

Constraint Hubs Deployment for Efficient Machine-Type Communications

Miloud Bagaa^{ib}, *Member, IEEE*, Tarik Taleb^{ib}, *Senior Member, IEEE*,
Ali Chelli^{ib}, *Member, IEEE*, and Hamed Hellaoui

Abstract—Massive Internet of Things (mIoT) is an important use case of 5G. The main challenge for mIoT is the huge amount of uplink traffic as it dramatically overloads the radio access network (RAN). To mitigate this shortcoming, a new RAN technology has been suggested, where small cells are used for interconnecting different devices to the network. The use of small cells will alleviate congestion at the RAN, reduce the end-to-end (E2E) delay, and increase the link capacity for communications. In this paper, we devise three solutions for deploying and interconnecting small cells that would handle mIoT traffic. A realistic physical model is considered in these solutions. The physical model is based on a composite fading channel that captures path loss, fast fading, shadowing, and interference to derive the signal-to-interference-plus-noise ratio. The three solutions consider two conflicting objectives, namely the cost and the E2E delay for deploying and backhauling small cells. The first solution minimizes the cost while the second reduces the E2E delay. The third solution uses bargaining game theory for reducing both the cost and the E2E delay. The proposed solutions are evaluated through simulations. The obtained results demonstrate the efficiency of each solution in achieving its design goals.

Index Terms—Relay node placement, wireless sensor network, connectivity, SINR model.

I. INTRODUCTION

NOWADAYS, massive Internet of Things (mIoT) and Machine Type Communications (MTC) technologies have been gaining a lot of ground due to the tremendous number of applications where they can be applied. These applications are for example intelligent transportation services, smart cities and industrial monitoring systems. Recent studies claim that by 2020 more than 50 billions of MTC devices would be connected to the network for serving these applications.

Manuscript received March 8, 2018; revised August 6, 2018; accepted September 19, 2018. Date of publication October 9, 2018; date of current version December 10, 2018. This work was supported in part by 6Genesis and AoF CSN Projects under Grant 318927 and Grant 311654 and in part by the Aalto 5G meets Industrial Internet (5G@II) Project. The associate editor coordinating the review of this paper and approving it for publication was X. Zhou. (*Corresponding author: Miloud Bagaa.*)

M. Bagaa and H. Hellaoui are with the Department of Communications and Networking, School of Electrical Engineering, Aalto University, FI-00076 Aalto, Finland (e-mail: miloud.bagaa@aalto.fi; hamed.hellaoui@aalto.fi).

T. Taleb is with the Department of Communications and Networking, School of Electrical Engineering, Aalto University, FI-00076 Aalto, Finland, also with the Faculty of Information Technology and Electrical Engineering, Oulu University, 90014 Oulu, Finland, and also with the Department of Computer and Information Security, Sejong University, Seoul 05006, South Korea (e-mail: tarik.taleb@aalto.fi).

A. Chelli is with the Faculty of Engineering and Science, University of Agder, 4898 Grimstad, Norway (e-mail: ali.chelli@uia.no).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TWC.2018.2873293

Thereby, mIoT and MTC technologies are one of the important use-cases that should be enabled by the next fifth generation mobile networks. Indeed, the communications in 5G mobile networks are classified into three main categories namely [1]: *i*) mMTC (massive Machine Type Communications); *ii*) eMBB (enhanced Mobile Broad Band); *iii*) URLLC (Ultra Reliable and Low latency Communications).

The deployment of high number of MTC devices will create a high overhead on the cellular mobile network, which can be an obstacle facing 5G mobile network for achieving its objectives. Indeed, these MTC devices would create massive signaling overhead at different Evolved Packet System (EPS), i.e. at the Radio Access Network (RAN) and the Core Network (CN), due to the huge amount of data/control traffic generated by different MTC devices. From another perspective, most of the MTC devices are powered-on using only constrained power supply that limits their communication capabilities. Each device is unable to communicate with a far away base-station (i.e., eNodeB). Thanks to the new RAN technology, where small (nano and pico) cells would be used, the use of mIoT and MTC technologies can be enabled in 5G mobile networks. The operators will deploy a set of small cells (i.e., Hubs), where each one ensures the communication for a limited number of users and MTC devices. Therefore, the signaling in the network will be mitigated, as well as the MTC devices will be attached to the closest Hubs, which extends their lifetime.

In this paper, we are interested in the optimal positioning of different Hubs for interconnecting MTC devices with respect to low cost and end-to-end delay. The first objective, low cost, aims at reducing the price related to deploying Hub nodes along with the used communication technology. The end-to-end delay refers to the delay between the MTC devices and the anchor, through the Hub nodes. As depicted in Fig. 1, a realistic scenario is considered in this paper, where each of the Hubs has limited candidate positions to be placed [2]. We give also more flexibility for interconnecting the placed Hubs using different communication technologies for reducing the cost and the end-to-end delay. While different wired or wireless technologies can be used to interconnect Hub nodes together or with anchor nodes, only wireless technology could be used to connect MTC devices to the network via a Hub/anchor node. In order to model the wireless link between an MTC device and a Hub/anchor node, we consider a realistic physical model that takes into account all the impacts of propagation on the transmitted signal. The physical model is based on a composite fading channel that captures path loss,

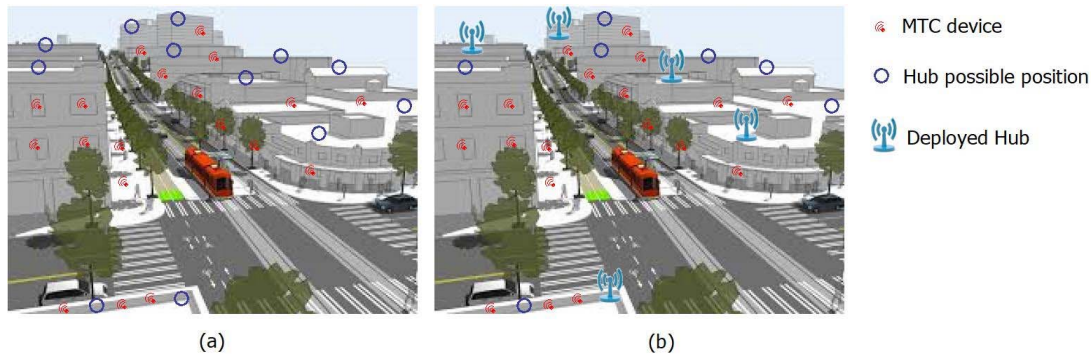


Fig. 1. Envisioned architecture.

fast fading, shadowing and interference to derive the signal-to-interference-plus-noise ratio (SINR). The cumulative density function (CDF) of the latter is calculated and used to determine the outage probability –for every communication links and in each directions– as a function of the receiver’s sensitivity threshold. The outage probability is then used to calculate the end-to-end delay between MTC devices with different Hubs positions. In this paper besides the reduction of the costs (i.e, cost of Hubs and the interconnecting technologies), we take also into account the reduction of end-to-end delay. The achieved results demonstrate the efficiency of the proposed solutions in terms of cost, end-to-end delay and network life time.

The rest of the paper is organized as follows. An overview on the literature of the small cells network backhauling (Hubs interconnection) is provided in Section II. The problem formulation and the envisioned architecture have been presented in Section III, followed by the description of the proposed solutions in Section IV. Section V provides the simulation results of different solutions. Finally, Section VI concludes the paper.

II. RELATED WORKS

This section focuses on the review of existing works for small cells network backhauling that has attracted high attention during the last years. Authors in [3] have proposed two solutions for ensuring reliable and low cost backhaul network for existing small cells. Both solutions assume that the small cells are already existing and try to interconnect them to the root node (macro cell). These solutions are based on minimum spanning tree (MST) and the shortest path (TSP) for designing reliable backhaul infrastructure at low costs. Both solutions are fully based on 2.4 GHz Wi-Fi wireless technology band and aim primarily at choosing aggregator nodes that satisfy mesh topology characteristics. The use of mesh topology is to ensure the network reliability. Unfortunately, the proposed solutions did not investigate the impact of other wireless linking alternatives, whose characteristics may impact the chosen design and could enhance its cost-efficiency. A new framework has been proposed in [4] for backhauling the small cells networks by exploring the carrier aggregation of mmWave and conventional frequency bands. In the proposed framework, the small cells are backhauled using fiber optic broadband

and wireless communications by allocating their frequency resources via aggregated bands. The optimization problem was formulated as a matching game problem. A solution that solves this problem has been proposed for guaranteeing the stability and the required QoS. The proposed approach gives practical recommendations for designing small cells wireless backhaul, especially for cases where some of the small cells already have access to operator broadband services. The main limitation in their proposed solution is the cost optimization procedure that focused on bands selection for the same linking technology family (mmWave and sub-6 GHz bands), rather than making selection among a set of wireless technologies.

Authors in [5] proposed practical methods for small cell backhaul planning. In this paper two scenarios were considered: In the first use-case, they have considered only fiber optic broadband, whereas in the second use-case multiple-communications technologies were considered. For the first scenario, two methods for finding a cost-efficient links with a tree and ring topology, respectively, have been considered. For heterogeneous multi-technology backhaul networks, a mixed integer linear programming problem was proposed. This problem was not solved and was only approached by heuristics based on fiber-only solutions. Authors in [6] have studied the minimum cost cell planning problem in a heterogeneous network. Then, a solution has been proposed for interconnecting different users to the small cells, relays and macro cells while ensuring their requirements with minimal cost. In the proposed solution only the communications between the users, and different cells have been considered. Moreover, only one wireless communication technology has been used in this solution, whereby the interferences effect has been ignored. Authors in [7] have considered the problem of wireless backhauling for interconnecting different devices in the network. The authors assume that a macro cell consists of one macro cell base-station, a set of small cells base-stations and a set of gateways. The different gateways are interconnected using FTTC (Fibre-to-the-Curb) technologies. The small cells are interconnected through multi-hops connection to the gateways. The authors, besides analyzing the capacity of 5G ultra-dense network, propose a new algorithm, named MAN, that interconnects the small cells to different gateways.

All the works previously mentioned assume that the small cells and Hubs are already deployed in the network. Then, they

want to build an efficient small cell backhaul planning for interconnecting the different small cells to the root nodes (macro cells). In contrast to these works, some works in the literature have considered the placement of small cells and Hubs in the network. Authors in [8] have investigated the problem for finding the optimal number of small cells for enabling 5G ultra-dense cellular networks regarding backhaul capacity. The proposed framework uses Monte Carlo simulations for finding the optimal number of small cells that should be deployed in the network. The proposed framework focused on finding the optimal number of small cells rather than designing their connections.

Most of the aforementioned works have only taking into account the techniques for backhauling already deployed small cells, but did not consider the deployment of new small cells. The works that tackled small cells placements problems did not consider the interconnection of these cells to the macro cell base-station using different wired and wireless communications technologies. Most of these works do not take into account network features, such as traffic, bandwidth and delay, when interconnecting different Hubs to the anchors. Moreover, the wireless interferences generated by different MTC devices are not considered in these works. In addition, these works did not take into account the cost of the links for interconnecting different small cells. In order to overcome the limitations of these works, we have proposed in this paper, a new framework that considers *i*) the deployment of both small cells and Hubs, and *ii*) the interconnections between them in a unified manner. The small cells and Hubs can be deployed in a limited set of positions. The proposed framework considers the interferences generated by different MTC devices, as well as the cost of links for interconnecting the different Hubs to the anchors. This framework consists of three different solutions that deploy and interconnect Hubs to the anchors using multiple communication technologies. These solutions aim at minimizing the cost by deploying the lowest number of Hubs using communication technologies with the cheapest prices, as well as minimizing the data latency when forwarding the data from MTC devices to the anchor nodes by using high quality links between different nodes.

III. PROBLEM FORMULATION AND NETWORK MODEL

A. Envisioned Architecture

As depicted in Fig. 1, we consider an area that contains a set of MTC devices, \mathcal{S} which have been already deployed by different users. We denote the set of candidate positions by \mathcal{Y} , where the Hubs can be placed. We denote by \mathcal{A} the set of existing eNodeBs in the network that are directly connected to the core network (CN) operator. \mathcal{A} will serve as data anchors for ensuring the Internet connection to different MTC devices. In the network, we should have at least one anchor node ($|\mathcal{A}| = 1$). In Fig. 1(a), the blue circles represent the possible positions of Hubs, whereas the red dots represent the MTC devices. In Fig. 1(b), the added nodes which are in cyan color represent the Hubs. As we have mentioned in the introduction, the MTC devices in \mathcal{S} can be far away from the anchor nodes. In order to interconnect \mathcal{S} to \mathcal{A} , this paper

suggests three solutions that augment the network by putting extra Hubs, called \mathcal{H} , in \mathcal{Y} .

