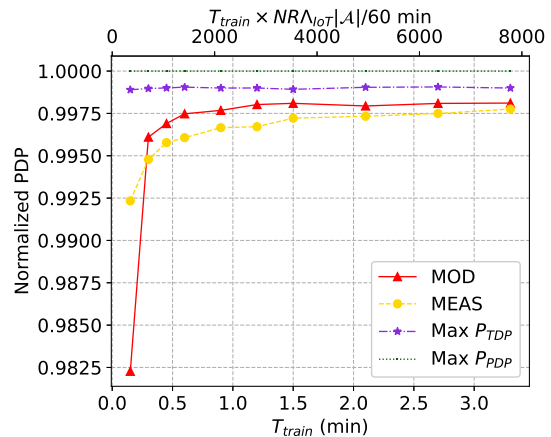

Fig. 7: The performance of the MOD and MEAS approach with respect to the length of the training time T_{train}. The PDP is normalized by the PDP of the theoretical P_{PDP}-maximizing assignment.
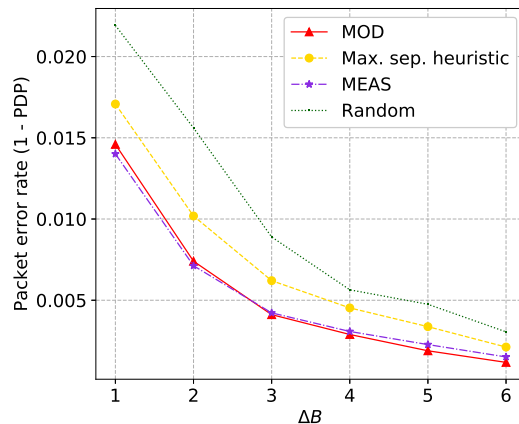

Fig. 8: The PDP of the tested placement and assignment algorithms with respect to the density of the IoT devices in the network.

Fig. 9: The improvement in PDP after ΔB new BSs are added to the network for different placement algorithms.

PDP. This heuristic may not always work, as UNB LPWANs require diversity in terms of BS reception when the decoding probability is low.

3) MOD assignment: No additional temporary BSs $\hat{B}$ are used in training. Furthermore, we set $S_m = 10 \ \forall m$.

4) MEAS assignment.

5) Optimal assignment that maximizes the probability of decoding a transmission, P_{TDP}. The optimal assignment for maximum transmission decoding rate can be obtained in simulation. We rerun every MC realization, with its particular sequence of pseudo-random events, for all possible M^B assignments, and select the assignment that gives the highest average packet or transmission decoding rate.

6) Optimal assignment that maximizes the probability of decoding a packet, P_{PDP}. The solutions are obtained through the same approach as for optimal P_{TDP} assignment.

First, we evaluate the performance of the algorithms with respect to the density of IoT devices. Here, we set the number of BSs to $B = 12$ and use high IoT device density, since with fewer BSs and low density, the band assignment problem becomes simpler and the gap between the heuristics and proposed solutions is smaller. The results are shown in Fig. 6. Typically, the PDP in real UNB networks such as Sigfox is $>95\%$, therefore our simulation setting is realistic in that regard. Random assignment performs significantly worse than either of our proposed approaches (MOD or MEAS). With assignment obtained using either of the proposed algorithms, the BS infrastructure can support around 25 more devices per km^2 in high density scenarios while maintaining the same PDP as random assignment, and around 5 more devices compared to the maximum separation heuristic.

