Ad Hoc Networks 68 (2018) 58–69

Contents lists available at ScienceDirect

Ad Hoc Networks

journal homepage: www.elsevier.com/locate/adhoc

Federated IoT services leveraging 5G technologies at the edge

I. Farris[a], A. Orsino[b], L. Militano[a], A. Iera[a], G. Araniti[a,c,*]

[a] University Mediterranea of Reggio Calabria, DIIES Department, 89124, Reggio Calabria, Italy
[b] Tampere University of Technology, Finland
[c] Peoples' Friendship University of Russia (RUDN University), Russia

ARTICLE INFO

Article history:
Received 18 April 2017
Revised 12 August 2017
Accepted 22 September 2017
Available online 23 September 2017

Keywords:
IoT
Multi-access edge computing
5G
Narrowband-IoT

ABSTRACT

The Internet of Things (IoT) ecosystem is evolving towards the deployment of integrated environments, wherein heterogeneous devices pool their capacities together to match wide-ranging user and service requirements. As a consequence, solutions for efficient and synergistic cooperation among objects acquire great relevance. Along this line, this paper focuses on the adoption of the promising MIFaaS (Mobile-IoT-Federation-as-a-Service) paradigm to support delay-sensitive applications for high-end IoT devices in next-to-come fifth generation (5G) environments. MIFaaS fosters the provisioning of IoT services and applications with low-latency requirements by leveraging cooperation among private/public clouds of IoT objects at the edge of the network. A performance assessment of the MIFaaS paradigm in a cellular 5G environment based on both Long Term Evolution (LTE) and the recent Narrowband IoT (NB-IoT) is presented. Obtained results demonstrate that the proposed solution outperforms classic approaches, highlighting significant benefits derived from the joint use of LTE and NB-IoT bandwidths in terms of increased number of successfully delivered IoT services.

1. Introduction

In the forthcoming fifth generation (5G) cellular infrastructure, the Internet of Things (IoT) transmissions are deemed to reach 7 billion units by 2025 [1]. Besides low-end devices, such as wearable sensors and cameras, characterized by strict resource limitations (e.g., limited memory and processing capabilities), also high-end devices equipped with high computational capabilities (e.g., autonomous cars, drones, robots, virtual reality-based systems) are increasingly populating the IoT [2], opening new market possibilities for highly integrated IoT applications [3]. This evolution brings along with it challenging requirements, especially in terms of increased bandwidth, mobility, and low latency, which need new solutions in both the radio access and the core network infrastructure [4]. To meet these constraints, network and service operators are looking with interest at distributed cloud infrastructures involving the edge of the network and devices on the ground. Not by chance, recently Multi-access Edge Computing (MEC) [5] has attracted the industry attention in view of supporting applications and services with reduced latency and improved QoS, thanks to features such as dense geographical distribution, proximity to consumers, high mobility support, and open platform provision. To this aim, this paper proposes the integration of the core functionalities of *MIFaaS (Mobile-IoT-Federation-as-a-Service)* [6], a recently introduced paradigm based on the synergistic cooperation among private/public sets of IoT objects, with the reference ETSI MEC architecture [7]. The final objective is to evaluate the performance enhancement achievable in integrated IoT scenarios, such as smart cities, including both low-end and high-end IoT devices. Compared to classic collaborative strategies in Mobile Cloud Computing (MCC) [8], where *the single device can establish opportunistic cooperation* by only exploiting its own resources, in MIFaaS *the federation process considers the pool of devices managed by each involved IoT Cloud Provider (ICP).* This new paradigm of cooperation allows to satisfy the exacting demands of high-end IoT services.

A further objective of this paper is to investigate how 5G-oriented radio access technologies (i.e., Long Term Evolution, LTE, and Narrowband-IoT, NB-IoT [9]) and their possible joint use perform in the context of the cited MIFaaS paradigm. To this aim, we have modeled the problem of allocating radio resources in a joint LTE and NB-IoT system, accounting for the specific latency constraints of the served tasks. Given the NP-hard complexity of the investigated problem, we have proposed a heuristic to efficiently allocate the available radio resources to the requested IoT tasks based on their utility. To validate our proposal, an extended simulative campaign has been conducted using Matlab. In doing this, we have focused our attention on three main *high-end IoT*

* Corresponding author.
E-mail addresses: ivan.farris@unirc.it (I. Farris), antonino.orsino@tut.fi (A. Orsino), leonardo.militano@unirc.it (L. Militano), antonio.iera@unirc.it (A. Iera), araniti@unirc.it (G. Araniti).

https://doi.org/10.1016/j.adhoc.2017.09.002
1570-8705/

use cases which present strict requirements in terms of scalability and latency, namely: *(i)* Vehicular-to-Everything (V2X) communications, *(ii)* e-Health services, and *(iii)* video surveillance. As the results show, a wise allocation of tasks among cooperating IoT devices and the exploitation of LTE/NB-IoT resources can significantly improve the service provisioning.

The remainder of the paper is structured as follows. The research background and motivation are presented in the next Section. In Section 3 we define the proposed framework for the ICPs federation in a cellular MEC environment. In Section 4 we present enhanced cellular-oriented spectrum configurations to support heterogeneous IoT tasks. Finally, a performance evaluation is provided in Section 5, whereas conclusive remarks are drawn in Section 6.