As depicted in Fig. 1(b), a set of Hubs would be placed and interconnected using different technologies that ensure the required quality of services (QoS) with the lowest cost. Note that the Hubs operate in half duplex mode. In this paper, we do not take into account the network reliability, therefore each MTC device $s \in \mathcal{S}$ would be connected to the Internet using only one anchor node $A \in \mathcal{A}$. Formally, s can be directly connected to A or via multi-hops path by using a set of Hubs from \mathcal{H} . Each Hub in \mathcal{H} is responsible for forwarding the data from a subset of \mathcal{S} , as well as for forwarding the commands/requests from the data anchors \mathcal{A} to these devices. We denote by $\mathcal{P}_{s,A}$ the path that connects the MTC device $s \in \mathcal{S}$ to the anchor node $A \in \mathcal{A}$. $\mathcal{P}_{s,A}$ is constituted with a set of Hubs $H \subseteq \mathcal{H}$. For each $u \in \mathcal{P}_{s,A}$: *i*) we denote by $Suc(u) \in \mathcal{P}_{s,A}$ (the successor of u), the u 's one step neighbor towards A ; *ii*) While, we denote by $Pre(u) \in \mathcal{P}_{s,A}$ (the predecessor of u), the u 's one step neighbor towards s . MTC devices are designed to forward the generated data to dedicated servers, while they receive the commands and requests from those servers. Each Hub u will receive the data from $Pre(u)$, which can be an MTC device or another Hub node, and transmit the received data to its successor $Suc(u)$. Each Hub node $r \in \mathcal{P}_{s,A}$ keeps forwarding the data until it is received by the anchor A . The latter will forward the revived data to the dedicated servers. Moreover, the Hub node u receives the commands/requests from its $Suc(u)$ and forwards them to its $Pre(u)$ until their received by s .

While the MTC devices use only one wireless communication technology, such as LTE high frequency (i.e., 1.9 GHz – 2.6 GHz), to connect to the Hubs \mathcal{H} or anchors \mathcal{A} , \mathcal{H} can use multiple wired and wireless communication technologies to interconnect among themselves or to \mathcal{A} . Let \mathcal{L} denote the set of communication technologies that can be used for interconnecting different Hubs and anchor nodes. \mathcal{L} includes both technologies wired and wireless communications.

Each communication technology $\ell \in \mathcal{L}$ is characterized by a specified link capacity, communication range, end-to-end delay and cost that can be varied according to the positions of Hubs and anchor nodes. Table I summaries the different notations used in this paper. As mentioned in Table I, according to the Hub's candidate position $i \in \mathcal{Y}$ and another Hub's candidate position or an anchor node $j \in \mathcal{Y} \cup \mathcal{A}$, we define the following parameters for the communication technology ℓ : *i*) $\Gamma_{i,j}^\ell$ denotes the cost of deploying the technology ℓ between i and j . While the cost of wired communication is varied according to the positions of i and j , the wireless communication does not have any cost. For a wireless communication $\ell \in \mathcal{L}$, $\Gamma_{i,j}^\ell$ equals to 0. Meanwhile, for the wired communication, the longer the distance between i and j is, the higher the cost becomes; *ii*) $\pi_{i,j}^\ell$ denotes the link capacity between i and j using technology ℓ ; *iii*) $\rho_{i,j}^\ell$ denotes the end-to-end delay between i and j using technology $\ell \in \mathcal{L}$. Actually, each Hub node $i \in \mathcal{H}$ has a traffic load generated by: *i*) its attached MTC devices, and/or *ii*) the received messages from its predecessors Hubs. In contrast to the user devices, the uplink data traffic in MTC devices is more important than the downlink

TABLE I
SUMMARY OF NOTATIONS

Notation	Description
\mathcal{S}	The set of MTC devices in the network.
\mathcal{Y}	The set of Hubs candidate positions in the network.
\mathcal{A}	The set of anchor nodes in the network. Formally, these nodes have a connection with the core network.
\mathcal{H}	The set of Hubs in the network should be added to the network for connecting MTC devices to the anchor nodes.
\mathcal{L}	The set of technology that can be used for backhauling network nodes.
$\pi_{i,j}^\ell$	The link capacity between Hub nodes placed in positions $i \in \mathcal{Y}$ and $j \in \mathcal{Y}$ using technology $\ell \in \mathcal{L}$.
$\rho_{i,j}^\ell$	The end-to-end delay between Hub nodes placed in positions $i \in \mathcal{Y}$ and $j \in \mathcal{Y}$ using technology $\ell \in \mathcal{L}$.
τ	The cost of deploying one anchor node.
$\Gamma_{i,j}^\ell$	The cost of using technology ℓ between nodes $i, j \in \mathcal{S} \cup \mathcal{Y} \cup \mathcal{B}$.
λ_i	The amount of traffic generated by MTC device $i \in \mathcal{S}$.
Ψ_i	A variable that shows the amount of traffic should be aggregated and forwarded by a hub node $i \in \mathcal{Y}$.
Π_i	A variable that shows the link capacity required by a hub node $i \in \mathcal{Y}$ to handle the amount of received traffic without creating network congestion.
$\eta^\ell(i)$	The neighbors of node i if the technology ℓ is used. Formally, $\eta^\ell(i) \subseteq \mathcal{Y} \cup \mathcal{A}$.

data traffic. For a simple request or command a huge data would be generated and forwarded from these devices to the dedicated servers. For this reason, the proposed solutions are focusing on how to enhance the QoE of uplink data traffic with the lowest cost.

The problem herein, in the paper, consists of finding the optimal deployment strategy of Hub nodes in \mathcal{Y} for interconnecting the MTC devices to the anchor nodes. The placed Hubs ensure bidirectional communication between the MTC devices \mathcal{S} and the anchor nodes \mathcal{A} by conveying the traffic generated at the MTC devices to the anchor nodes and vice versa. A constrained Hubs placement problem is considered, where the possible positions at which the Hubs can be placed are limited. Note that the number of possible deployment positions \mathcal{Y} is larger than the number of Hubs \mathcal{H} that will be eventually introduced in the network. The optimal choice of Hub nodes positions should satisfy several objectives including: Firstly, a good QoS should be ensured by offering reliable communication link between the MTC devices and the anchor nodes. Secondly, we aim to interconnect the MTC devices to the Hubs that have a good link in order to prevent the re-transmissions, and thus extend the network lifetime. Thirdly, the number of deployed Hubs should be minimized and the cheapest communication technologies should be used in order to reduce as much as possible the solution cost. In order to further enhance the communication reliability, an automatic repeat request (ARQ) scheme is considered for forwarding the information. The ARQ scheme allows re-sending a packet until successful reception or a maximum number of re-transmissions M is reached. We assume that each MTC device $s \in \mathcal{S}$ has a transmission power P_s .

B. Communication Model

In the network under investigation, the MTC devices can be connected to the anchor node directly or via one or several Hub nodes. The communication between the MTC device and other nodes in the network (either an anchor node or a Hub node) uses only wireless technology. On the other hand Hub nodes are interconnected together and with the anchor node using either wireless or wired technologies. These wireless technologies include microwave, 802.11n and 802.11ac, while

the wired technologies include Optical-Ethernet1, Optical-Ethernet2, and Optical-Ethernet3. Note that for the wired technologies exact values of their characteristics such as the link capacity, the delay and the communication range can be found in the literature. Therefore, this section will focus only on the modeling of the wireless link connecting an MTC device to an anchor/Hub node.

Let us denote by u the transmitting node (the MTC device), and by v the receiving node (the Hub/anchor node). The channel gain between these two nodes is referred to as $\alpha_{u,v}$. This channel gain $\alpha_{u,v}$ takes into account the shadowing effect, the fast fading and the path loss. The channel gain $\alpha_{u,v}$ can be expressed as [9]

$$\alpha_{u,v} = f_{u,v} \sqrt{c s_{u,v} d_{u,v}^{-\eta}}, \quad (1)$$

where c is a constant, $d_{u,v}$ is the distance between the node u and the node v , $s_{u,v}$ is the shadow-fading factor which follows a log-normal distribution, and the symbol η denotes the path loss exponent. The fast fading is captured by $f_{u,v}$ which has a Nakagami block-fading distribution, where $f_{u,v}$ remains constant over one block¹ but changes independently from one block to another. The overall distribution of the channel gain $\alpha_{u,v}$ is the product of two random variables: *i*) a first random variable capturing the shadowing and following a log-normal distribution and *ii*) a second random variable accounting for the fast fading and following a Nakagami distribution.

In the proposed model, we take into consideration the interference from the other MTC devices in the environment on the receiving node v . The fading coefficient from node t to node v is referred to as $\alpha_{t,v}$. The received signal at a destination node, v , can be expressed as

$$y_v = \alpha_{u,v} \sqrt{P_u} x_u + \sum_{t \in \mathcal{N}, t \neq u,v} \alpha_{t,v} \sqrt{P_t} x_t + n_v, \quad (2)$$

where P_u and P_t are the transmission powers at node u and at node t , respectively, and \mathcal{N} is the set of all the transmitting nodes in the network. The symbols transmitted by node u and node t are referred to as x_u and x_t , respectively.

¹A block corresponds to the time duration necessary to send one packet.

The term n_v is a zero-mean additive white Gaussian noise with variance N_0 .

It is worth mentioning that in our channel model, we take into account all the physical phenomena that occur for a signal transmitted over the wireless channel. More specifically, we consider the impact of fast fading, path loss and shadowing on the desired signal and the interference signal. These facts result in a realistic physical model that can capture the behavior of real world wireless channels.

On the other hand, we model the fast fading and shadowing using generic and well-accepted models in the literature. In particular, we assume a Nakagami distribution and a log-normal distribution for the fast fading and the shadowing, respectively. Note that both distributions are well-accepted in the literature and have been shown to fit real world measurements for the wireless channel in various scenarios. The Nakagami distribution is quite generic and includes the Rayleigh and Rice distributions as special cases. In the literature, most authors consider much simpler fading channels to obtain closed-form expressions for various performance metrics, such as the outage probability.

We denote by $\gamma_{u,v}$ the instantaneous received signal-to-noise ratio for the link (u, v) , whereas $\gamma_{t,v}$ stands for the instantaneous interference-to-noise ratio from interferer t . These two ratios can be written as

$$\gamma_{u,v} = P_u \alpha_{u,v}^2 / N_0 \quad (3)$$

$$\gamma_{t,v} = P_t \alpha_{t,v}^2 / N_0. \quad (4)$$

The desired signal and the interfering signal experience both a composite log-normal shadowing and Nakagami fast fading. The expression of the PDF of $\gamma_{u,v}$ and $\gamma_{t,v}$ can be expressed as [10]

$$p_{\gamma_{u,v}}(\gamma) = \int_0^\infty \frac{m^m \gamma^{m-1}}{\Omega^m \Gamma(m)} \exp\left(-\frac{m\gamma}{\Omega}\right) \times \left\{ \frac{10/\ln(10)}{\sqrt{2\pi\sigma^2\Omega}} \exp\left[-\frac{(10\log_{10}\Omega - \mu)^2}{2\sigma^2}\right] \right\} d\Omega, \quad (5)$$

where m stands for the Nakagami shape factor, Ω refers to the received power, μ is the average received power, and σ represents the shadow standard deviation which has a typical value of 8 dB. The average received signal-to-noise ratio at node v can be expressed as

$$\mu_{u,v} = \frac{P_u}{N_0} \left(\frac{d_0}{d_{u,v}} \right)^\eta, \quad (6)$$

where d_0 represents a reference distance typically set to 1 m. Using the change of variable $x = (10\log_{10}(\Omega) - \mu)/\sqrt{2}\sigma$, we obtain

$$p_{\gamma_{u,v}}(\gamma) = \frac{m^m \gamma^{m-1}}{\Gamma(m)\sqrt{\pi}} \int_{-\infty}^\infty \exp\left(-\frac{m\gamma}{\Phi_{u,v}(x)}\right) \frac{\exp(-x^2)}{(\Phi_{u,v}(x))^m} dx, \quad (7)$$

where $\Phi_{u,v}(x) = 10^{(\sqrt{2}\sigma x - \mu_{u,v})/10}$. Utilizing the Hermite-Gauss quadrature integration, we can compute efficiently the

integral in (7) as

$$p_{\gamma_{u,v}}(\gamma) = \frac{m^m \gamma^{m-1}}{\Gamma(m)\sqrt{\pi}} \sum_{i=1}^N w_i \exp\left(-\frac{m\gamma}{\Phi_{u,v}(x_i)}\right) \times \frac{1}{(\Phi_{u,v}(x_i))^m}, \quad (8)$$

where $x_i (i = 1, \dots, N)$ are the zeros of the N th-order Hermite polynomial and w_i are weight factors tabulated in [11, Table 25.10]. The PDF $p_{\gamma_{u,v}}(\gamma)$ of $\gamma_{u,v}$ can be obtained from (8) by substituting m, μ, σ with $m_{u,v}, \mu_{u,v}, \sigma_{u,v}$. Similarly, the PDF $p_{\gamma_{t,v}}(\gamma)$ of $\gamma_{t,v}$ can be obtained from (8) by substituting m, μ, σ with $m_{t,v}, \mu_{t,v}, \sigma_{t,v}$.