We also analyze the dependence of the MOD and MEAS approach on the length of the training T_{train}. This is important because, when training is performed, the BSs need to move

away from their optimal band assignment which may lower the current PDP and so the training phase should be as short as possible. The results are shown in Fig. 7. We included a scaled axis, where T_{train} is scaled by $NR\Lambda_{\text{IoT}}|\mathcal{A}|/60$ min, which is equivalent to the expected number of IoT transmissions in $\mathcal{A}$ over time T_{train}. The higher the number of transmissions received by the BSs, the more accurate will the estimates of ADP and JDP be. From Fig. 7, it can be observed that the MOD approach requires less training time to achieve its highest PDP compared to the MEAS approach. This is because in the training stage of the MOD approach fewer parameters need to be estimated than in the training stage of the MEAS approach and the number of training phases $|\mathcal{L}|$ for the MOD approach is smaller. Since T_{train} is equally split between the $|\mathcal{L}|$ training phases, each training phase lasts longer in MOD approach and so the parameters estimated during each training phase contain less error. Moreover, the MOD and MEAS approach have a very similar performance to the theoretical best assignment that maximizes P_{TDP}. That indicates that the objective function in (P3) is a tight lower bound to the objective function in (P2). Furthermore, the theoretical best assignment that maximizes P_{TDP} is very close in performance to the theoretical best assignment that maximizes P_{PDP}. This indicates that the objective function in (P2) is a tight lower bound to the objective function in (P1).

C. BS band assignment and placement results

The following algorithms are evaluated for both BS placement and BS assignment to bands in terms of the PDP:

1) Random placement with optimized band assignment: The ΔB BSs are randomly placed within the C candidate locations. With the placement of ΔB BSs decided, the band assignment of $B + \Delta B$ is done using the MEAS approach.

2) Maximum BS separation heuristic: As described in Sec. VIII-B

3) MOD assignment: No additional temporary BSs $\hat{B}$ are used in training and $S_m = 10 \; \forall m$.

4) MEAS assignment: This approach is not practical for placement of BSs since it involves placing a temporary BS at each of the C candidate locations, but we include it as a benchmark for the MOD approach.

The theoretical optimal assignment for maximum packet decoding rate or maximum transmission decoding rate cannot be obtained numerically in simulation as with BS assignment in Sec. VIII-B since searching for the optimal assignment and placement exhaustively is intractable due to the number of possible solutions for BS placement and assignment.

The number of candidate locations is $C = 30$ and their locations are uniformly and randomly distributed in each MC iteration. The number of new BSs being installed is $\Delta B = 1$. We evaluate the performance with respect to the density of IoT devices and the results are shown in Fig. 8. Before addition of the BSs the frequency assignment of BSs is obtained using the MEAS approach. Putting additional BSs significantly improves the PDP, however, with optimized placement of the new BSs, additional improvements can be

achieved. Using the MOD approach additional 20 devices per km^2 can be supported while maintaining a similar PDP as with random placement, and around 5 more compared to the maximum separation heuristic. Furthermore, the MOD approach is similar in PDP to the MEAS approach which is a good demonstration of the practical usefulness of the MOD approach. As a reminder, the MEAS approach results are obtained by placing temporary BSs at each of the C candidate locations to estimate ADPs and JDPs while the MOD approach predicts the ADPs and JDPs at candidate locations using only the measurements from the currently installed B BSs. Despite of this, the PDP obtained when using MOD placement and assignment is the same as the PDP when using the MEAS placement and assignment. In Fig. 9 we increase the number of candidate positions to $C = 50$ and vary ΔB from 1 to 6. As ΔB increases, the difference between random and optimized placement becomes smaller. Therefore, optimized placement becomes more important when fewer new BSs can be afforded to be installed in the IoT network. Furthermore, we can see that the proposed solutions can achieve the same PDP as the random heuristic with two fewer BSs, and one fewer BS compared to the maximum BS separation heuristic. Additionally, we observe that MOD approach outperforms the MEAS approach for some ΔB, which can be attributed to the MEAS approach not having sufficient training time to estimate ADP and JDP accurately enough.

IX. CONCLUSIONS

In this work, we have addressed BS infrastructure management and expansion in multi-band UNB IoT networks. Namely, we devise two algorithms for maximization of packet delivery through optimal placement of BSs and optimal assignment of BSs to multiplexing bands. We derive a model-based and a measurement-based algorithm. The model-based approach has a simpler training stage than the measurement-based algorithm, however, it is based on certain assumptions that the IoT network and the environment should match. In the development of the model-based algorithm, we derive models of average decoding probability and joint decoding probability using stochastic geometry which we verify through simulation. Furthermore, our assignment and placement algorithms offer significant improvement in packet decoding probability over baseline approaches and in certain cases closely match the theoretical best performance. Future work will focus on optimizing the access policy of IoT devices in a multi-band UNB network and will also consider collaborative decoding of transmission by several BSs through fusion of their received signals.