2. Research background and motivation

2.1. IoT and cloud

The concept of federation among devices and systems has attracted a high interest within the IoT research communities. The FIESTA-IoT project [10] is an example that promotes interconnectivity among IoT platforms based on federation, unification, and semantic interoperability across multiple IoT domains. However, solutions that exclusively rely on the adoption of Cloud infrastructures and services will unlikely be able to keep up with the expected scalability and network congestion issues in future IoT scenarios. To the purpose of reducing latency for time-critical services, the Fog computing [11,12] and Mobile Edge Computing (MEC) [5] paradigms, along with the Cloudlet concept [13], have recently promoted the idea of relying on a middle tier between end-devices and the remote Cloud. This tier consists of purpose-built servers, access points (APs) or base stations, to offer fast access to Cloud services. This approach enables the offloading of the network/cloud infrastructure, which is currently inadequate to sustain the growing demand for massive ICT services.

On the other hand, recent works focusing on Mobile Cloud Computing (MCC) [8] have proposed infrastructure-less cooperative solutions to create local mobile Cloud [14,15]. In particular, these solutions rely on a distributed cooperation to create collaborative computing platforms wherein devices share their resources. The idea of a resource coordinator elected among the mobile nodes is proposed in [16] to manage the matching between application requests and resources at the devices. In [17], the authors propose the refactoring of the Cloudlet in a controller so that nearby devices can be configured into a coordinated Cloud computing service. However, the aforementioned papers do not consider the heterogeneity of IoT devices in terms of capabilities and resource constraints.

Indeed, the MIFaaS paradigm, we adopt in the present work, moves cooperation from the level of *single device* to the one of *groups of devices* (where each group is owned by a single entity, such as the ICP). In this case, the objective is to increase the performance in terms of satisfied local requests. In addition, MIFaaS exploits Fog Computing at the network edges to increase the reliability of the delivered services. In a previous work [6], we have introduced an initial idea of edge-assisted collaboration model for private IoT clouds and shown the improved performance when compared to alternative device-oriented approaches. In this paper we extend our original idea in order *to cope with low-latency and mobility challenges typical of high-end IoT applications in forthcoming 5G access segments at the edge of a cellular network*.

2.2. Cellular IoT

The efforts of academic, industrial, and standardization bodies are pushing towards the fulfillment of requirements of typical high-end IoT applications by leveraging the next-to-come fifth generation (5G) wireless systems [18–20]. Indeed, the common objective is to design a cellular IoT ecosystem able to support ultra-reliable and ultra-low latency IoT transmissions over licensed bands (e.g., LTE). The idea behind the use of cellular technologies is to provide effective inter-networking of IoT devices through the enhanced coverage, high data rate, low latency, low cost per bit, and high spectrum efficiency that these may offer.

In this context, the Third Generation Partnership Project (3GPP) has introduced, in its latest Release 13, a number of novel key features to support machine-type communications (MTC) [21]. In September 2015, 3GPP decided to standardize *Narrowband IoT* (NB-IoT) as a new narrowband radio technology able to address IoT requirements. This new technology is expected to provide support for a massive number of low-throughput devices, low delay sensitivity, ultra-low device cost, and low device power consumption. Moreover, NB-IoT is also expected to help current legacy cellular systems (i.e., LTE) to offload part of the traffic offered by high-end IoT applications.

The commercial launch of NB-IoT is expected to be around the first quarter of 2017 [22] and it will be released in a form of a software update for the network operators, making it fully backward compatible with existing 3GPP devices and infrastructure. The technology can be either deployed "in-band" by using the resource blocks within a normal LTE carrier (an LTE operator can deploy NB-IoT inside an LTE carrier by allocating one of the Physical Resource Blocks (PRB) of 180 kHz to NB-IoT), or in the unused resource blocks within a LTE carriers guard-band[1], or in "standalone" manner for deployments in dedicated spectrum [23]. As reported in the white paper from Nokia [1], the maximum data rate values (i.e., by considering the overall bandwidth) in terms of instantaneous peak rate provided by the NB-IoT technology are: *170 kbps (down-link, DL)* and *250 kbps (up-link, UL)*. Furthermore, the data rate available for the single tone in downlink and uplink is 680 bits and 1000 bits, respectively. According to these features, it is evident that communication requirements of IoT-based services (typically transferring small data packets) will be easily satisfied. For a summary of the main NB-IoT transmissions characteristics, the reader is referred to the analysis presented in [22].

Even though there is a general consensus that NB-IoT will boost drastically the handling of new application and services driven by next-to-come 5G systems, still some open concerns exist on how to enable devices to efficiently use the NB-IoT spectrum and meet latency and reliability requirements. A possible solution is represented by a joint use and management of LTE and NB-IoT spectra, exploiting the peculiarities of both systems for the services to which they are most suitable.

3. Federation paradigm for cellular IoT cloud providers

We consider an ICP as a set of heterogeneous devices, either belonging to the same private user/company or deployed by a public authority/service provider forming a local Cloud of resources. We introduce a software module called *ICP controller*, that is created when an ICP owner registers the devices, which is in charge of managing resources and services of the corresponding ICP. This module has an aggregate vision of the status of the offered services and available resources (e.g., computation, storage, and sensing) of the ICP. Furthermore, to manage federations of different ICPs on demand, a *MIFaaS Orchestrator* has to be deployed at the edge node. In particular, it is in charge of implementing policies for the efficient formation and maintenance of federations among ICPs,

[1] For instance, for an LTE bandwidth of 10 MHz (i.e., 56 resource blocks – RBs) 6 RBs are reserved for Guard SubCarriers and can be used for NB-IoT.