In our physical model, the effects of path loss, shadowing, and fast fading are taken into account. The instantaneous received SINR at node v can be determined as [10]

$$\text{SINR}_{u,v} = \frac{P_u \alpha_{u,v}^2}{N_0 + \sum_{t \in \mathcal{N}}^{t \neq u,v} P_t \alpha_{t,v}^2} = \frac{\gamma_{u,v}}{1 + \sum_{t \in \mathcal{N}}^{t \neq u,v} \gamma_{t,v}}. \quad (9)$$

Theorem 1: $\forall u \in \mathcal{S}$ and $\forall v \in \mathcal{Y} \cup \mathcal{A}$, v fails to receive correctly a packet transmitted from u if $\text{SINR}_{u,v}$ falls below a threshold γ_{th} , i.e. $\text{SINR}_{u,v} < \gamma_{\text{th}}$. This event is known as an outage event and occurs with a probability, $\mathcal{P}_{u,v}$, which can be expressed as

$$\mathcal{P}_{u,v} = \sum_{n=1}^N \frac{w_n}{\sqrt{\pi}} - \sum_{n=1}^N \frac{w_n}{\sqrt{\pi}} \sum_{i_1=1}^N \sum_{i_2=1}^N \cdots \sum_{i_L=1}^N \left(\prod_{l=1}^L \frac{w_{i_l}}{\sqrt{\pi}} \right) \sum_{l=1}^L \sum_{j=1}^m \frac{\beta_{l,j} (-1)^j}{(j-1)!} \sum_{p=1}^P \lambda_p (\Phi_l)^j (\theta_p)^{j-1} \frac{\Gamma\left(m, \frac{m\gamma_{\text{th}}(\Phi_l \theta_p + 1)}{\Phi_{u,v}(x_n)}\right)}{\Gamma(m)}. \quad (10)$$

Proof: See Appendix. \square

To the best of our knowledge, the expression of the outage probability in **Theorem 1** constitutes a novel result that was not derived in the literature before. From (10), we can conclude that the outage probability $\mathcal{P}_{u,v}$ increases proportionally with the receiver's sensitivity threshold γ_{th} . In fact, the second term in the right hand side of (10) decreases with γ_{th} . The receiver's sensitivity threshold γ_{th} indicates the threshold that the SINR should exceed so that the packet is successfully received. As the sensitivity threshold of the receiver increases, its ability to detect weak signals decreases. The larger the receiver's sensitivity threshold is, the lower the probability of successful packet reception becomes.

C. Delay Model

This section is devoted to the delay analysis of MTC systems. The MTC devices in the network are equipped with a buffer to store the packets before their transmission. The use of buffers improves the control of packet flow and reduces the network congestion. As mentioned in Section III-A, the MTC device u can attach to v (i.e., an anchor node or Hub node). The packets are generated at the MTC devices according to a Poisson distribution with a rate λ_u . For the packets received at v , the arrival of these packets follows a Poisson distribution with a rate λ_u . This section focus on the analyses

of the average sojourn time in the buffer denoted as $\mathcal{D}_{u,v}^{[\lambda_u]}$ and the average waiting time for a data packet referred to as $\mathcal{W}_{u,v}^{[\lambda_u]}$.

The average waiting time $\mathcal{W}_{u,v}^{[\lambda_u]}$ in the buffer of u for a data packet is the average time spent by a packet in the buffer of u which begins from the generation of that packet at the buffer of u until the start of its transmission. Note that the successful reception of a packet at the destination v occurs after a random number of retransmissions. To quantify the delay associated with the retransmission events, we measure the average sojourn time $\mathcal{D}_{u,v}^{[\lambda_u]}$ of a packet in the buffer of u , which is defined as the average time elapsed from the generation of the packet until its successful reception at the next destination node v . The packet's sojourn time in the buffer can be evaluated using the Pollaczek-Khinchin equation as [12]

$$\mathcal{D}_{u,v}^{[\lambda_u]} = \mathcal{W}_{u,v}^{[\lambda_u]} + \mathbb{E}(T_{u,v})T_F, \quad (11)$$

where T_F is the time required for a single transmission of a given packet and $\mathbb{E}(T_{u,v})$ is the average number of retransmissions for the packets sent from u to v . For the ARQ scheme, the packet is retransmitted until successful reception at the receiver v or a maximum number of retransmissions M is reached. In case of reception failure after M retransmissions the packet is discarded. The number of retransmissions $T_{u,v}$ varies randomly according to the conditions of the fading channel between the nodes u and v . The average number of retransmissions $\mathbb{E}(T_{u,v})$ can be expressed as [13]

$$\begin{aligned} \mathbb{E}(T_{u,v}) &= 1 + \sum_{m=1}^{M-1} P(F^1, \dots, F^m) = 1 + \sum_{m=1}^{M-1} (\mathcal{P}_{u,v})^m \\ &= \sum_{m=0}^{M-1} (\mathcal{P}_{u,v})^m = \frac{1 - (\mathcal{P}_{u,v})^M}{1 - \mathcal{P}_{u,v}}, \end{aligned} \quad (12)$$

where $P(F^1, \dots, F^m)$ is the probability of a reception failure at the $1, \dots, m^{\text{th}}$ retransmissions. In our channel model, we consider a block-fading channel where the channel remains constant over one retransmission but changes independently from one ARQ round to another. At each ARQ round the channel realization is an output of a random generator. Note that these channel realizations are independent and identically distributed. This implies that the SINRs at different retransmissions are independent and have the same distributions. Consequently, the event of reception failure at each step are independent and have equal probabilities, thus $P(F^1, \dots, F^m) = (\mathcal{P}_{x,y})^m$. Using (12), we can conclude that the average number of retransmissions $\mathbb{E}(T_{u,v})$ increases as the outage probability $\mathcal{P}_{u,v}$ and the value of M increase.

The average waiting time $\mathcal{W}_{u,v}^{[\lambda_u]}$ for a data packet can be obtained as [12]

$$\mathcal{W}_{u,v}^{[\lambda_u]} = \frac{\lambda_u \mathbb{E}(T_{u,v}) T_F^2}{2(1 - \rho)} + \frac{T_F}{2}, \quad (13)$$

where ρ is a parameter which should satisfy the following stability condition

$$\rho = \lambda_u \mathbb{E}(T_{x,y}) T_F < 1. \quad (14)$$

The term $\mathbb{E}(T_{u,v}^2)$ represents the second-order moment of the number of retransmissions $T_{u,v}$ and can be expressed as [13]

$$\mathbb{E}(T_{u,v}^2) = 1 + \sum_{m=1}^{M-1} (2m+1) P(F^1, \dots, F^m). \quad (15)$$

Using [14, eq. (0.113)] and the equality $P(F^1, \dots, F^m) = (\mathcal{P}_{u,v})^m$, we can further simplify the expression of $\mathbb{E}(T_{u,v}^2)$ as

$$\begin{aligned} \mathbb{E}(T_{u,v}^2) &= 1 + \sum_{m=1}^{M-1} (2m+1) (\mathcal{P}_{u,v})^m = \sum_{m=0}^{M-1} (2m+1) (\mathcal{P}_{u,v})^m \\ &= \frac{1 - (2M-1) (\mathcal{P}_{u,v})^M}{1 - \mathcal{P}_{u,v}} + \frac{2\mathcal{P}_{u,v} (1 - (\mathcal{P}_{u,v})^{M-1})}{(1 - \mathcal{P}_{u,v})^2}. \end{aligned} \quad (16)$$

From (16), it can be concluded that the second-moment $\mathbb{E}(T_{u,v}^2)$ of the number of retransmissions increases as the outage probability and M increase. Using (13), it can be clearly seen that the average waiting time $\mathcal{W}_{u,v}^{[\lambda_u]}$ is proportional to the second-moment $\mathbb{E}(T_{u,v}^2)$. Similarly, it can be deduced from (11) that the sojourn time $\mathcal{D}_{u,v}^{[\lambda_u]}$ is proportional to both the first-moment, $\mathbb{E}(T_{u,v})$, and the second-moment, $\mathbb{E}(T_{u,v}^2)$, of the number of retransmissions. Since both $\mathbb{E}(T_{u,v})$ and $\mathbb{E}(T_{u,v}^2)$ increase proportionally with the outage probability and M , consequently a larger average sojourn time $\mathcal{D}_{u,v}^{[\lambda_u]}$ and a longer average waiting time $\mathcal{W}_{u,v}^{[\lambda_u]}$ are experienced by the packets as the outage probability $\mathcal{P}_{u,v}$ and the value of M increases.

IV. EFFICIENT HUBS PLACEMENT ALGORITHMS FOR INTERCONNECTING MTC DEVICES

In this section, we present three Hubs placement solutions for interconnecting different MTC devices to the anchor nodes. For getting optimal configurations, the three solutions have formalized through linear integer programming. While these solutions mainly target the problem of constrained Hubs deployment, the linear integer programming of each of them differ from each other. In the first solution, dubbed Cost efficient Hubs deployment for connecting MTC devices (C-MTC), the objective is to minimize the cost by deploying lower number of Hubs and using cheaper communication technologies, while the second solution, dubbed Delay aware Hubs deployment for connecting MTC devices (D-MTC), aims at reducing the end-to-end delay. While the two first solutions take into consideration solely one objective independently of the other, the third solution, dubbed Fair Hubs deployment for efficient MTC applications (F-MTC), aims at finding, using bargaining game, a fair trade-off between the cost and end-to-end delay when placing Hubs and interconnecting MTC devices to the anchor nodes. The following paragraph defines the different variables used in the three solutions.

Each wireless communication has limited communication range, which makes the number of successors/predecessors of Hub or MTC devices very limited. In order to lightweight the complexity of the proposed solutions, we aim at reducing the search space by limiting the number of variables.

As mentioned in table I, we denote by $\eta^\ell(i)$ (i 's neighbors) the set of $\mathcal{Y} \cup \mathcal{A}$ (anchors or Hub's positions) that can be selected as successors/predecessors of the node i when the technology ℓ is used. Formally, $\eta^\ell(i)$ is defined as the set of nodes that fall within a specified distance to i . In addition to \mathcal{L} , $\Gamma_{i,j}^\ell$, $\pi_{i,j}^\ell$ and $\rho_{i,j}^\ell$ aforementioned in section III-A, we defined the following variables: 1) Ψ_i that is a real variable that represents the amount of data traffic that should be aggregated and forwarded by a Hub node $i \in \mathcal{Y}$; 2) Π_i is a real variable that represents the link capacity required by a Hub node $i \in \mathcal{Y}$ to handle the amount of received traffic without creating network congestion. A congestion may occur in the network if $\Pi_i < \Psi_i$; 3) For each MTC device $i \in \mathcal{S}$ and a Hub or anchor node $j \in \mathcal{Y} \cup \mathcal{A}$, we define $\mathcal{Y}_{i,j}$, which is a Boolean decision variable. $\mathcal{Y}_{i,j}$ equals to 1 if the node i is attached to the anchor j . Otherwise, it equals to 0. 4) For each communication technology ℓ , and Hub position $i \in \mathcal{Y}$ and its neighbor $j \in \eta^\ell(i)$, we define $\mathcal{X}_{i,j}^\ell$, which is a Boolean decision variable. $\mathcal{X}_{i,j}^\ell$ equals to 1 if the node i selects the node j as successor using technology ℓ . Otherwise, it equals to 0. 5) \mathcal{T}_i is a real variable that represents the end-to-end delay of a node $i \in \mathcal{Y} \cup \mathcal{A}$; 6) For each $(i, j) \in (\mathcal{Y} \cup \mathcal{A})^2$ and $\ell \in \mathcal{L}$, we define $\mathcal{F}_{i,j}^\ell$, which is an integer variable that introduces the notion of flow in order to force the optimization function to ensure the end-to-end connectivity between \mathcal{X} and \mathcal{A} without creating any cycle; 7) $\forall i \in \mathcal{Y}$: \mathcal{Z}_i is a Boolean decision variable equals to 1 if a Hub node is deployed in the location $i \in \mathcal{Y}$. Otherwise, it equals to 0. If a location i is empty, then $\Psi_i = 0$, $\Pi_i = 0$ and $\mathcal{T}_i = 0$. Moreover, for each $j \in \eta^\ell(i)$, $\ell \in \mathcal{L}$, we have $\mathcal{X}_{i,j}^\ell = 0$ iff $\mathcal{Z}_i = 0$. Formally, $\mathcal{Z}_i = 0 \Leftrightarrow (\Psi_i = 0 \wedge \Pi_i = 0 \wedge \mathcal{T}_i = 0 \wedge \mathcal{X}_{i,j}^\ell = 0)$.