APPENDIX

A. Proof of Proposition 1

Since we assume $\Pr(\beta(r) = m)$ is constant across m, we can rewrite the transmission decoding probability as

$P_{\text{TDP}}(\mathbf{X}) = \sum_m \Pr\left(\bigcup_b [\mathbf{X}]_{b,m} \, \text{SINR}_{b,m}(r) > \tau\right)$. Using the exclusion-inclusion principle we can express this function as:

$$P_{\text{TDP}}(\mathbf{X}) = \sum_m \left(\sum_b \Pr\left([\mathbf{X}]_{b,m} \, \text{SINR}_{b,m}(r) \geq \tau\right) - \right.$$

$$\sum_{v < b} \Pr\left([\mathbf{X}]_{b,m} \, \text{SINR}_{b,m}(r) \geq \tau, [\mathbf{X}]_{v,m} \, \text{SINR}_{v,m}(r) \geq \tau\right) +$$

$$\left. \dots + (-1)^{M-1} \Pr\left(\bigcap_{b=1}^{B} [\mathbf{X}]_{b,m} \, \text{SINR}_{b,m}(r) \geq \tau\right)\right) \quad (13)$$

The above sum can be truncated up to the second-order terms

$$P_{\text{TDP}}(\mathbf{X}) \approx \sum_m \left(\sum_b [\mathbf{X}]_{b,m} \Pr\left(\text{SINR}_{b,m}(r) \geq \tau\right) - \right.$$

$$\left. \sum_{b < v} [\mathbf{X}]_{b,m} [\mathbf{X}]_{v,m} \Pr\left(\text{SINR}_{b,m}(r) \geq \tau, \text{SINR}_{v,m}(r) \geq \tau\right)\right)$$

$$(14)$$

The approximation above is a concave function and is also a lower bound on the objective function in (P2). The latter follows from Bonferroni inequalities, since we are taking a second order approximation of the expression in (13) [27]. In (14), $Pr\left(\text{SINR}_{b,m}(r) \geq \tau\right) \equiv \mathbb{E}_r\{A_{b,m}(r)\}$ and $Pr\left(\text{SINR}_{b,m}(r) \geq \tau, \text{SINR}_{v,m}(r) \geq \tau\right) \equiv \mathbb{E}_r\{A_{b,m}(r)A_{v,m}(r)\}$. The objective function can be proven to be concave by calculating the Hessian of the expression over $\text{vec}(\mathbf{X})$. The Hessian matrix will be a symmetric block diagonal matrix $\text{diag}(\mathbf{A}_1, \dots, \mathbf{A}_M)$, where each matrix $\mathbf{A}_m$ is a covariance matrix $\mathbb{E}_r\{\mathbf{a}(r)\mathbf{a}(r)^T\}$, of a random vector $\mathbf{a}(r) = [A_{1,1}(r), A_{1,2}(r), A_{B+C,M}(r)]^T$. Since covariance matrices are known to be positive semi-definite, including their sampled versions, the Hessian matrix is negative semi-definite and therefore the expression in (14) is concave.