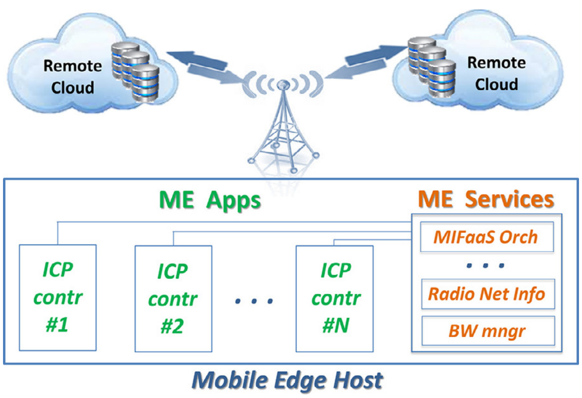


Fig. 1. MIFaaS paradigm.

rapidly scaling the dimension of the federation based on the ICP needs. The MIFaaS Orchestrator implements discovery and registration of resources belonging to new local ICPs, and updates their status by accounting for their utilization and network connectivity.

In ecosystems characterized by high-end IoT services, the need for providing fast decisions requires that the proposed core modules of the MIFaaS architecture have to be deployed as close as possible to the ICPs. To this aim, the emergent cloud-enabled 5G edge cells [24] can represent the ideal environment to deploy both the ICP controllers and the MIFaaS Orchestrator module. To design the integrated architecture of the MIFaaS paradigm in next 5G systems, we explicitly refer to the ETSI MEC guidelines [7]. In particular, the Mobile Edge Hosts (ME Hosts), e.g., cloud-enhanced 5G access points, contain the virtualization infrastructure which provides computing, storage, and network resources for the execution of user Mobile Edge Applications (ME Apps). Furthermore, the ICP controller can be considered as an ME App that an ICP can request to activate. The ME host itself can offer Mobile Edge Services (ME services) to provide additional functionalities to the ME Apps. The ETSI MEC documents specify some basic ME services, such as: *(i)* Radio Network Information, which exposes up-to-date network information to the ME Apps; *(ii)* Location, which offers location information of specific UEs currently served by the radio node associated with the ME host; *(iii)* Bandwidth Manager, which allows allocation of bandwidth to certain traffic routed to and from ME apps and the prioritization of specific traffic. Accounting for these services, we envisage the MIFaaS Orchestrator to be introduced as an additional ME service available to the ICP controllers (see Fig. 1), to cooperate and obtain the desired IoT services.

According to the MIFaaS paradigm, whenever an application running in a device generates a new task request, the ICP controller checks if the relevant ICP has enough resources to serve the request. If this request cannot be directly satisfied or the cost to serve the request is too high, then the ICP controller sends a collaboration request to the MIFaaS Orchestrator of its ME host, which is responsible to take decisions about the creation of federations based on the requests from all the ICP controllers running locally. The MIFaaS Orchestrator checks the status of all the ICPs interested in a federation, and, based on the collected information, decides which resources among those available should resolve the task requests from the different ICPs over a given federation duration (hereafter referred to as *federation period*). Then the list of tasks the ICP devices can provide is created and sent to the LTE scheduler, which will schedule the requests so to meet the desired latency constraints over the overall federation period.

3.1. Federation management at the edge of the network

We consider multiple mobile ICPs under the coverage of a single LTE femtocell. The proposed paradigm creates and orchestrates one or more federations of mobile ICPs to maximize the number of executed tasks while meeting the desired Service Level Agreement (SLA) levels (in terms of latency). The required SLA can be selected according to the reliability requirements of the envisioned IoT services. Accounting for potential mission-critical IoT services, in our analysis we consider the most demanding scenarios. In this case, the task is considered successfully performed only if all the relevant interactions are performed within a specific given time deadline during the federation period. To this aim, the connected ICPs inform the orchestrator about the set of devices capabilities and the set of task requests. Based on this information, possible federations of ICPs are created after having efficiently mapped the service requests onto the available resources. This requires an implementation of a task allocation solution at the MIFaaS Orchestrator. We consider a simple algorithmic solution, based on a greedy approach to map as many task requests as possible to the resources of the involved ICPs, by giving preference to those devices that guarantee the lowest latency for the requested service (in the logic of high-end services provision). In particular, the following steps are defined: *(i)* the adopted task allocation algorithm checks among already allocated tasks if any of the corresponding allocated resources can be reused to also serve the new request (e.g., sensing services); *(ii)* if service reuse is not possible then a suitable device for the service is searched by ordering the available nodes according to the expected latency and the number of currently served tasks; *(iii)* the device with the lowest latency and number of executed tasks is selected to increase the probability of tasks meeting the SLAs and at the same time guarantee some load balancing; *(iv)* the algorithm allocates the task to the corresponding ICP owning that device. Based on the task allocation, the ICP controllers update the resource availability status relevant to all their involved devices.

Accounting for the possible allocated services, the selection of the best configuration of federated ICPs can be naturally modeled as a coalition formation problem (in partition form) where each player (i.e., an ICP) wishes to maximize the associated value in the coalition it belongs to. For each possible federation to form, we associate a single value that gives the measure of how preferred this is with respect to the other federations. Such a value is considered as the difference between the utility the players obtain in cooperation minus a cost term associated to the sharing of their resources. The utility term for any ICP depends on the number of tasks being executed over the number of requested tasks, whereas the cost term is a measure of the amount of resources used to execute all assigned tasks over the total amount of available resources at the ICP. Concerning the analysis shown in this paper, we equally weight the costs and the benefits in the utility even tough other solutions may be considered to better account for the ICPs' willingness to share their own resources. The game is in partition form since the amount of network resources available to a coalition in a cellular setting depends on the resource requests coming from the other coalitions in a given partition [25]. Once the utility is computed for every possible coalition, the MIFaaS orchestrator selects the best ICP federations to form according to an iterative application of so-called *switch operations* that lead to a stable partition (please refer to [6] for more details).