A. C-MTC: Cost Efficient Hubs Deployment for Connecting MTC Devices

In this subsection, we describe C-MTC that aims at minimizing the Hubs' deployment cost. Basically, the cost would be minimized by reducing both the number of Hubs and the cost of the communication technologies used for interconnecting the Hubs to \mathcal{A} . The C-MTC is formulated as follows:

$$\min \left(\sum_{i \in \mathcal{Y}} \tau \times \mathcal{Z}_i + \sum_{i \in \mathcal{Y}} \sum_{\ell \in \mathcal{L}} \sum_{j \in \eta^\ell(i)} \Gamma_{i,j}^\ell \times \mathcal{X}_{i,j}^\ell \right) \quad (17)$$

$$\text{s.t. } \forall i \in \mathcal{S} : \sum_{j \in \mathcal{Y} \cup \mathcal{A}} \mathcal{Y}_{i,j} = 1 \quad (18)$$

$$\forall i \in \mathcal{Y} : \sum_{\ell \in \mathcal{L}} \sum_{j \in \eta^\ell(i)} \mathcal{X}_{i,j}^\ell \leq \mathcal{Z}_i \quad (19)$$

$$\forall i \in \mathcal{Y} : \Psi_i = \sum_{j \in \mathcal{S}} \mathcal{Y}_{j,i} \times \lambda_j + \sum_{\ell \in \mathcal{L}} \sum_{j \in \eta^\ell(i) \cap \mathcal{Y}} \mathcal{X}_{j,i}^\ell \times \Psi_j \quad (20)$$

$$\forall i \in \mathcal{Y} : \Pi_i = \sum_{\ell \in \mathcal{L}} \sum_{j \in \eta^\ell(i)} \mathcal{X}_{i,j}^\ell \times \pi_{i,j}^\ell \quad (21)$$

$$\forall i \in \mathcal{Y} : \Psi_i \leq \Pi_i. \quad (22)$$

$$\sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{A}} \mathcal{Y}_{i,j} + \sum_{i \in \mathcal{Y}} \sum_{\ell \in \mathcal{L}} \sum_{j \in \mathcal{A}} \mathcal{F}_{i,j}^\ell = |\mathcal{S}|. \quad (23)$$

$$\forall i \in \mathcal{Y} : \sum_{\ell \in \mathcal{L}} \sum_{j \in \eta^\ell(i)} \mathcal{F}_{i,j}^\ell = \sum_{j \in \mathcal{S}} \mathcal{Y}_{j,i} + \sum_{\ell \in \mathcal{L}} \sum_{j \in \eta^\ell(i)} \mathcal{F}_{j,i}^\ell. \quad (24)$$

$$\forall i \in \mathcal{Y}, \quad \forall \ell \in \mathcal{L}, \quad \forall j \in \eta^\ell(i) : 0 \leq F_{i,j}^\ell \leq |\mathcal{S}| \times \mathcal{X}_{i,j}^\ell. \quad (25)$$

$$\forall i \in \mathcal{S}, \quad \forall j \in \mathcal{A} : \mathcal{Y}_{i,j} \times \mathcal{D}_{i,j}^{[\lambda_i]} \leq \mathcal{T}_i. \quad (26)$$

$$\forall i \in \mathcal{Y}, \quad \forall j \in \mathcal{S} : \mathcal{Y}_{j,i} \times \mathcal{D}_{j,i}^{[\lambda_j]} + \sum_{k \in \eta^\ell(i)} \sum_{\ell \in \mathcal{L}} \rho_{i,k}^\ell \mathcal{X}_{i,k} \leq \mathcal{T}_i. \quad (27)$$

$$\forall i \in \mathcal{Y}, \quad \forall j \in \eta^\ell(i) \cap \mathcal{Y}, \quad \forall \ell \in \mathcal{L} : \mathcal{T}_j \times \mathcal{X}_{j,i}^\ell + \sum_{\ell' \in \mathcal{L}} \sum_{k \in \eta^{\ell'}(i)} \rho_{i,k}^{\ell'} \times \mathcal{X}_{i,k} \leq \mathcal{T}_i. \quad (28)$$

$$\forall i \in \mathcal{S} \cup \mathcal{Y} : \mathcal{T}_i \leq \zeta_{\mathcal{D}}. \quad (29)$$

The objective function (17) allows minimizing the cost for interconnecting MTC devices to \mathcal{A} . The cost, herein, is divided in two parts: $i)$ $\sum_{i \in \mathcal{Y}} \tau \times \mathcal{Z}_i$, which represents the cost of the Hubs' deployment; $ii)$ $\sum_{i \in \mathcal{Y}} \sum_{\ell \in \mathcal{L}} \sum_{j \in \eta^\ell(i)} \Gamma_{i,j}^\ell \times \mathcal{X}_{i,j}^\ell$, which represents the cost of the communication technologies used for interconnecting different Hubs in the network. Meanwhile, the constraints are used to ensure the following conditions: Constraint (18) ensures that each MTC device should be assigned only to one access point, which can be a Hub or an anchor node. Constraint (19) ensures that each candidate position $i \in \mathcal{Y}$ which participate in forwarding the traffic is considered as a Hub node. Moreover, it ensures also that i should select only one successor using different technologies. Constraints (20), (21) and (22) allow preventing congestion in the network. In constraint (20), we compute the amount of traffic Ψ_i handled by a Hub i and should be forwarded to its successor. It consists of two main parts: $i)$ $\sum_{j \in \mathcal{S}} \mathcal{Y}_{j,i} \times \lambda_j$, which represents the amount of traffic generated by the MTC devices attached to i , and $ii)$ $\sum_{\ell \in \mathcal{L}} \sum_{j \in \eta^\ell(i) \cap \mathcal{Y}} \mathcal{X}_{j,i}^\ell \times \Psi_j$, which denotes the sum of traffic forwarded by i 's predecessors ($pre(i)$). The amount of link capacity between the Hub i and its successor is represented by constraint (21). Whereas, constraint (22) specifies that the handled traffic by i should be lower than the link capacity between it and its successor. These constraints will force the optimization function to select the adequate technology that carries all the traffic without creating any congestion in the network.

As to the constraints (23), (24) and (25), they ensure the network connectivity of each MTC device to an anchor node. To this end, each MTC device mimics packet flow by generating only one packet that will be routed from the MTC devices to the anchor nodes, which in turn forwards the same number of received packets from its predecessors. The mimicked flow should be forwarded within the constructed topology (i.e. each node sends the flow to its successor only). While constraint (23) ensures that the number of packets received by \mathcal{Y} equals to the number of MTC devices \mathcal{S} in the network, constraint (24) guarantees that each Hub forwards the packet flow of its MTC devices as well as to the flow received

from its predecessors. To force the generated flow to be routed only within the constructed topology, constraint (25) is employed. The constraints (26), (27) and (28) allow modeling the end-to-end delay in the network. We denote by \mathcal{T}_i the maximum end-to-end delay of a node $i \in \mathcal{S} \cup \mathcal{Y}$, and compute \mathcal{T}_i for each $i \in \mathcal{S}$ using constraint (26), where \mathcal{T}_i should be greater than or equals to $\mathcal{D}_{i,j}^{[\lambda_i]}$ (the average sojourn time in the buffer of the MTC device i) if node i selects node j as successor. By introducing constraint (27), we ensure that the maximum end-to-end delay \mathcal{T}_i of a Hub $i \in \mathcal{Y}$ is greater than the end-to-end delay of any MTC devices, that select i as a successor, plus the time required for i to forward the received data to its successor (i.e., $\sum_{k \in \eta^{\ell'}(i)} \sum_{\ell' \in \mathcal{L}} \rho_{i,k}^{\ell'} \times \mathcal{X}_{i,k}$).

Similarly, constraint (28) ensures that the maximum end-to-end delay of a Hub $i \in \mathcal{Y}$ is greater than its predecessors plus the time required for i to forward the received data to its successor. Finally, constraint (29) ensures that the maximum end-to-end delay in the network does not exceed a predefined threshold $\zeta_{\mathcal{D}}$.

However the proposed optimization problem ((17)–(29)) is not linear due to the constraints defined by (20) and (28). In order to simplify the optimization problem, we convert it to a linear integer programming by doing the following transformations on (20) and (28). Constraint (20) can be transformed to linear constraint as follows: First, we add the following variables to the optimization problem: $\Psi_{j,i}^{\ell}$ is a real number $\forall i \in \mathcal{Y}, \forall \ell \in \mathcal{L}, j \in \eta^{\ell}(i) \cap \mathcal{Y}$. Each variable $\Psi_{j,i}^{\ell}$ would replace the expression $\mathcal{X}_{j,i}^{\ell} \times \Psi_j$, which is not linear. Second, we update (20) by the following constraints:

$$\forall i \in \mathcal{Y} : \Psi_i = \sum_{j \in \mathcal{S}} \mathcal{Y}_{j,i} \times \lambda_j + \sum_{\ell \in \mathcal{L}} \sum_{j \in \eta^{\ell}(i) \cap \mathcal{Y}} \Psi_{j,i}^{\ell} \quad (30)$$

$$\forall i \in \mathcal{Y}, \forall \ell \in \mathcal{L}, \forall j \in \eta^{\ell}(i) \cap \mathcal{Y} : \Psi_{j,i}^{\ell} - \Psi_j \leq (1 - \mathcal{X}_{j,i}^{\ell}) \times \mathcal{M} \quad (31)$$

$$\forall i \in \mathcal{Y}, \forall \ell \in \mathcal{L}, \forall j \in \eta^{\ell}(i) \cap \mathcal{Y} : \Psi_j - \Psi_{j,i}^{\ell} \leq (1 - \mathcal{X}_{j,i}^{\ell}) \times \mathcal{M} \quad (32)$$

$$\forall i \in \mathcal{Y}, \forall \ell \in \mathcal{L}, \forall j \in \eta^{\ell}(i) \cap \mathcal{Y} : \Psi_{j,i}^{\ell} \leq \mathcal{X}_{j,i}^{\ell} \times \mathcal{M}, \quad (33)$$

where \mathcal{M} is a big real number ($\mathcal{M} \simeq \infty$).

Constraint (30) is equivalent to the constraint (20); $\mathcal{X}_{j,i}^{\ell} \times \Psi_j$ is substituted by $\Psi_{j,i}^{\ell}$. Meanwhile, constraints (31), (32) and (33) ensure that $\Psi_j = \Psi_{j,i}^{\ell}$ if $\mathcal{X}_{j,i}^{\ell} = 1$. Otherwise, $\Psi_{j,i}^{\ell} = 0$. If $\mathcal{X}_{j,i}^{\ell} = 1$, then two statements will hold: First, from (31) and (32), we have $\Psi_j = \Psi_{j,i}^{\ell}$; Second, from (33), we have $\Psi_{j,i}^{\ell} \leq \infty$, which is always true. Similarly, if $\mathcal{X}_{j,i}^{\ell} = 0$, then we have from (31) and (32) Ψ_j and $\Psi_{j,i}^{\ell}$ can take any values, while (33) sets the value of $\Psi_{j,i}^{\ell}$ to zero.

Likewise, constraint (28) will be transformed to linear constraint as follows: we first add the following variables to the optimization problem: $T_{j,i}^{\ell}$ is a real number $\forall i \in \mathcal{Y}, \forall \ell \in \mathcal{L}, j \in \eta^{\ell}(i) \cap \mathcal{Y}$. Each variable $T_{j,i}^{\ell}$ would replace the expression $\mathcal{T}_j \times \mathcal{X}_{j,i}^{\ell}$. Then, we update (28) by the following constraints:

$$\forall i \in \mathcal{Y}, \forall \ell \in \mathcal{L}, \forall j \in \eta^{\ell}(i) \cap \mathcal{Y} : T_{j,i}^{\ell} + \sum_{\ell' \in \mathcal{L}} \sum_{k \in \eta^{\ell'}(i)} \rho_{i,k}^{\ell'} \times \mathcal{X}_{i,k} \leq \mathcal{T}_i. \quad (34)$$

$$\forall i \in \mathcal{Y}, \forall \ell \in \mathcal{L}, \forall j \in \eta^{\ell}(i) \cap \mathcal{Y} : T_{j,i}^{\ell} - \mathcal{T}_j \leq (1 - \mathcal{X}_{j,i}^{\ell}) \times \mathcal{M} \quad (35)$$

$$\forall i \in \mathcal{Y}, \forall \ell \in \mathcal{L}, \forall j \in \eta^{\ell}(i) \cap \mathcal{Y} : \mathcal{T}_j - T_{j,i}^{\ell} \leq (1 - \mathcal{X}_{j,i}^{\ell}) \times \mathcal{M} \quad (36)$$

$$\forall i \in \mathcal{Y}, \forall \ell \in \mathcal{L}, \forall j \in \eta^{\ell}(i) \cap \mathcal{Y} : T_{j,i}^{\ell} \leq \mathcal{X}_{j,i}^{\ell} \times \mathcal{M} \quad (37)$$

Constraint (34) is equivalent to the constraint (28); $\mathcal{T}_j \times \mathcal{X}_{j,i}^{\ell}$ is replaced by $T_{j,i}^{\ell}$. The constraints (35), (36) and (37) ensure that $\mathcal{T}_j = T_{j,i}^{\ell}$ if j is a predecessor of i (i.e., $\mathcal{X}_{j,i}^{\ell} = 1$). Otherwise, $T_{j,i}^{\ell} = 0$. If $\mathcal{X}_{j,i}^{\ell} = 1$, then two statements would hold: First, from (35) and (36), we have $\mathcal{T}_j = T_{j,i}^{\ell}$; Second, from (37), we have $T_{i,j} \leq \infty$, which is always true. At the other hand, if $\mathcal{X}_{j,i}^{\ell} = 0$, then from (35) and (36), \mathcal{T}_j and $T_{j,i}^{\ell}$ can take any values, while (37) sets the value of $T_{j,i}^{\ell}$ to zero. Based on the aforementioned analysis, the optimization problem would be transformed to the following linear program

$$\begin{aligned} \text{(OP1)} : \quad & \min \left(\sum_{i \in \mathcal{Y}} \tau \times \mathcal{Z}_i + \sum_{i \in \mathcal{Y}} \sum_{\ell \in \mathcal{L}} \sum_{j \in \eta^{\ell}(i)} \Gamma_{i,j}^{\ell} \times \mathcal{X}_{i,j}^{\ell} \right) \\ \text{S.t.} \quad & (18), (19), (30), (31), (32), (33), (21), \\ & (22), (23), (24), (25), (26), (27), (29), \\ & (34), (35), (36), (37). \end{aligned} \quad (38)$$

B. D-MTC: Delay Aware Hubs Deployment for Connecting MTC Devices

In this subsection, we provide a quick description of D-MTC solution. Unlike C-MTC solution, D-MTC aims at increasing the QoE and reduce the end-to-end delay in the network when deploying the Hubs. Basically, the end-to-end delay will be minimized, and by consequent the QoE increases, by selecting a good link quality between MTC devices and their successors, as well as the use of high quality links between different Hubs. D-MTC seeks the optimal deployment of Hubs by applying linear integer programming. More precisely, min-max approach is employed to minimize the maximum end-to-end delay in the network. D-MTC is formulated as follows

$$\text{(OP2)} : \quad \min \max_{\forall i \in \mathcal{Y} \cup \mathcal{A}} \mathcal{T}_i \quad (39)$$

$$\text{S.t.} \quad (18), (19), (30), (31), (32), (33), (21),$$

$$(22), (23), (24), (25), (26), (27),$$

$$(34), (35), (36), (37), \text{ and}$$

$$\sum_{i \in \mathcal{Y}} \tau \times \mathcal{Z}_i + \sum_{i \in \mathcal{Y}} \sum_{\ell \in \mathcal{L}} \sum_{j \in \eta^{\ell}(i)} \Gamma_{i,j}^{\ell} \times \mathcal{X}_{i,j}^{\ell} \leq \zeta_{\mathcal{C}}. \quad (40)$$

The main goal of the objective function (39) is to minimize the end-to-end delay by minimizing the maximum data transfer of each node. While the first constraints ((18)-(36)) are similar to the ones presented in OP1, the constraint (40) ensures that the cost for interconnecting \mathcal{S} to \mathcal{A} do not exceed a predefined threshold $\zeta_{\mathcal{C}}$.