B. Proof of Theorem 1

The complementary cumulative distribution (CCDF) of $\text{SINR}_{i,b,m}$ is

$$\Pr(\text{SINR}_{i,b,m} \geq \tau \mid p_{i,b})$$

$$\overset{(a)}{=} \Pr\left(h_{i,b} \geq s_{i,b}\left(\sum_{j \in \Phi} f_j p_{j,b}^{-\alpha} + \sum_{j' \in \Phi_{I,m}} \hat{P}_I f_{j'} p_{j',b}^{-\alpha}\right)\right)$$

$$\overset{(b)}{=} \exp\left(s_{i,b}\left(\sum_{j \in \Phi} f_j p_{j,b}^{-\alpha} + \sum_{j' \in \Phi_{I,m}} \hat{P}_I f_{j'} p_{j',b}^{-\alpha}\right)\right)$$

where, to obtain (a), we have assumed that the $\text{SINR}_{i,b,m}$ is dominated by interference over the noise (we set $\hat{P}_N \to 0$) and set $s_{i,b} = \tau p_{i,b}^\alpha$; (b) follows from h_i being modelled as $\text{Exp}(1)$. Using a similar derivation as in [2, Eq. 39], we can arrive at the expression

$$\Pr(\text{SINR}_{i,b,m} \geq \tau \mid p_{i,b}) = \exp\left(-\epsilon_m \tau^{\frac{2}{\alpha}} p_{i,b}^2\right),$$

where $\epsilon_m = \pi\left(\lambda + \hat{P}_I^{\frac{2}{\alpha}} \lambda_{I,m}\right) \frac{2\pi/\alpha}{\sin(2\pi/\alpha)}$. The above expression provides us with a complimentary CDF (CCDF) of the

$\text{SINR}_{i,b,m}$ for a device i with a given separation between the receiver b and IoT device j, $p_{i,b}$. Note that in the derivation of the expression, we assumed that interference comes from IoT and incumbent devices that extend beyond $\mathcal{A}$ such that $\max_{j \in \Phi_{I,m}} p_{j,b} \to \infty$ and $\max_{j' \in \Phi_m} p_{j',b} \to \infty$. This approximation is reasonable if the BS b is very far away from furthest away devices that are causing interference.

C. Proof of Theorem 3

The distance of the device i to BS b is $r_{i,b}^2 = r_i^2 + \frac{d_{b,v}^2}{4} + p_i d_{b,v} \cos(\theta_i)$ and the distance to BS v is $r_{i,v}^2 = r_i^2 + \frac{d_{b,v}^2}{4} - p_i d_{b,v} \cos(\theta_i)$. We define $s_{i,b} = \tau p_{i,b}^\alpha$ and $s_{i,v} = \tau p_{i,v}^\alpha$. The conditional joint decoding probability is then

$$\Pr\left(\text{SINR}_{i,b,m} \geq \tau, \text{SINR}_{i,v,m} \geq \tau \mid i\right)$$

$$\overset{(a)}{=} \exp\left(s_{i,b}\left(\sum_{j \in \Phi_{IoT}} f_j p_{j,b}^{-\alpha} + \sum_{j' \in \Phi_{I,m}} \hat{P}_I f_{j'} p_{j',b}^{-\alpha}\right) + \right.$$

$$\left. s_{i,v}\left(\sum_{j \in \Phi_{IoT}} f_{j,v} p_{j,v}^{-\alpha} + \sum_{j' \in \Phi_{I,m}} \hat{P}_I f_{j',v} p_{j',v}^{-\alpha}\right)\right)$$

$$\overset{(b)}{=} \exp\left(\lambda \int_{\phi_j} \int_{p_j} \left(p_j - \frac{p_j}{\left(1 + s_{i,b} p_{j,b}^{-\alpha}\right)\left(1 + s_{i,v} p_{j,v}^{-\alpha}\right)}\right)\right.$$

$$\left. + \lambda_{I,m} \int_{\phi_{j'}} \int_{p_{j'}} \left(p_{j'} - \frac{p_{j'}}{\left(1 + s_{i,b} p_{j',b}^{-\alpha}\right)\left(1 + s_{i,v} p_{j',v}^{-\alpha}\right)}\right)\right)$$

where (a) follows from the assumption that the fading RVs at BS b and v are independent; (b) follows from the expression for Laplace–Stieltjes transform of the Rayleigh RV and the PGFL of a HPPP. Evaluating the integral to a closed form expression is not possible to the best of our knowledge, therefore we adopt an approach similar to [28] to obtain the upper bound on the conditional joint outage probability.