It is worth observing that the service provisioning in a device clearly requires a not negligible activation time for delay-sensitive applications. Therefore, we assume that the involved devices initially setup the relevant service instance and during the federation period the requesting ICP can send multiple task requests (e.g., an application can require either execution of offloaded micro-services or sensing samples every x ms). In this way, the requesting ICP can receive fast responses from the cooperating ICPs during the *federation period*.

As concerns the communication among devices and corresponding ICPs, we assume that the data exchange between federated ICPs happens through the edge node so that interoperability is guaranteed by implementing appropriate syntactical and semantic translations. Clearly, in case of mobility of the ICPs, different edge nodes might need to communicate to transfer the task requests/responses between collaborating ICPs. In the LTE setting considered in this work, the communication between two adjacent femtocells is easily supported through communications over the X2 interface. Of course, the communications involving the forwarding over the X2 interface between the two ME hosts (i.e., the femtocells) will introduce some delay in the service delivery.

4. 5G spectrum allocation for MIFaas

As described in Section 3, once the federation and the task allocation is determined for the involved ICPs, the MIFaaS Orchestrator can further support the federated ICPs by exploiting ME services, such as *Radio Network Information* and *Bandwidth Manager* and manage the data transmission relevant to the involved tasks over the cellular radio resources.

Thus, our goal is to solve the task and spectrum resource allocation while accounting for the task deadlines and the constraints for the available radio resources. Indeed, we have a given amount of spectrum resources with different capacities (the LTE RBs and the NB-IoT tones) available to support MIFaaS tasks. Each task has a specific deadline, a time interval between subsequent requests, and an expected number of messages to be sent during the federation period. The objective is to schedule the task transmissions such that the deadline constraints are met, the constraints on the radio resources are satisfied, and the overall number of solved tasks is maximized (i.e., we note that a task is solved only if all its messages are sent successfully within the given task deadline).

In [26] a multiple robot task allocation with task deadline constraints is studied. This problem is close to our formulation and is modeled as an extension of the Generalized Assignment Problem (GAP), which has NP-hard complexity. However, compared to [26], our problem formulation has further challenges: *(i)* the same task may request different resources over time; *(ii)* a task may request multiple messages to be sent over time which are non-necessarily consecutive; and *(iii)* the resources have different capacities since we have LTE and NB-IoT bands offering different data rate performance.

4.1. Problem formulation

Let us consider a set of tasks $\mathcal{T}_n$ associated to ICP $n \in \mathcal{N}$. For each task t we have a corresponding utility $u_{n,t}$ obtained by the requesting ICP when the task is successfully performed. Each task is allocated for the so-called *federation period* and during this time several interactions are foreseen (e.g., multiple information instances are requested, or multiple computations are made over time). Thus, for each task t of ICP n we define a pool of interactions during the federation period as $\mathcal{I}_{n,t}$. For a task t to be considered successfully performed, all interactions in $\mathcal{I}_{n,t}$ must be performed within a given time deadline, i.e., $D_{n,t}$. Let $y_{n,t,i}$ be a binary variable having value of 1 when the ith interaction of the tth task of ICP n is successfully served within the deadline $D_{n,t}$. We define $x_{n,t}$ as: $x_{n,t} = \prod_{i \in \mathcal{I}_{n,t}} y_{n,t,i}$ so that $x_{n,t}$ is a binary variable having value 1 when the tth task of ICP n is successfully served (all its interactions meet the deadline). More specifically, each interaction i of task t of ICP n includes a set of message exchanges which we refer to as $\mathcal{M}_{n,t,i}$. To identify when a specific interaction is successfully served, we use the variable $y_{n,t,i} = \prod_{m \in \mathcal{M}_{n,t,i}} z_{n,t,i,m}$, with $z_{n,t,i,m}$ being a binary variable having value 1 when the mth message of the ith interaction for the tth task of ICP n is successfully delivered. Thus, the ith interaction of the tth task for ICPS n is considered successfully served only if all the relevant messages are sent over the cellular radio interface. Since we consider only frequency domain scheduling in the LTE system, whenever a message is scheduled, it will be transmitted in a single Transmission Time Interval (TTI) of duration 1 ms. Under this assumption, we define $z_{n,t,i,m} = a^{LTE}_{n,t,i,m} + a^{NB}_{n,t,i,m} \leq 1$, with the two terms indicating whether the specific message is scheduled over LTE or NB-IoT resources respectively.

We define the time duration (expressed in ms) of the ith interaction of the tth task of ICP n as $d_{n,t,i} = \sum_{m \in \mathcal{M}_{n,t,i}} z_{n,t,i,m} \cdot TTI + ET_{n,t,i}$, where $ET_{n,t,i}$ is the execution time for the interaction i of task t for ICP n. Within each *federation period* the number of available time slots is equal to F. In each TTI, the set of available frequency resources that can be allocated in LTE and NB-IoT bands are respectively indicated with $\mathcal{R}^{LTE}$ for a total of R^{LTE} resources and $\mathcal{R}^{NB}$ for a total of R^{NB} resources. The amount of data that can be sent by a device of ICP n (it is reasonable to assume that all its devices have the same channel quality towards the considered femtocell) in a single frequency resource in a specific time slot s out of the F total time slots, is respectively $C^{LTE}_{n,s}$ and $C^{NB}_{n,s}$ (we assume this is not depending on n as all the ICPs adopt MCS = 1 for transmissions).