C. F-MTC: Fair Hubs Deployment for Efficient MTC Applications

The F-MTC solution aims to find a fair trade-off between the conflicting objectives, i.e., the cost and the end-to-end

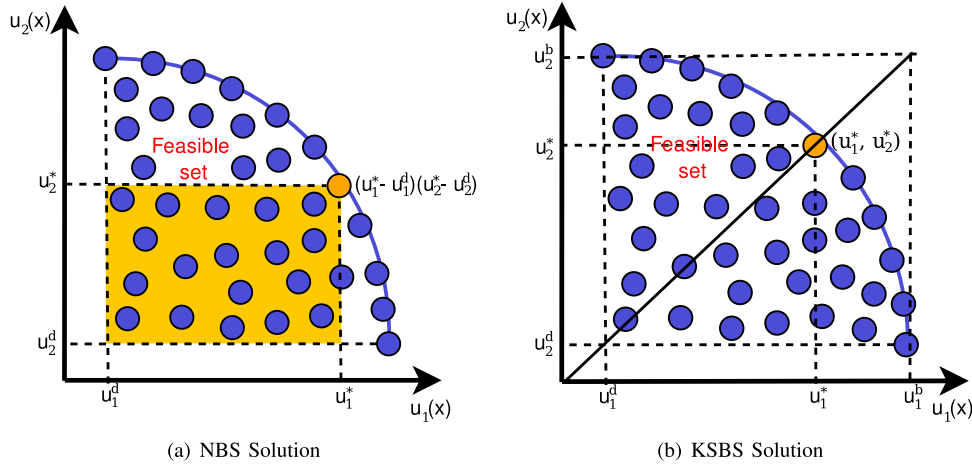


Fig. 2. NBS and KSBS solutions.

delay. A bargaining game would be used to find the fair trade-off. From one point of view, the deployment of high number of Hubs taking into account the shortest path between every MTC device and the anchor node and the use of good communication technologies will reduce the end-to-end delay. However, from another point of view, it will increase the cost for the users significantly. Whereas, the deployment of low number of Hubs and the use of the cheapest communication technologies will negatively impact the QoE and the end-to-end delay. Therefore, in order to find a fair trade off between the cost and end-to-end delay, we have introduced the F-MTC approach using the concept of bargaining game theory (a form of the cooperative games) where both the cost and end-to-end delay are considered as two players that would like to barter goods. In what follows, we would like to introduce first, the concept of the cooperative games, whereby we will present two games, namely: the Nash Bargaining Model (NBS) and the Kalai and Smorodinsky Bargaining (KSBS) Model which is an enhanced form of the NBS model.

1) *Cooperative Games*: In cooperative games, the players are assumed to attain either most desirable point when negotiation succeeds or disagreement point when negotiation fails. We consider a two persons game who would like to barter goods, each one of them wants to increase his benefits. Let \mathcal{P} the vector payoffs of these players. Formally, $\mathcal{P} = \{(u_1(x), u_2(x)), x = (x_1, x_2) \in X\}$, where X is the set of the two players' strategies. $u_1(x)$ and $u_2(x)$ represent the players utility functions respectively. In [15], Nash bargaining model (NBS) is presented, which is a cooperative game with non-transferable utility. This means that the utility scales of the players are measured in non-comparable units. Nash bargaining game is based on two elements assumed to be given and known to the players. First, the set of vector payoffs \mathcal{P} achieved by the players if they agree to cooperate. \mathcal{P} should be a convex and compact set. Second, the threat point, $d = (u_1^d, u_2^d) \in \mathcal{P}$, which represents the pair of utility whereby the two players fail to achieve an agreement. In NBS, we aim to find a fair and reasonable point, $(u_1^*, u_2^*) = f(\mathcal{P}, u_1^d, u_2^d) \in \mathcal{P}$. Based on Nash theory, a set of axioms are

defined that lead to $f(\mathcal{P}, u_1^d, u_2^d)$ achieves a unique optimal solution (u_1^*, u_2^*) [15]. Moreover, the unique solution (\bar{u}, \bar{v}) , satisfying the above axioms, is proven to be the solution of the following optimization problem:

$$\max(u_1(x) - u_1^d)(u_2(x) - u_2^d) \quad (41)$$

$$s.t. (u_1(x), u_2(x)) \in \mathcal{P} \quad (42)$$

$$(u_1(x), u_2(x)) \geq (u_1^d, u_2^d) \quad (43)$$

An enhanced solution of Nash bargaining game, named KSBS, is proposed by Kalai and Smorodinsky [16]. KSBS aims to enhance the fairness between the players more by sharing the same utility fraction r among them. KSBS preserves the same Nash bargaining axioms except the independence of irrelevant alternatives. In addition to that, it has a new axiom called monotonicity. In contrast to the Nash bargaining game, KSBS, in addition to the disagreement point $d = (u_1^d, u_2^d) \in \mathcal{P}$, needs the ideal point for both players $x^b = (u_1^b, u_2^b)/x^b \in \mathcal{P}$, which is the best utility that both players can achieve separately without bargaining. Kalai and Smorodinsky proof that the unique solution that satisfies KSBS's axioms is the solution of the following optimization problem:

$$\max r \quad (44)$$

$$s.t. (u_1(x), u_2(x)) \in \mathcal{P} \quad (45)$$

$$r = \frac{u_1(x) - u_1^d}{u_1^b - u_1^d} \quad (46)$$

$$r = \frac{u_2(x) - u_2^d}{u_2^b - u_2^d} \quad (47)$$

Fig. 2 shows how KSBS game enhances NBS in terms of fairness and Pareto-optimality. As shown in Fig. 2(a), NBS aims to increase as much as possible the size of the orange rectangle. However, this strategy may favor one player than the other. As shown in Fig. 2(b), KSBS enhances the trade-off between the two players by sharing the same utility fraction $r = \frac{u_1^* - u_1^d}{u_1^b - u_1^d} = \frac{u_2^* - u_2^d}{u_2^b - u_2^d}$. In order to increase the benefit of each player, KSBS increases as much as possible the fraction r .

2) *F-MTC Description*: We denote by $d = (u_C^d, u_D^d)$ and $b = (u_C^b, u_D^b)$ the threat and best points of the KSBS game that solves F-MTC. As we have mentioned earlier, ζ_C and ζ_D denote the maximum threshold values of the cost and end-to-end delay. In KSBS game, both players (C-MTC and D-MTC) should bargain for increasing their benefits, which is the opposite to their utility function defined by the optimization problems (OP1) and (OP2). In order to use KSBS game to ensure a fair trade-off between the end-to-end delay and the cost, as depicted in Fig. 2(b), we need to change the utility function of both players to be a optimization problem. The utility function of the first player C-MTC is updated using the following optimization problem (OP3) as follows:

$$\begin{aligned} \text{OP3: } \max & (\zeta_C - \sum_{i \in \mathcal{Y}} \tau \times \mathcal{Z}_i - \sum_{i \in \mathcal{Y}} \sum_{\ell \in \mathcal{L}} \sum_{j \in \eta^\ell(i)} \Gamma_{i,j}^\ell \times \mathcal{X}_{i,j}^\ell) \\ \text{S.t, (18), (19), (30), (31), (32), (33), (21),} \\ & (22), (23), (24), (25), (26), (27), (29), \\ & (34), (35), (36), (37). \end{aligned} \quad (48)$$

Meanwhile, the utility function of the second player D-MTC is updated using the following optimization problem (OP4) as follows:

$$\begin{aligned} \text{OP4: } \max & \min_{i \in \mathcal{Y} \cup \mathcal{A}} (\zeta_D - \mathcal{T}_i) \\ \text{S.t, (18), (19), (30), (31), (32), (33), (21),} \\ & (22), (23), (24), (25), (26), (27), (34), (35), \\ & (36), (37), (40) \end{aligned} \quad (49)$$

In what follows, we will show how d and b would be computed. Let $\dot{\mathcal{X}}$ and $\ddot{\mathcal{X}}$ the two matrices of $\mathcal{X}_{i,j}^\ell$ variables obtained by resolving OP3 and OP4, respectively. Moreover, let $\dot{\mathcal{T}}$ and $\dot{\mathcal{Z}}$ two vectors obtained by resolving OP3, whereas $\ddot{\mathcal{T}}$ and $\ddot{\mathcal{Z}}$ two vectors obtained by resolving OP4. Then, $d = (u_C^d, u_D^d)$ and $b = (u_C^b, u_D^b)$ would be computed as follows:

- 1) $u_C^d = \zeta_C - \sum_{i \in \mathcal{Y}} \tau \times \dot{\mathcal{Z}}_i - \sum_{i \in \mathcal{Y}} \sum_{\ell \in \mathcal{L}} \sum_{j \in \eta^\ell(i)} \Gamma_{i,j}^\ell \times \dot{\mathcal{X}}_{i,j}^\ell$
- 2) $u_C^b = \zeta_C - \sum_{i \in \mathcal{Y}} \tau \times \ddot{\mathcal{Z}}_i - \sum_{i \in \mathcal{Y}} \sum_{\ell \in \mathcal{L}} \sum_{j \in \eta^\ell(i)} \Gamma_{i,j}^\ell \times \ddot{\mathcal{X}}_{i,j}^\ell$
- 3) $u_D^d = \min_{i \in \mathcal{Y} \cup \mathcal{A}} (\zeta_D - \dot{\mathcal{T}}_i)$
- 4) $u_D^b = \min_{i \in \mathcal{Y} \cup \mathcal{A}} (\zeta_D - \ddot{\mathcal{T}}_i)$

The fair Pareto optimal solution F-MTC will be formulated as follows:

$$\begin{aligned} \max & r \\ \text{S.t, (18), (19), (30), (31), (32), (33), (21), (22),} \\ & (23), (24), (25), (26), (27), (29), (34), (35), \\ & (36), (37), (40), \\ & r = \frac{u_C(x) - u_C^d}{u_C^b - u_C^d}, \\ & r = \frac{u_D(x) - u_D^d}{u_D^b - u_D^d}, \text{ and} \\ & u_C(x) = \zeta_C - \sum_{i \in \mathcal{Y}} \tau \times \mathcal{Z}_i - \sum_{i \in \mathcal{Y}} \sum_{\ell \in \mathcal{L}} \sum_{j \in \eta^\ell(i)} \Gamma_{i,j}^\ell \times \mathcal{X}_{i,j}^\ell \end{aligned} \quad (50)$$

$$\forall i \in \mathcal{Y} \cup \mathcal{A} : u_D(x) \geq \zeta_D - \mathcal{T}_i \quad (52)$$

Then, the maximum end-to-end delay $\mathcal{D}_{\mathcal{F}}$ and the cost $\mathcal{C}_{\mathcal{F}}$ of F-MTC will be computed as follows:

$$\mathcal{D}_{\mathcal{F}} = \zeta_D - u_D(x). \quad (53)$$

$$\mathcal{C}_{\mathcal{F}} = \zeta_C - u_C(x). \quad (54)$$

V. SIMULATION

In this section, we evaluate the proposed solutions C-MTC, D-MTC and F-MTC. The three solutions have been evaluated using python, extended package for graph theory called Networkx and Gurobi Optimizer software. We have compared the proposed solutions to two base-line approaches that use weighting method. The two base-line approaches have been formulated as follow:

$$\begin{aligned} \min & \alpha \times (\max_{i \in \mathcal{Y} \cup \mathcal{A}} \mathcal{T}_i) + (1 - \alpha) \times (\sum_{i \in \mathcal{Y}} \tau \times \mathcal{Z}_i \\ & + \sum_{i \in \mathcal{Y}} \sum_{\ell \in \mathcal{L}} \sum_{j \in \eta^\ell(i)} \Gamma_{i,j}^\ell \times \mathcal{X}_{i,j}^\ell) \end{aligned}$$

$$\begin{aligned} \text{S.t, (18), (19), (30), (31), (32), (33), (21), (22), (23),} \\ (24), (25), (26), (27), (29), (34), (35), (36), (37), (40). \end{aligned}$$

While we have fixed the value of α to 0.2 for the first solution, named W-0.2, we have fixed the value of α to 0.8 for the second solution named W-0.8. The different solutions are evaluated in terms of the following metrics: *i*) End-to-end delay, which is defined as the maximum time required for all MTC devices to forward all their packets to the anchor nodes; *ii*) The cost, which is defined as the prices of deployed Hubs and used technologies for interconnecting MTC devices to the Anchor nodes; *iii*) The execution time, which is defined as the time needed to execute each solution. For each solution we measure the difference between its finishing and starting time. Before starting the execution of each solution, we obtained the time of the system (t_1) in seconds, then when the execution of that solution is finished, we again obtained the time of the system (t_2). Then, we measure the execution time as the subtraction of the starting time from the finishing time ($t_2 - t_1$). In the simulation results, each plotted point represents the average of 100 executions. The plots are presented with 95% confidence interval. The evaluation of different algorithms is performed by varying: *i*) The number of MTC devices, and *ii*) The number of Hubs' positions. In the evaluation of different protocols, the receiver's sensitivity γ_{th} has been fixed to $10^{-9.4}$. We conducted three set of experiments. Firstly, we vary the number of MTC devices $|\mathcal{S}|$ while fixing the number of Hub's candidate positions $|\mathcal{Y}|$ to 50 and the receiver's sensitivity γ_{th} to $10^{-9.4}$. Secondly, we vary the number of Hub's candidate positions $|\mathcal{Y}|$ while fixing the number of MTC devices $|\mathcal{S}|$ to 80 and the receiver's sensitivity γ_{th} to $10^{-9.4}$. Finally, we vary the receiver's sensitivity γ_{th} while fixing the number of MTC devices $|\mathcal{S}|$ to 80 and the number of Hub's candidate positions $|\mathcal{Y}|$ to 50.