First, we will derive the expression for conditional decoding probability over the decoding probability of one of the BSs $\frac{\Pr(\text{SINR}_{i,v,m} \geq \tau \mid \text{SINR}_{i,b,m} \geq \tau)}{\Pr(\text{SINR}_{i,v,m} \geq \tau)} = \frac{\Pr(\text{SINR}_{i,b,m} \geq \tau, \text{SINR}_{i,v,m} \geq \tau)}{\Pr(\text{SINR}_{i,b,m} \geq \tau)\Pr(\text{SINR}_{i,v,m} \geq \tau)}$. Following the same derivation as in the expression above we can arrive at the product of two integrals:

$$\frac{\Pr(\text{SINR}_{i,v,m} \geq \tau \mid \text{SINR}_{i,b,m} \geq \tau)}{\Pr(\text{SINR}_{i,v,m} \geq \tau)} =$$

$$\exp\left(\lambda \int_{\phi_j} \int_{p_j} \left(\frac{s_{i,b} p_{j,b}^{-\alpha}}{1 + s_{i,b} p_{j,b}^{-\alpha}}\right)\left(\frac{s_{i,v} p_{j,v}^{-\alpha}}{1 + s_{i,v} p_{j,v}^{-\alpha}}\right)\right) \times$$

$$\exp\left(\lambda_{I,m} \int_{\phi_j'} \int_{p_j'} \left(\frac{s_{i,b} p_{j',b}^{-\alpha}}{1 + s_{i,b} p_{j',b}^{-\alpha}}\right)\left(\frac{s_{i,v} p_{j',v}^{-\alpha}}{1 + s_{i,v} p_{j',v}^{-\alpha}}\right)\right)$$

If we let $d_{b,v} \to 0$, we can obtain an upper bound on the expression above and then evaluate the resulting integral to obtain:

$$\frac{\Pr(\text{SINR}_{i,v,m} \geq \tau \mid \text{SINR}_{i,b,m} \geq \tau)}{\Pr(\text{SINR}_{i,v,m} \geq \tau)} \leq e^{\left(\left(-\frac{2}{\alpha}+1\right)\epsilon_m \tau^{\frac{2}{\alpha}} r^2\right)}.$$

Using the above expression we can then also derive the upper bound to the joint decoding probability:

$$\Pr\left(\text{SINR}_{i,b,m} \geq \tau, \text{SINR}_{i,v,m} \geq \tau\right) = \Pr(\text{SINR}_{i,b,m} \geq \tau)$$
$$\times \Pr(\text{SINR}_{i,v,m} \geq \tau) \frac{\Pr(\text{SINR}_{i,v,m} \geq \tau \mid \text{SINR}_{i,b,m} \geq \tau)}{\Pr(\text{SINR}_{i,v,m} \geq \tau)} \leq$$
$$\exp\left(\left(-\frac{2}{\alpha} - 1\right)\epsilon_m \tau^{\frac{2}{\alpha}} r^2\right) \exp\left(-\frac{1}{2}\epsilon_m \tau^{\frac{2}{\alpha}} d_{b,v}^2\right).$$

REFERENCES

[1] U. Raza, P. Kulkarni, and M. Sooriyabandara, "Low power wide area networks: An overview," *IEEE Commun. Surv. Tutorials*, vol. 19, no. 2, pp. 855–873, 2017.

[2] G. Hattab and D. Cabric, "Spectrum sharing protocols based on ultra-narrowband communications for unlicensed massive IoT," in *2018 IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN)*, pp. 1–10, IEEE, 2018.

[3] G. Hattab, P. Popovski, and D. Cabric, "Spectrum sharing for massive access in ultra-narrowband IoT systems," *IEEE J. Sel. Areas Commun.*, 2020.

[4] E. Krijestorac, G. Hattab, P. Popovski, and D. Cabric, "Band assignment in ultra-narrowband (UNB) systems for massive IoT access," *arXiv preprint arXiv:2010.04307*, 2020.