With the objective to maximize the total utility under the given constraints, the problem can be formulated as follows (please refer to Table 1 for the list of notations used in the paper):

$$
\begin{aligned}
&\max \sum_{n \in \mathcal{N}} \sum_{t \in \mathcal{T}_n} u_{n,t} \cdot x_{n,t} \\
&\text{subject to:} \\
&y_{n,t,i} \cdot d_{n,t,i} \leq D_{n,t}, \ \ \forall n \in \mathcal{N}, \forall t \in \mathcal{T}_n, \forall i \in \mathcal{I}_{n,t} \\
&\sum_{p \in \mathcal{R}^{LTE}} c^{LTE}_{n,t,i,m,s,p} \cdot C^{LTE}_{n,s} \geq r^{LTE}_{t,i,m} \cdot L_{n,t,i,m}, \ \ \forall n \in \mathcal{N}, \forall t \in \mathcal{T}_n, \forall i \in \mathcal{I}_{n,t}, \\
&\quad \forall m \in \mathcal{M}_{n,t,i}, \forall s \in \mathcal{F} \\
&\sum_{p \in \mathcal{R}^{NB}} c^{NB}_{n,t,i,m,s,p} \cdot C^{NB}_{n,s} \geq r^{NB}_{t,i,m} \cdot L_{n,t,i,m}, \ \ \forall n \in \mathcal{N}, \forall t \in \mathcal{T}_n, \forall i \in \mathcal{I}_{n,t}, \\
&\quad \forall m \in \mathcal{M}_{n,t,i}, \forall s \in \mathcal{F} \\
&\sum_{n \in \mathcal{N}} \sum_{t \in \mathcal{T}_n} \sum_{i \in \mathcal{I}_{n,t}} \sum_{m \in \mathcal{M}_{n,t,i}} c^{LTE}_{n,t,i,m,s,p} \leq R^{LTE}, \ \ \forall s \in \mathcal{F}, \forall p \in \mathcal{R}^{LTE} \\
&\sum_{n \in \mathcal{N}} \sum_{t \in \mathcal{T}_n} \sum_{i \in \mathcal{I}_{n,t}} \sum_{m \in \mathcal{M}_{n,t,i}} c^{NB}_{n,t,i,m,s,p} \leq R^{NB}, \ \ \forall s \in \mathcal{F}, \forall p \in \mathcal{R}^{NB},
\end{aligned} \tag{1}
$$

where the constraints for the problem are as follows:

Constraint 1: the ith interaction of the tth task for ICP n is successfully served only if the relevant duration $d_{n,t,i}$ is smaller than the task deadline $D_{n,t}$.

Constraint 2: for each message $r^{LTE}_{t,i,m}$ transmitted over the LTE band, the sum of capacities related to the allocated radio resources must guarantee that all data relative to the message can be transmitted.

Constraint 3: for each message $r^{NB}_{t,i,m}$ transmitted over the NB-IoT band, the sum of capacities related to the allocated radio resources must guarantee that all data relative to the message can be transmitted.

Constraint 4: for each TTI in the federation period, the number of radio resources allocated for the LTE band must be less than the total available resources.

Constraint 5: for each TTI in the federation period, the number of radio resources allocated for the NB-IoT band must be less than the total available resources.

Table 1
List of notations.

Parameter	Definition
$\mathcal{N}$	Set of ICPs
$\mathcal{T}_n$	Set of tasks associated to each ICP $n \in \mathcal{N}$
$u_{n,\,t}$	Utility associated to the execution of task t for ICP n
$D_{n,\,t}$	Deadline for task t (in ms) of ICP n
$x_{n,\,t}$	Binary variable indicating if task t of ICP n is successfully served
$\mathcal{I}_{n,t}$	Set of interactions associated to task t of ICP n
$y_{n,\,t,\,i}$	Binary variable indicating if interaction i of task t of ICP n is successfully served
$d_{n,\,t,\,i}$	Time duration (in ms) for interaction i of task t of ICP n
$\mathcal{M}_{n,t,i}$	Set of messages to be exchanged for interaction i of task t of ICP n
$z_{n,\,t,\,i,\,m}$	Binary variable indicating if message m of interaction i of task t of ICP n is successfully served
$ET_{n,\,t,\,i}$	Execution time for interaction i of task t of ICP n
$L_{n,\,t,\,i,\,m}$	Length **(in bytes)** of the mth message related to i-th interaction of the tth task of ICP n
$a^{LTE}_{n,t,i,m}/a^{NB}_{n,t,i,m}$	Binary variable indicating if message m of interaction i of task t of ICP n is successfully served over either LTE or NB-IoT resources
$c^{LTE}_{n,t,i,m,s,p}/c^{NB}_{n,t,i,m,s,p}$	Binary variable indicating if the pth radio resource in either LTE or NB-IoT bands, in time slot s is used to deliver the mth message of the ith interaction for the tth task of ICP n
F	Number of TTIs in a federation period
$\mathcal{R}^{LTE}$ / $\mathcal{R}^{NB}$	Set of frequency resources units to be allocated in each TTI for either LTE or NB-IoT bands
$C^{LTE}_{n,s}$ / $C^{NB}_{n,s}$	Capacity of frequency resources for ICP n, in time slot s for either LTE or NB-IoT bands

4.2. A heuristic for MIFaas tasks allocation over 5G spectrum resources

Given the complexity of the problem formulated in Section 4.1 (i.e., Eq. (1)), which is a more complicated version of the NP-hard problem studied in [26], we consider a heuristic solution to efficiently allocate the available radio resources. In particular, the proposed heuristic is based on the intuition that the NB-IoT radio resources are expected to be the more suitable for small amount of data, whereas the legacy LTE is better to be exploited for larger messages given the larger transport block size with respect to NB-IoT.