In the simulation, \mathcal{S} , \mathcal{Y} and \mathcal{A} are uniformly placed in a square area of length $D = 5 \text{ Km}^2$. The communication range of \mathcal{Y} is related to the technology $\ell \in \mathcal{L}$. Formally, a Hub's candidate position i is a neighbor to another Hub's candidate position j using technology ℓ iff the Euclidean distance

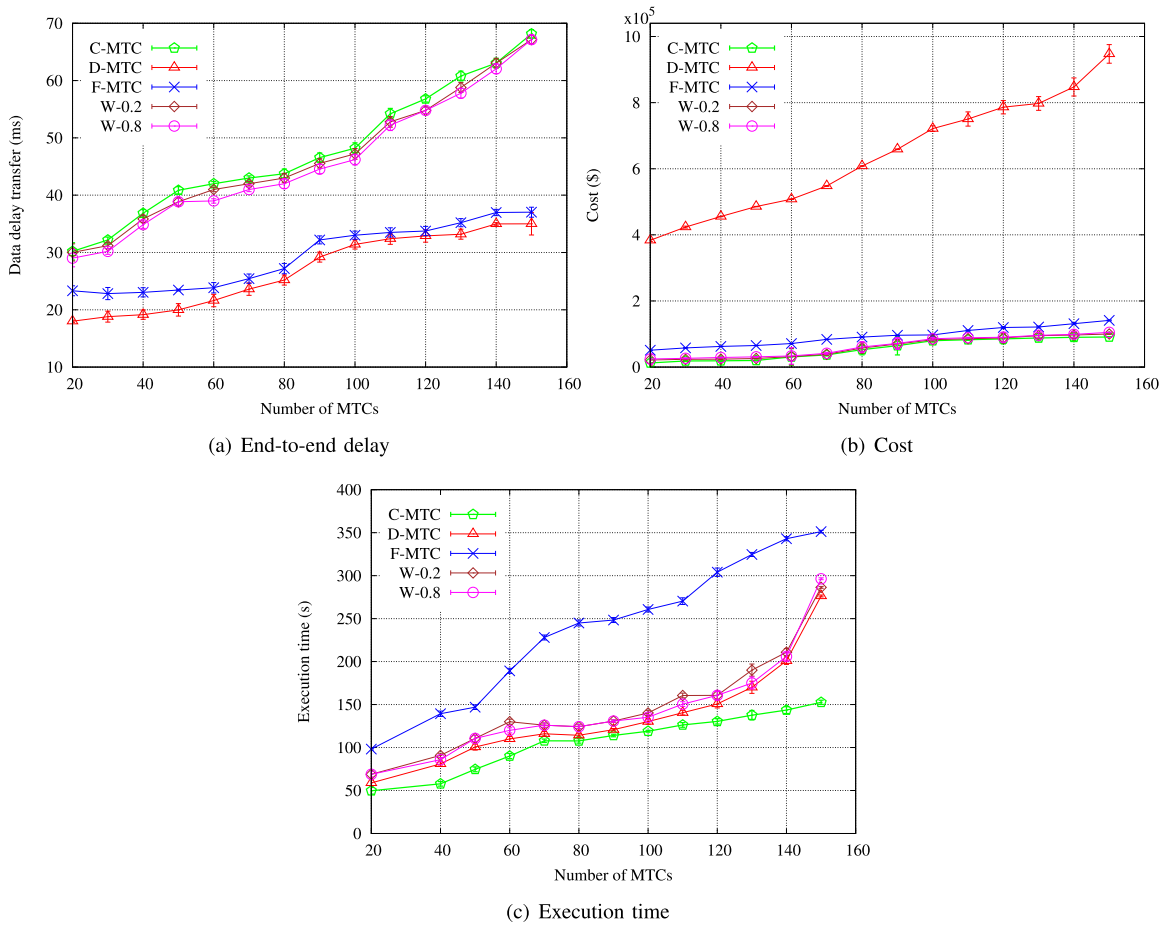


Fig. 3. Performance of the proposed solutions as a function of number of MTC devices.

between the two nodes doesn't exceed technology range. For interconnecting the Hubs, two sets of technologies have been considered: *i*) Wired based on fiber optic technology including Optical-Ethernet1, Optical-Ethernet2, Optical-Ethernet3 and GPON; *ii*) wireless based on Microwave (MW) and most recent IEEE 802.11 including MW Standard 6-42GHz, MW V-band, MW E-band, 802.11n and 802.11ac. Accurate measurements of different technologies have been used in the simulation including the band-width capacity, the end-to-end delay, the communication range and the cost.

Fig. 3 illustrates the performance of different solutions as a function of number of MTC devices $|S|$. The first observation we can draw from this figure is that the increase in the number of MTC devices has a negative impact on the end-to-end delay, the cost and the execution time. The higher number of MTC devices is, the more probability for MTC devices get attach to faraway anchor nodes, and then the higher end-to-end delay gets. Moreover, the increase in the number of MTC devices will increase the number of Hubs should be deployed and the communication technologies should be used, which affect the cost. Fig. 3(a) shows that D-MTC exhibits C-MTC and F-MTC in terms of end-to-end delay. We observe also that both base-line approaches W-0.2 and W-0.8 have similar performances as C-MTC in terms of the end-to-end delay. This can be explained as follows: Based on the observation that the

objective cost dominates the end-to-end delay objective; The objective cost has higher values than the end-to-end delay. Then, whatever the values of α is, the weighting method will opt to minimize the cost that has a great impact on its objective function.

Whereas, Fig. 3(b) depicts that C-MTC exhibits both D-MTC and F-MTC in terms of cost. This is due to the fact that the key objective of D-MTC and C-MTC are the minimization of the end-to-end delay and the cost, respectively. Fig. 3(a) and 3(b) illustrate the fair trade-off achieved by F-MTC between the two conflicting objectives, i.e, the cost and end-to-end delay. From these figures, we observe that F-MTC always finds an optimal trade-off between the end-to-end delay and the cost. From Fig. 3(a), we observe that F-MTC performs similarly to D-MTC in terms of end-to-end delay. Moreover, in the most of time, F-MTC has the same performances like D-MTC in terms of end-to-end delay. Fig. 3(b) depicts that F-MTC has similar performances as C-MTC in terms of cost. This demonstrates that F-MTC successfully achieves the key objective of its design. From Fig. 3(b), W-0.2 and W-0.8 have similar performances as C-MTC in terms of the cost. Thus, both solutions have better performance in terms of costs than D-MTC and F-MTC.

Fig. 3(c) shows the execution time of different solutions. From this figure, the execution time of different solution does

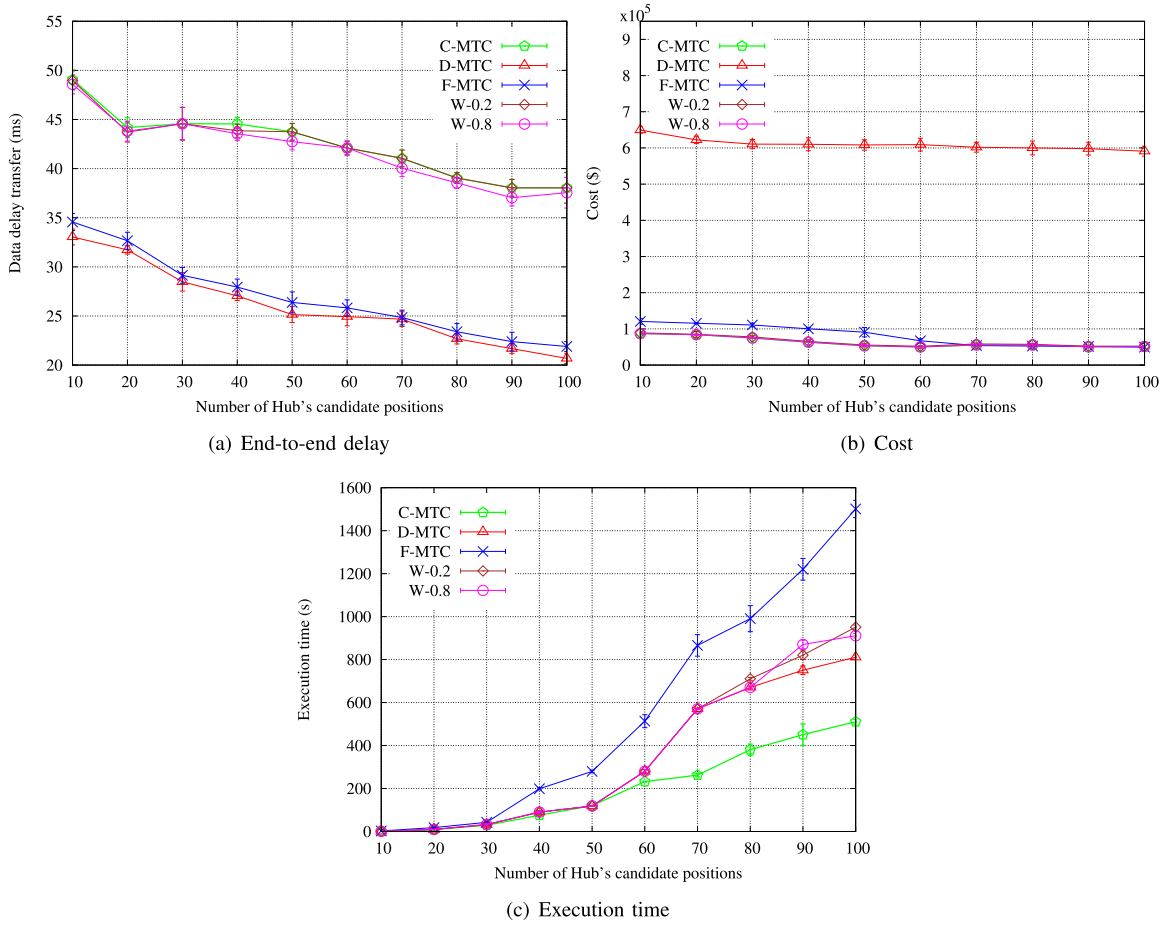


Fig. 4. Performance of the proposed solutions as a function of number of Hub devices.

not exceed 300s, which is reasonable for a deployment that should be used for a while. We observe also that W-0.2 and W-0.8 require more execution time than C-MTC and D-MTC. This is due to the fact that the number of constraints in F-MTC exceeds that of both C-MTC and D-MTC. From this figure, we observe also F-MTC requires more execution time than C-MTC and D-MTC, which is a reasonable cost in comparing to the gain achieved by F-MTC.

Fig. 4 shows the performances of different solutions as a function of number of Hub's candidate positions $|\mathcal{Y}|$. The main observation that we can draw from this figure is $|\mathcal{Y}|$ does not have a great impact on the end-to-end delay and the cost. The end-to-end delay is slightly enhanced when the number of \mathcal{Y} is increased. Clearly, the more number of \mathcal{Y} is, the more possibility to choose closer Hub nodes to MTC devices becomes, and then the higher quality of links between MTC and Hub nodes get. Meanwhile, the cost is not much affected by $|\mathcal{Y}|$. This is because the size of deployed area is small, which makes the number of deployed Hubs and used communication technologies remain the same regardless of the number of \mathcal{Y} . Fig. 4(a) shows that D-MTC and F-MTC outperform C-MTC in term of end-to-end delay, whereas Fig. 4(b) depicts that C-MTC and F-MTC outperform D-MTC in terms of cost. From these figures, we observe that both base-line approaches, W-0.2 and W-0.8, have similar performances as C-MTC in terms of the cost. From these

figures, F-MTC performs similarly to D-MTC in terms of data latency and similarly to C-MTC in terms of cost. Fig. 4(c) shows the performances of different solutions in terms of execution time. From this figure, the execution time of different solutions does not exceed 1600s, which remains reasonable cost for a deployment that should be used for a while.