[5] X. Ling and K. L. Yeung, "Joint access point placement and channel assignment for 802.11 wireless LANs," *IEEE Trans. Wireless Commun.*, vol. 5, no. 10, pp. 2705–2711, 2006.

[6] S. Uygungelen, G. Auer, and Z. Bharucha, "Graph-based dynamic frequency reuse in femtocell networks," in *2011 IEEE 73rd Veh. Tech. Conf. (VTC Spring)*, pp. 1–6, IEEE, 2011.

[7] L. Narayanan and S. M. Shende, "Static frequency assignment in cellular networks," *Algorithmica*, vol. 29, no. 3, pp. 396–409, 2001.

[8] H. S. Chae, J. Gu, B.-G. Choi, and M. Y. Chung, "Radio resource allocation scheme for device-to-device communication in cellular networks using fractional frequency reuse," in *The 17th Asia Pacific Conf. Comm.*, pp. 58–62, IEEE, 2011.

[9] E. Ahmed, A. Gani, S. Abolfazli, L. J. Yao, and S. U. Khan, "Channel assignment algorithms in cognitive radio networks: Taxonomy, open issues, and challenges," *IEEE Commun. Surv. Tutorials*, vol. 18, no. 1, pp. 795–823, 2014.

[10] B. Ousat and M. Ghaderi, "Lora network planning: Gateway placement and device configuration," in *2019 IEEE Int. Congr. Internet Things (ICIOT)*, pp. 25–32, IEEE, 2019.

[11] N. Matni, J. Moraes, D. Rosário, E. Cerqueira, and A. Neto, "Optimal gateway placement based on fuzzy C-means for low power wide area networks," in *2019 IEEE Latin-American Conf. Comm. (LATINCOM)*, pp. 1–6, IEEE, 2019.

[12] H. Tian, M. A. Weitnauer, and G. Nyengele, "Optimized gateway placement for interference cancellation in transmit-only LPWA networks," *Sensors*, vol. 18, no. 11, p. 3884, 2018.

[13] "Technical overview - Sigfox USA," Jun 2020.

[14] "WAVIot NB-Fi base station datasheet," Jun 2020.

[15] "Telensa technical specifications," May 2021.

[16] J. Finnegan and S. Brown, "A comparative survey of LPWA networking," *arXiv preprint arXiv:1802.04222*, 2018.

[17] Y. Zhong, G. Mao, X. Ge, and F.-C. Zheng, "Spatio-temporal modeling for massive and sporadic access," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 3, pp. 638–651, 2020.

[18] B. Vejlgaard, M. Lauridsen, H. Nguyen, I. Z. Kovács, P. Mogensen, and M. Sorensen, "Interference impact on coverage and capacity for low power wide area IoT networks," in *2017 IEEE Wireless Comm. Netw. Conf. (WCNC)*, pp. 1–6, IEEE, 2017.

[19] M. Jünger, T. M. Liebling, D. Naddef, G. L. Nemhauser, W. R. Pulleyblank, G. Reinelt, G. Rinaldi, and L. A. Wolsey, *50 Years of integer programming 1958-2008: From the early years to the state-of-the-art*. Springer Science & Business Media, 2009.

[20] M. Haenggi, *Stochastic geometry for wireless networks*. Cambridge University Press, 2012.

[21] J. M. Steele, "Le Cam's inequality and poisson approximations," *The American Mathematical Monthly*, vol. 101, no. 1, pp. 48–54, 1994.

[22] S. Sigfox, "Sigfox technical overview," 2017.

[23] J. J. Moré, "The Levenberg-Marquardt algorithm: implementation and theory," in *Numerical analysis*, pp. 105–116, Springer, 1978.

[24] P. Slavik, "Improved performance of the greedy algorithm for partial cover," *Information Processing Letters*, vol. 64, no. 5, pp. 251–254, 1997.

[25] S. K. Bera, S. Gupta, A. Kumar, and S. Roy, "Approximation algorithms for the partition vertex cover problem," *Theor. Comput. Sci.*, vol. 555, pp. 2–8, 2014.