The proposed heuristic, as reported in Algorithm 1, foresees that for each task we compute the ratio between: *(i)* the utility for the requesting ICP when the task is successfully performed; and *(ii)* the radio spectrum cost, which considers all the data expected to be exchanged during the federation period. For every ICP, we create a queue of tasks to be allocated over the cellular spectrum which is sorted in descending order according to the computed ratio. Then, the scheduler progressively selects the task to be allocated the available radio spectrum resources among the ICPs' queues. We adopt a simple Round Robin policy, in order to guarantee fairness among the ICPs.

Algorithm 1: Task allocation to cellular radio resources.

```
Data: Set of ICPs 𝒩
Result: Radio Spectrum Allocation
Phase I - ICP Task Ordering:
for all n ∈ 𝒩 do
    for all t ∈ 𝒯n do
        Compute the ratio between utility and radio spectrum cost for
        the t-th task;
    end
    Order the queue of tasks in descending order of the computed
    ratio;
end
Phase II - Task Scheduling Selection:
MTS=true;
while MTS==true do
    MTS=false;
    for all n ∈ 𝒩 do
        if 𝒯n ≠ ∅ then
            Select the tth task on top of the queue 𝒯n;
            Allocate cellular resources for the tth task;
            𝒯n ← 𝒯n \ t;
            if still resources available then
                MTS==true;
            end
        end
    end
end
```

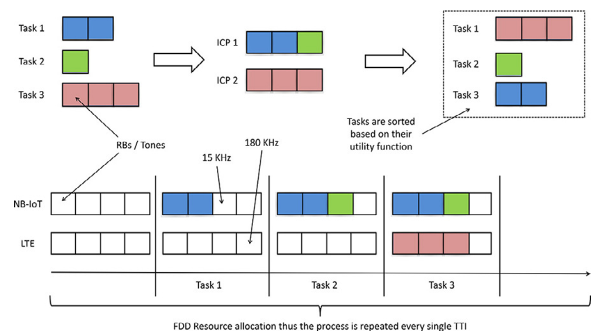


Fig. 2. The proposed tasks allocation over the radio spectrum.

In our study, we consider three spectrum configurations: (i) *only classic LTE* radio resources available for MIFaaS transmissions; (ii) *only NB-IoT* resources available; (iii) *both LTE and NB-IoT* resources available. In this latter case, starting from the tasks at the top of the ordered list, the relevant messages are first allocated to the NB-IoT resources and only when no more NB-IoT resources are available we forward the allocation to the LTE band. With this approach we expect enhanced possibilities to support MIFaaS services with respect to standard LTE systems without NB-IoT resources available (see the performance evaluation campaign in next Section)[2]. For the sake of clarification, the proposed tasks allocation when considering both LTE and NB-IoT spectrum is illustrated in Fig. 2.

5. Performance evaluation in 5G high-end scenarios

5.1. Smart resource allocation for MIFaas: A study case

Among the novel applications expected to be supported by 5G systems, the case of ultra-reliable-low latency communications (URLLC) with strict requirements in terms of latency and reliability is gaining the attention of both academia and industrial communities. Indeed, for URLLC scenarios, low latency (in the level of

[2] Clearly, also alternative policies could be adopted, for instance to consider other priorities on the task allocations.

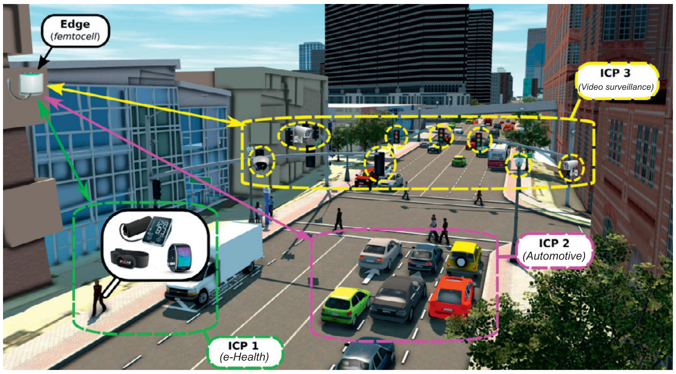


Fig. 3. Exemplary MIFaaS case study for urban environment.

milliseconds) and reliability (five nines, and beyond) together with zero mobility interruption are of utmost importance. This is, for instance, the case of high-end services, such as connected cars, remote surgeries, mobile communications and entertainment.

We focus our analysis on a low-latency 5G scenario sketched in Fig. 3, envisioning the cooperation of ICPs for augmented cognitive assistance in urban environment. Indeed, in this scenario we may have different high-end IoT devices that provide a broad range of services. For instance, a *video surveillance* system may own a number of cameras dislocated across the city (i.e., on walls or traffic lights) for monitoring and security purposes. Furthermore, as part of an e-Health system, various users may move in the city equipped with *wearable devices* able to constantly monitor their health conditions and trigger appropriate alerts/actions according to dangerous situations. Finally, according to the paradigm of Vehicular-to-Everything (V2X) communications, we may consider the presence of *vehicles* which are grouped in fleets, or involved in platooning systems, to provide common digital-enhanced transportation services. Noteworthy, vehicles are equipped with computational and storage capabilities, making them able to find the fastest routes across the city, avoid traffic jams, and guarantee access to up-to-date security warnings.