Based on the formulation of different optimization problem, all the solutions have the same number of variables that are defined as follows: *i*) $|\mathcal{S}| \times |\mathcal{A}| + |\mathcal{Y}| \times (1 + |\mathcal{S}| + |\mathcal{L}| \times \mathcal{K})$ binary variables; *ii*) $|\mathcal{Y}| \times |\mathcal{L}| \times |\mathcal{K}|$ integer variables; *iii*) $|\mathcal{Y}| + |\mathcal{S}| \times (|\mathcal{Y}| + |\mathcal{A}|)$ real variables. Meanwhile, the number of constraints in each solution are defined as follows: *i*) For C-MTC solution, we have at most $1 + |\mathcal{S}| \times (2 + |\mathcal{A}| + |\mathcal{Y}|) + 2|\mathcal{Y}| \times (3 + 4 \times |\mathcal{L}| \times \mathcal{K})$ constrains, whereby \mathcal{K} is the maximum number of neighbors of any nodes using any technologies; *ii*) For D-MTC, we have at most $2 + |\mathcal{S}| \times (1 + |\mathcal{A}| + |\mathcal{Y}|) + |\mathcal{Y}| \times (5 + 8 \times |\mathcal{L}| \times \mathcal{K})$; *iii*) F-MTC requires the parallel executions of both Algorithms C-MTC and D-MTC. Moreover, F-MTC requires the execution of **OP4** that has at most $5 + |\mathcal{S}| \times (2 + |\mathcal{A}| + |\mathcal{Y}|) + |\mathcal{Y}| \times (7 + 8 \times |\mathcal{L}| \times \mathcal{K})$ constraints. From the optimization problems, we can observe that \mathcal{Y} has higher impact on the number of constraints, in all the solutions, than \mathcal{S} . For this reason, from Figures 3(c) and 4(c), we observe that \mathcal{Y} has a higher impact than \mathcal{S} on the execution time.

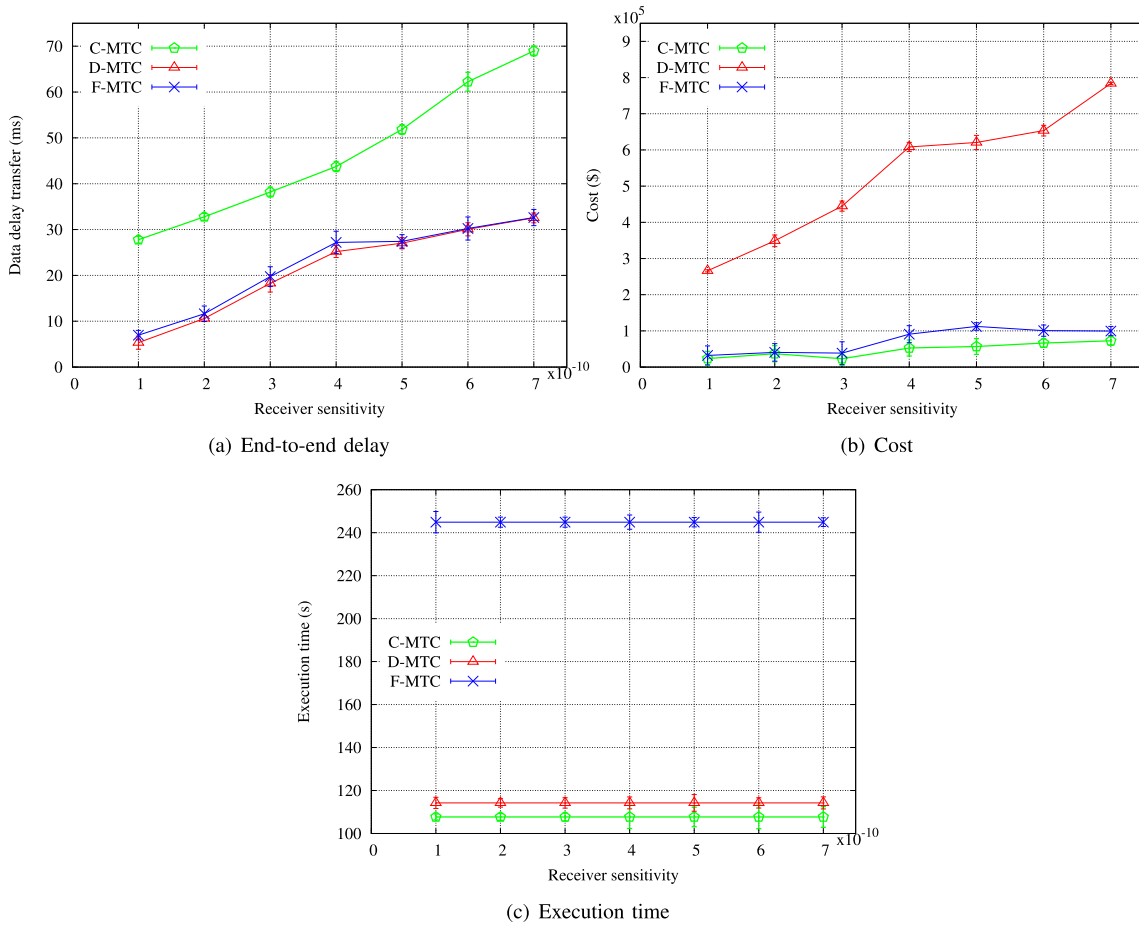


Fig. 5. Performance of the proposed solutions as a function of receiver's sensitivity.

Fig. 5 shows the impact of receiver's sensitivity, on different solutions, in terms of end-to-end delay, cost and execution time. In this figure, we omit the performances of W-0.2 and W-0.8 solutions as they have exactly the same performances like C-MTC in terms of cost and end-to-end delay. The first observation that we can draw from Fig. 5(a) is that an increase in the receiver's sensitivity will have a negative impact on end-to-end delay. From (11) and (12), the outage probability has a negative impact on end-to-end delay between MTC devices and Hubs. The increase in receiver's sensitivity will negatively affect the outage probability that in turn affects the end-to-end delay. Fig. 5(a) and Fig. 5(b) show that F-MTC and D-MTC outperform C-MTC in terms of data latency, whereas F-MTC and C-MTC have better performances than D-MTC in terms of cost. Fig. 5(b) shows that the receiver's sensitivity has a negative impact on D-MTC in terms of cost. This can be explained as follows: When the receiver's sensitivity becomes higher, D-MTC will attach MTC devices to the closest Hubs which do not affect the end-to-end delay. This will increase the number of Hubs deployed and then affect the cost negatively. In contrast to D-MTC, F-MTC, beside the end-to-end delay, takes into account the cost. The achieved results demonstrate the efficiency of F-MTC in achieving the key objective of its design. Fig. 5(c) shows the performances of different solutions in terms of execution time. From this figure, the execution

time of different solutions does not exceed 242s, which is reasonable. We observe also that receiver's sensitivity does not have any impact on the execution time. This can be explained as follows: the receiver's sensitivity does not affect the number of variables and constraints in all the solutions.

VI. CONCLUSION

Enabling mIoT and MTC applications is one of the main objectives of 5G network. This kind of applications is characterized by an important amount of up-link traffic that will create a high overhead on EPS. To deal with this problem, small cells technology has been proposed, where each small cell is responsible for a small number of MTC devices. In this paper, we suggested three solutions for deploying and interconnecting small cells by taking into account wireless communication condition and network features, such as traffic and communication bandwidth. While the first solution, named C-MTC, reduces the cost when deploying and interconnecting small cells, the second one, named D-MTC, reduces as much as possible the end-to-end delay when deploying and interconnecting small cells. The last solution, named F-MTC, uses bargaining game to ensure a fair trade-off between the cost and the end-to-end delay for deploying and interconnecting small cells. The simulation results have proven the efficiency of each solution in achieving its key design objectives.

APPENDIX
PROOF OF THEOREM 1

In this appendix, we derive the proof for the outage probability between two nodes u and v in the network. By definition, the link $u - v$ is in outage if $\text{SINR}_{u,v}$ falls below a threshold level γ_{th} [10]. This event occurs with a probability $\mathcal{P}_{u,v}$. In order to determine an expression for the outage probability $\mathcal{P}_{u,v}$, we need first to compute the CDF of $\text{SINR}_{u,v}$. First, we recall that the expression of the SINR is given by

$$\text{SINR}_{u,v} = \frac{\gamma_{u,v}}{1 + \sum_{t \neq u,v} \gamma_{t,v}}. \quad (\text{A.1})$$

The random variable $\gamma_{u,v}$ which represents the SNR of the desired signal and accounts for both the shadowing and the fast fading have the following PDF

$$p_{\gamma_{u,v}}(\gamma) = \frac{m^m \gamma^{m-1}}{\Gamma(m) \sqrt{\pi}} \sum_{n=1}^N \frac{w_n}{(\Phi_{u,v}(x_n))^m} \exp\left(-\frac{m\gamma}{\Phi_{u,v}(x_n)}\right). \quad (\text{A.2})$$

The CDF of $\gamma_{u,v}$ can be deduced from its PDF as

$$\begin{aligned} F_{\gamma_{u,v}}(z) &= \int_0^z p_{\gamma_{u,v}}(\gamma) d\gamma = \sum_{n=1}^N \frac{w_n}{\sqrt{\pi}} \int_0^z \frac{m^m \gamma^{m-1}}{\Gamma(m) (\Phi_{u,v}(x_n))^m} \\ &\quad \exp\left(-\frac{m\gamma}{\Phi_{u,v}(x_n)}\right) d\gamma \\ &= \sum_{n=1}^N \frac{w_n}{\sqrt{\pi}} \left(1 - \frac{\Gamma\left(m, \frac{mz}{\Phi_{u,v}(x_n)}\right)}{\Gamma(m)}\right), \end{aligned} \quad (\text{A.3})$$

where $\Gamma(\cdot, \cdot)$ is the lower incomplete Gamma function.

Next, our aim is to derive the PDF of the denominator in (A.1). Without loss of generality, we assume that the number of interferers for a given node v is equal to L . For ease of notation, we denote by $Y_{l,v} = \gamma_{l,v}$, thus, the interference term can be written as $\sum_{t \neq u,v} \gamma_{t,v} = \sum_{l=1}^L Y_{l,v} = Y$. The PDF of $Y_{l,v}$ can be deduced from (8) as

$$p_{Y_{l,v}}(\gamma) = \frac{m^m \gamma^{m-1}}{\Gamma(m) \sqrt{\pi}} \sum_{i=1}^N w_i \exp\left(-\frac{m\gamma}{\Phi_{l,v}(x_i)}\right) \frac{1}{(\Phi_{l,v}(x_i))^m}. \quad (\text{A.4})$$

The moment generating function (MGF) can be determined as

$$\begin{aligned} \mathcal{M}_{Y_{l,v}}(s) &= \mathbb{E}(e^{sY_{l,v}}) = \int_0^\infty e^{sY_{l,v}} p_{Y_{l,v}}(\gamma) d\gamma \\ &= \sum_{i=1}^N \frac{w_i}{\sqrt{\pi}} \left(1 - s \frac{\Phi_{l,v}(x_i)}{m}\right)^{-m}. \end{aligned}$$

The MGF of $Y = \sum_{l=1}^L Y_{l,v}$ can be obtained as

$$\begin{aligned} \mathcal{M}_Y(s) &= \prod_{l=1}^L \mathcal{M}_{Y_{l,v}}(s) \\ &= \prod_{l=1}^L \left(\sum_{i=1}^N \frac{w_i}{\sqrt{\pi}} \left(1 - s \frac{\Phi_{l,v}(x_i)}{m}\right)^{-m} \right) = \prod_{l=1}^L \sum_{i=1}^N a_{l,i} \end{aligned} \quad (\text{A.5})$$

where $a_{l,i} = \frac{w_i}{\sqrt{\pi}} \left(1 - s \frac{\Phi_{l,v}(x_i)}{m}\right)^{-m}$

$$\begin{aligned} &\prod_{l=1}^L \sum_{i=1}^N a_{l,i} \\ &= \underbrace{\sum_{i_1=1}^N \sum_{i_2=1}^N \cdots \sum_{i_L=1}^N a_{1,i_1} a_{2,i_2} \cdots a_{L,i_L}}_{L \text{ terms}} \\ &= \sum_{i_1=1}^N \sum_{i_2=1}^N \cdots \sum_{i_L=1}^N \frac{w_{i_1}}{\sqrt{\pi}} \left(1 - s \frac{\Phi_{1,v}(x_{i_1})}{m}\right)^{-m} \\ &\quad \frac{w_{i_2}}{\sqrt{\pi}} \left(1 - s \frac{\Phi_{2,v}(x_{i_2})}{m}\right)^{-m} \cdots \frac{w_{i_L}}{\sqrt{\pi}} \left(1 - s \frac{\Phi_{L,v}(x_{i_L})}{m}\right)^{-m} \\ &= \sum_{i_1=1}^N \sum_{i_2=1}^N \cdots \sum_{i_L=1}^N \left(\prod_{l=1}^L \frac{w_{i_l}}{\sqrt{\pi}} \right) \prod_{l=1}^L (1 - \Phi_l s)^{-m}, \end{aligned} \quad (\text{A.6})$$

where $\Phi_l = \frac{\Phi_{l,v}(x_{i_l})}{m}$. Using the fractional decomposition in [17, eq. (11)], we can write

$$\prod_{l=1}^L (1 - \Phi_l s)^{-m} = \sum_{l=1}^L \sum_{j=1}^m \frac{\beta_{l,j}}{(s - 1/\Phi_l)^j}, \quad (\text{A.7})$$

with

$$\beta_{l,j} = \left(\frac{-1}{\Phi_l}\right)^m \sum_{\tau(l,j)} \prod_{k=1, k \neq l}^L \binom{m + q_k - 1}{q_k} \frac{(\Phi_k)^{q_k}}{\left(1 - \frac{\Phi_k}{\Phi_l}\right)^{m+q_k}}, \quad (\text{A.8})$$

where $\tau(l, j)$ is the set of L -tuples such that $\tau(l, j) = \{(q_1, \dots, q_L) : q_k \in \mathbb{N}, q_l = 0, \sum_{k=1}^L q_k = m - j\}$, with \mathbb{N} the set of nonnegative integers. For the case where $m = 1$, which corresponds to Rayleigh fading channel, we obtain the following expression

$$\prod_{l=1}^L (1 - \Phi_l s)^{-1} = \sum_{l=1}^L \frac{\beta_l}{(s - 1/\Phi_l)^j}, \quad (\text{A.9})$$

with

$$\beta_l = \left(\frac{-1}{\Phi_l}\right) \prod_{k=1, k \neq l}^L \frac{1}{\left(1 - \frac{\Phi_k}{\Phi_l}\right)}. \quad (\text{A.10})$$