[26] "What is LoRaWAN® specification," May 2021.

[27] C. Bonferroni, "Statistical class theory and probability calculation," *Publ. Higher Inst. of Econ. Commercial Sci. Florence*, vol. 8, pp. 3–62, 1936.

[28] R. K. Ganti and M. Haenggi, "Spatial and temporal correlation of the interference in ALOHA ad hoc networks," *IEEE Commun. Lett.*, vol. 13, no. 9, pp. 631–633, 2009.

Enes Krijestorac received a B.S. degree in Electrical Engineering from New York University, Abu Dhabi in 2018, graduating *summa cum laude*. He received his M.S. degree from University of California, Los Angeles, in 2021, also in Electrical Engineering. He is currently pursuing a Ph.D. degree at the University of California, Los Angeles, US. His research interests include UAV assisted wireless communication, wireless channel prediction and sensing, distributed computing over wireless channels and machine learning. He was a recipient of the Dean's Scholar Award and the PhD Preliminary Exam Fellowship at the UCLA Electrical and Computer Engineering Department.

Ghaith Hattab received his PhD degree from the University of California, Los Angeles (UCLA). He received his M.Sc degree from Queen's University, Kingston ON, Canada, and his B.Sc degree from American University of Sharjah (AUS), Sharjah, United Arab Emirates (UAE), all in Electrical Engineering. He has held research internships at Nokia Bell Labs, IL, USA, and Qualcomm Research, CA, USA. He is currently with Apple Inc., USA. His research interests include system-level analysis, optimization, and signal processing for wireless communications, with emphasis on channel state feedback, spectrum sharing, and coexistence of wireless systems. He was the recipient of the prestigious Presidential Cup awarded by His Highness Sheikh Dr. Sultan Al Qassimi, the IEEE Kingston Section M.Sc Research Excellence Award, and the departmental fellowship at UCLA.

Petar Popovski (Fellow, 2016) is a Professor at Aalborg University, where he heads the section on Connectivity and a Visiting Excellence Chair at the University of Bremen. He received his Dipl.-Ing and M. Sc. degrees in communication engineering from the University of Sts. Cyril and Methodius in Skopje and the Ph.D. degree from Aalborg University in 2005. He received an ERC Consolidator Grant (2015), the Danish Elite Researcher award (2016), IEEE Fred W. Ellersick prize (2016), IEEE Stephen O. Rice prize (2018), Technical Achievement Award from the IEEE Technical Committee on Smart Grid Communications (2019), the Danish Telecommunication Prize (2020) and Villum Investigator Grant (2021). He was a Member at Large at the Board of Governors in IEEE Communication Society 2019-2021. He is currently an Editor-in-Chief of IEEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. He also serves as a Vice-Chair of the IEEE Communication Theory Technical Committee and the Steering Committee of IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING. Prof. Popovski was the General Chair for IEEE SmartGridComm 2018 and IEEE Communication Theory Workshop 2019. His research interests are in the area of wireless communication and communication theory. He authored the book "Wireless Connectivity: An Intuitive and Fundamental Guide", published by Wiley in 2020.

Danijela Cabric is Professor in Electrical and Computer Engineering at University of California, Los Angeles. She earned MS degree in Electrical Engineering in 2001, UCLA and Ph.D. in Electrical Engineering in 2007, UC Berkeley, Dr. Cabric received the Samueli Fellowship in 2008, the Okawa Foundation Research Grant in 2009, Hellman Fellowship in 2012 and the National Science Foundation Faculty Early Career Development (CAREER) Award in 2012 and the Qualcomm Faculty award in 2020 and 2021. She served as an Associate Editor of IEEE Transactions of Cognitive Communications and Networking, IEEE Transactions of Wireless Communications, IEEE Transactions on Mobile Computing and IEEE Signal Processing Magazine, and IEEE ComSoc Distinguished Lecturer. Her research interests are millimeter-wave communications, distributed communications and sensing for Internet of Things, and machine learning for wireless networks co-existence and security. She is an IEEE Fellow.