In the described scenario, different challenges emerge to guarantee latency and reliability constraints typical of high-end IoT-based applications. For instance, vehicles need real-time mobility information integrating as many sources as possible and they often lack the means to obtain reliable, accurate and updated data about traffic conditions in the city. On the other hand, they are able to process high amounts of data thanks to the on-board distributed processing and storage resources. The users equipped with e-Health wearable devices, instead, aim at monitoring their health condition by gathering information from their body and processing them through appropriate data analytics algorithms. Here the bottleneck is represented by the limited computing resources to guarantee the expected Quality of Experience (QoE). Finally, the video surveillance company deploys a number of cameras across the urban environment to generate videos in support of different applications such as: (i) monitoring the traffic jam, (ii) managing the travel demand by rationing the traffic during peak hours or during peak pollution events, (iii) providing real-time security information to drivers, or (iv) controlling traffic lights to grant priority to emergency vehicles. Given the heterogeneous nature of these applications, it is evident that a wide range of requirements and constraints have to be taken into account (e.g., latency, reliability, and efficient connectivity).

In our analysis, we consider that all the ICPs described above interact with each other to provide integrated applications and mutual support. Indeed the video surveillance company, the user carrying e-Health wearable devices, and the vehicles would significantly benefit from the MIFaaS federation. This feature, in fact, allows them to share their unused local resources, as it is the case of the computational and storage capabilities on the smart vehicles, and to experience a faster access to sensing resources that could not be available on the Cloud yet. In this context, we envision the following interactions:

- The vehicles may benefit from data streams provided by the person's e-Health wearable sensors and by the video surveillance company infrastructure in collecting the information needed to compute the fastest and safest routes. In turn, computing and storage facilities of the vehicles are offered.
- The user shares contextual information acquired with his personal sensing devices and, as reward, he may gain the access to computation facilities of the vehicles fleet to offload part of the application processing and data elaboration efforts.
- The video surveillance company offers its local video streams, and receives both support from the computation facilities of the vehicles fleet (to implement the local processing operations towards proper urban management) and context data from e-Health wearable of the users (to enhance the security level in urban environment).

5.2. Performance evaluation setup

The reported system-level simulations have been performed by using a framework developed in Matlab. For the purpose of this study, we consider a urban scenario covering an area of [500 $\times$ 500] m, where edge nodes (i.e., represented by LTE femtocells[3]) are scattered according to a grid topology with an equi-spatially inter-site distance of 100m. In addition, three different types of mobility settings are considered for the ICPs: (i) static, (ii) pedestrian, and (iii) vehicular mobility. Pedestrian movements are modeled according to the Levy Flight mobility model with an alpha parameter equal to $\alpha = 1$ [28]. Although Levy Flight is the most used mobility model to characterize human behavior in urban environment, it is not applicable when considering vehicles traffic patterns. Thus, for vehicular mobility we exploit the Random Direction mobility model [29] with an average speed of 45 km/h. Despite the Manhattan model is more frequently used to characterize vehicular patterns, most of the cities do not adhere to the Manhattan grid deployment (including, to a large extent, Manhattan island itself). Instead, the Random Direction mobility model can be easily adapted to act according to the Manhattan scheme.

In the reference scenario, the tasks that may be executed by the ICPs can be split into six different classes: *computation offloading, data caching*, and four different *sensing services* (e.g., video streaming, warning messages, temperature, and sensing for wearable devices). Each of them has different packet sizes (i.e., for both request and response messages), latency requirements, time intervals between two consecutive requests, and execution times. More details about the task configuration and relevant parameters are provided in Table 2.

Furthermore, we adopt the legacy *LTE* band as our benchmark, where applications are managed by using licensed bands (i.e., in our study we consider a 1.4 MHz bandwidth, corresponding to 6 RBs). In addition, we consider an *NB-IoT* guard band solution where radio resources are taken from those RBs that normally are used as guard bands within the standard LTE carrier (i.e., referred to our case we use the 6 "guard" RBs for the 10 MHz LTE bandwidth, which corresponds to 320 NB-IoT tones). Finally, we consider a hybrid solution (i.e., hereafter named *NB-IoT + LTE*) that ex-

[3] Noteworthy, edge nodes are interconnected through X2 LTE protocol and channels [27].

Table 2
Task configuration parameters.

ID Task	Task type	Latency thr	Pck size request	Pck size response	Time interval	Resource reqs	Execution time
1	Computation *(video processing)*	[30,60,100] ms	[200,250,600] Kbits	[200,250,600] Kbits	100 ms	[2, 4, 6] MFLOPS	25 ms
2	Storage *(video download)*	[30,60,100] ms	0.1 Kbits	[200,250,600] Kbits	100 ms	[10, 30, 50] MB	1 ms
3	Sensing *(video streaming)*	10 ms	0.1 Kbits	80 Kbits	10 ms	Available sensor	1 ms
4	Sensing *(warning messages)*	25 ms	0.1 Kbits	25 Kbits	25 ms	Available sensor	1 ms
5	Sensing *(blood pressure)*	50 ms	0.1 Kbits	0.5 kbits	50 ms	Available sensor	1 ms
6	Sensing *(heart rate)*	50 ms	0.1 Kbits	0.5 kbits	50 ms	Available sensor	1 ms

* Task ID 1 and 2 refer to the video surveillance service, Task ID 3 and 4 to the automotive service, whereas Task ID 5 and 6 refer to the e-Health service.

ploits both NB-IoT and LTE radio resources according to the heuristic discussed in Section 4.