Using the identity

$$\mathcal{L}^{-1} \left[\frac{1}{(s - 1/\Phi_l)^j} \right] = \begin{cases} \frac{(-1)^j}{(j-1)!} x^{j-1} e^{-\frac{x}{\Phi_l}} & \text{if } x \geq 0 \\ 0 & \text{otherwise,} \end{cases} \quad (\text{A.11})$$

where $\mathcal{L}^{-1}(\cdot)$ stands for the Inverse Laplace transform. Thus, we can write

$$\mathcal{L}^{-1} \left[\prod_{l=1}^L (1 - \Phi_l s)^{-m} \right] = \sum_{l=1}^L \sum_{j=1}^m \frac{\beta_{l,j} (-1)^j}{(j-1)!} x^{j-1} e^{-\frac{x}{\Phi_l}}. \quad (\text{A.12})$$

It follows that the PDF of Y can be obtained as

$$p_Y(y) = \sum_{i_1=1}^N \sum_{i_2=1}^N \cdots \sum_{i_L=1}^N \left(\prod_{l=1}^L \frac{w_{i_l}}{\sqrt{\pi}} \right) \sum_{l=1}^L \sum_{j=1}^m \frac{\beta_{l,j} (-1)^j}{(j-1)!} y^{j-1} e^{-\frac{y}{\Phi_l}}. \quad (\text{A.13})$$

Let $Y' = Y + 1$ then the PDF of Y' can be determined as

$$p_{Y'}(y) = p_Y(y-1) = \sum_{i_1=1}^N \sum_{i_2=1}^N \cdots \sum_{i_L=1}^N \left(\prod_{l=1}^L \frac{w_{i_l}}{\sqrt{\pi}} \right) \sum_{l=1}^L \sum_{j=1}^m \frac{\beta_{l,j} (-1)^j}{(j-1)!} (y-1)^{j-1} e^{-\frac{y-1}{\Phi_l}}. \quad (\text{A.14})$$

The SINR in (A.1) can be rewritten as $\text{SINR}_{u,v} = \frac{\gamma_{u,v}}{Y'}$. The outage probability $\mathcal{P}_{u,v}$ can be evaluated as

$$\begin{aligned} \mathcal{P}_{u,v} &= P(\text{SINR}_{u,v} \leq \gamma_{\text{th}}) = P\left(\frac{\gamma_{u,v}}{Y'} \leq \gamma_{\text{th}}\right) \\ &= \mathbb{E}_{Y'} [P(\gamma_{u,v} \leq \gamma_{\text{th}} y | Y' = y)] \\ &= \int_1^\infty F_{\gamma_{u,v}}(\gamma_{\text{th}} y) p_{Y'}(y) dy \\ &= \sum_{n=1}^N \frac{w_n}{\sqrt{\pi}} - \sum_{n=1}^N \frac{w_n}{\sqrt{\pi}} \sum_{i_1=1}^N \sum_{i_2=1}^N \cdots \\ &\quad \cdots \sum_{i_L=1}^N \left(\prod_{l=1}^L \frac{w_{i_l}}{\sqrt{\pi}} \right) \sum_{l=1}^L \sum_{j=1}^m \frac{\beta_{l,j} (-1)^j}{(j-1)!} \\ &\quad \int_1^\infty (y-1)^{j-1} e^{-\frac{y-1}{\Phi_l}} \frac{\Gamma\left(m, \frac{m\gamma_{\text{th}} y}{\Phi_{u,v}(x_n)}\right)}{\Gamma(m)} dy. \quad (\text{A.15}) \end{aligned}$$

In (A.15), the notation $\mathbb{E}_{Y'}(\cdot)$ stands for the expectation operation with respect to the random variable Y . Unfortunately, the integral in (A.15) cannot be evaluated in closed form. In the following, we use the Laguerre method to evaluate this integral numerically. For the integral having the form

$$\int_0^\infty e^{-x} f(x) dx = \sum_{p=1}^P \lambda_p f(\theta_p), \quad (\text{A.16})$$

where $\theta_p (p = 1, \dots, P)$ are the zeros of the P th-order Laguerre polynomial and λ_p are weight factors tabulated in [11, Table 25.9].

$$\begin{aligned} &\int_1^\infty (y-1)^{j-1} e^{-\frac{y-1}{\Phi_l}} \frac{\Gamma\left(m, \frac{m\gamma_{\text{th}} y}{\Phi_{u,v}(x_n)}\right)}{\Gamma(m)} dy \\ &= \int_0^\infty y^{j-1} e^{-\frac{y}{\Phi_l}} \frac{\Gamma\left(m, \frac{m\gamma_{\text{th}}(y+1)}{\Phi_{u,v}(x_n)}\right)}{\Gamma(m)} dy \\ &= \int_0^\infty (\Phi_l)^j y^{j-1} e^{-y} \frac{\Gamma\left(m, \frac{m\gamma_{\text{th}}(\Phi_l y+1)}{\Phi_{u,v}(x_n)}\right)}{\Gamma(m)} dy \\ &= \sum_{p=1}^P \lambda_p (\Phi_l)^j (\theta_p)^{j-1} \frac{\Gamma\left(m, \frac{m\gamma_{\text{th}}(\Phi_l \theta_p+1)}{\Phi_{u,v}(x_n)}\right)}{\Gamma(m)}. \quad (\text{A.17}) \end{aligned}$$

Finally, we obtain an analytical expression for the outage probability $\mathcal{P}_{u,v}$ as

$$\begin{aligned} \mathcal{P}_{u,v} &= \sum_{n=1}^N \frac{w_n}{\sqrt{\pi}} - \sum_{n=1}^N \frac{w_n}{\sqrt{\pi}} \sum_{i_1=1}^N \sum_{i_2=1}^N \cdots \sum_{i_L=1}^N \left(\prod_{l=1}^L \frac{w_{i_l}}{\sqrt{\pi}} \right) \\ &\quad \sum_{l=1}^L \sum_{j=1}^m \frac{\beta_{l,j} (-1)^j}{(j-1)!} \sum_{p=1}^P \lambda_p (\Phi_l)^j (\theta_p)^{j-1} \\ &\quad \frac{\Gamma\left(m, \frac{m\gamma_{\text{th}}(\Phi_l \theta_p+1)}{\Phi_{u,v}(x_n)}\right)}{\Gamma(m)}. \quad (\text{A.18}) \end{aligned}$$

For the case where $m = 1$:

$$\begin{aligned} \mathcal{P}_{u,v} &= \sum_{n=1}^N \frac{w_n}{\sqrt{\pi}} + \sum_{n=1}^N \frac{w_n}{\sqrt{\pi}} \sum_{i_1=1}^N \sum_{i_2=1}^N \cdots \sum_{i_L=1}^N \left(\prod_{l=1}^L \frac{w_{i_l}}{\sqrt{\pi}} \right) \\ &\quad \sum_{l=1}^L \beta_l \sum_{p=1}^P \lambda_p \Phi_l \frac{\Gamma\left(m, \frac{m\gamma_{\text{th}}(\Phi_l \theta_p+1)}{\Phi_{u,v}(x_n)}\right)}{\Gamma(m)}. \quad (\text{A.19}) \end{aligned}$$

REFERENCES

- [1] (2016). *Paper, G. W. 5G Mobile Communications Systems for 2020 and Beyond*. [Online]. Available: <http://5gmf.jp/wp/wp-content/uploads/2016/09/5gmfwfp101all.pdf>
- [2] M. Bagaa, A. Chelli, D. Djenouri, T. Taleb, I. Balasingham, and K. Kansanen, "Optimal placement of relay nodes over limited positions in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2205–2219, Apr. 2017.
- [3] Y. Ma and J. Zhang, "Optimisation of Wi-Fi mesh network for small cell backhaul," in *Proc. IEEE Int. Workshop High Mobility Wireless Commun.*, Nov. 2014, pp. 89–93.
- [4] O. Semiari, W. Saad, Z. Dawy, and M. Bennis, "Matching theory for backhaul management in small cell networks with mmWave capabilities," in *Proc. IEEE ICC*, Jun. 2015, pp. 3460–3465.
- [5] O. Grøndalen, O. Østerbø, G. Millstein, and T. Tjelta, "On planning small cell backhaul networks," in *Proc. IEEE EuCNC*, Jun./Jul. 2015, pp. 397–402.
- [6] W. Zhao and S. Wang, "Cell planning for heterogeneous cellular networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2013, pp. 1032–1037.
- [7] X. Ge, L. Pan, S. Tu, H.-H. Chen, and C.-X. Wang, "Wireless backhaul capacity of 5G ultra-dense cellular networks," in *Proc. IEEE 84th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2016, pp. 1–6.
- [8] X. Ge, S. Tu, G. Mao, and C. X. Wang, "5G ultra-dense cellular networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 1, pp. 72–79, Feb. 2016.
- [9] A. F. Molisch, *Wireless Communications*. Hoboken, NJ, USA: Wiley, 2005.
- [10] M. K. Simon and M.-S. Alouini, *Digital Communication Over Fading Channels*, 2nd ed. Hoboken, NJ, USA: Wiley, 2005.
- [11] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*. New York, NY, USA: Dover, 1964.
- [12] W. C. Chan, T.-C. Lu, and R.-J. Chen, "Pollaczek-Khinchin formula for the M/G/1 queue in discrete time with vacations," *IEE Proc.-Comput. Digit. Techn.*, vol. 144, no. 4, pp. 222–226, Jul. 1997.
- [13] A. Chelli, E. Zedini, M.-S. Alouini, J. R. Barry, and M. Pätzold, "Performance and delay analysis of hybrid ARQ with incremental redundancy over double Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6245–6258, Nov. 2014.
- [14] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, 7th ed. New York, NY, USA: Academic, 2007.
- [15] J. F. Nash, Jr., "The bargaining problem," *Econometrica*, vol. 18, no. 2, pp. 155–162, 1950. [Online]. Available: <http://www.jstor.org/stable/1907266>
- [16] E. Kalai and M. Smorodinsky, "Other solutions to Nash's bargaining problem," *Econometrica*, vol. 43, no. 3, pp. 513–518, May 1975. [Online]. Available: <http://www.jstor.org/stable/1914280>
- [17] X. W. Cui, Q. T. Zhang, and Z. M. Feng, "Outage performance for maximal ratio combiner in the presence of unequal-power co-channel interferers," *IEEE Commun. Lett.*, vol. 8, no. 5, pp. 289–291, May 2004.



Miloud Bagaa received the bachelor's, master's, and Ph.D. degrees from the University of Science and Technology Houari Boumediene, Algiers, Algeria, in 2005, 2008, and 2014, respectively. He is currently a Senior Researcher with the Communications and Networking Department, Aalto University. His research interests include wireless sensor networks, Internet of Things, 5G wireless communication, security, and networking modeling.



Ali Chelli (S'08–M'12) was born in Sfax, Tunisia. He received the B.Sc. degree in communications from the Ecole Supérieure des Communications de Tunis in 2005, and the M.Sc. and Ph.D. degrees in information and communication technology from the University of Agder, Norway, in 2007 and 2013, respectively. He was a Post-Doctoral Fellow with the King Abdullah University of Science and Technology from 2012 to 2013. Afterward, he joined the Department of Electronics and Telecommunications, Norwegian University of Science and Technology, Trondheim, as a Researcher, from 2014 to 2017. He is currently a Research Scientist with the University of Agder. His research interests include machine learning for human activity detection and wireless communication theory with focus on performance analysis of wireless sensor networks, cooperative relaying, vehicle-to-vehicle communications, game theory, and channel modeling.



Tarik Taleb received the B.E. degree (Hons.) in information engineering in 2001, and the M.Sc. and Ph.D. degrees in information sciences from Tohoku University, Sendai, Japan, in 2003 and 2005, respectively. He is currently a Professor with the School of Electrical Engineering, Aalto University, Espoo, Finland. He is also the Founder and the Director of the MOSAIC Lab. He is the Guest Editor-in-Chief for the IEEE JSAC series on Network Softwarization and Enablers.



Hamed Hellaoui received the Engineering and master's degrees in computer science from the Ecole Nationale Supérieure d'Informatique, Algeria. He is currently pursuing the Ph.D. degree with Aalto University, Finland. His research interests include unmanned aerial vehicles and Internet of Things.