To provide a comprehensive understanding of the benefits introduced by the MiFaaS solution, within the above considered spectrum configurations, the tasks of the involved ICPs are handled by considering four different modalities:

- *No cooperation* – (No Fed): no cooperation is implemented, i.e., each ICP operates as a stand-alone entity, by relying only on its own resources.
- *No cooperation* – (No Fed+Cloud+Edge): the single ICPs work as stand-alone entities without forming cooperative federations; they can exploit resources belonging to the available remote Cloud and to the constrained Edge node. In particular, the Cloud can especially offer support for delay-tolerant tasks, since the Round Trip Time to remote data center is typically not negligible.
- *Cooperation-based MIFaaS* – (Fed): this is the proposed paradigm providing federation of ICPs.
- *Cooperation-based MIFaaS* – (Fed+Cloud+Edge): this solution refers to the proposed paradigm of ICPs federation, with additional computation and storage resources provided by the traditional remote Cloud and by the resource-constrained edge node.

The main metrics for performance assessment are: (i) *percentage of served tasks*, defined as the percentage of MIFaaS tasks that are effectively and successfully served, and (ii) *transport block size utilization*, characterizing the ratio between the amount of data sent over the used radio resources to execute the tasks and the amount of data that can be transferred over the total available resources.

Further details on the simulation settings are reported in Table 3.

Table 3
Main simulation parameters.

Parameter	Value
Number of task requests per ICP	3
Number of sensing types	4
Sensing units per task request	[1–10]
Computation capacity per device	[1–10] MFLOPS
Storage resource units per device	[10–100] MBs
Sensing resource units per device	[0–1]
Computation capacity of Cloud	Unlimited
Storage resource units of Cloud	Unlimited
Computation capacity of Edge node	10 MFLOPS
Storage resource units of Edge node	100 MBs
Number of ICPs	[4–24]
Number of devices per ICP	3
Mobility model ICP pedestrian	Levy Flights ($\alpha = 1$)
Mobility model ICP vehicular	Random Direction
Vehicular speed ICP	45 km/h
Round Trip Time Remote Cloud	80 ms
Round Trip Time X2 inter-edge link	10 ms
Coverage diameter edge	100 m
NB-IoT bandwidth per RB	180 KHz
# NB-IoT tones (subcarriers)	288
# LTE RBs	6
Simulation runs	100

5.3. Simulative results

5.3.1. Static scenario

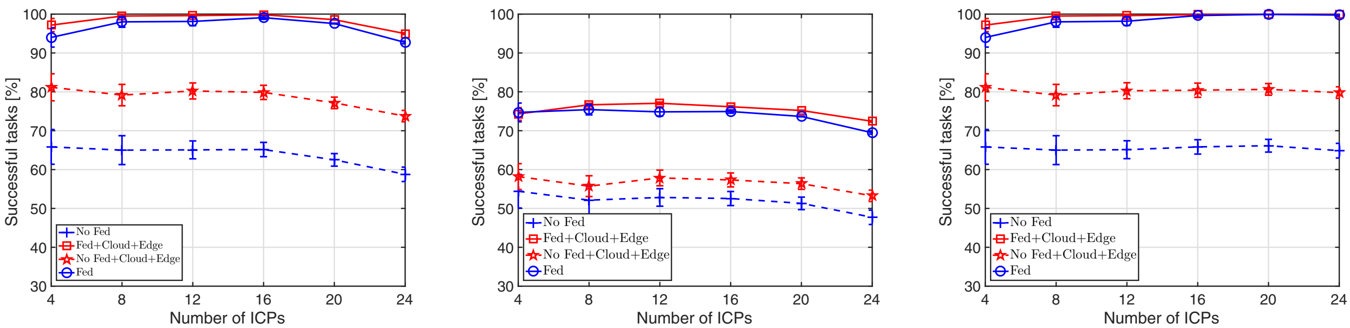
In the first analysis we present, the ICPs are deployed within the coverage of one edge node and do not move over time. As a consequence, no wireless channel fluctuations occur and we have a constant system-level performance during the whole federation period. The initial focus is on the percentage of successfully completed tasks when varying the number of ICPs (in the range [4–24]). As we can observe from Fig. 4, the proposed *MIFaaS-based* solutions (i.e., *Fed* and *Fed+Cloud+Edge*) outperform other solutions in all the three spectrum configurations. However, in case of NB-IoT resources only, as expected, the percentage of successful tasks is lower than the other cases, reaching around 75%. Whereas, it reaches 100% for the LTE and NB-IoT+LTE spectrum configurations. It is also possible to observe that in LTE and NB-IoT cases there is a general decrease in the performance with a number of ICPs greater than 16, due to the absence of sufficient radio spectrum resources. Differently, a joint use of LTE and NB-IoT bandwidth is able to guarantee a percentage of 100% also beyond 16 ICPs.

The trend observed in Fig. 4 is confirmed when varying the federation period from 5 to 30 s (i.e., see Fig. 5 where we considered a scenario with 24 ICPs). The joint NB-IoT+LTE spectrum usage successfully supports all the tasks even when the federation period becomes longer (e.g., 30 s). Noteworthy, the constant trend in all considered solution is due to the static nature of the ICPs whose channel quality conditions do not change during the considered time interval.

5.3.2. Pedestrian scenario

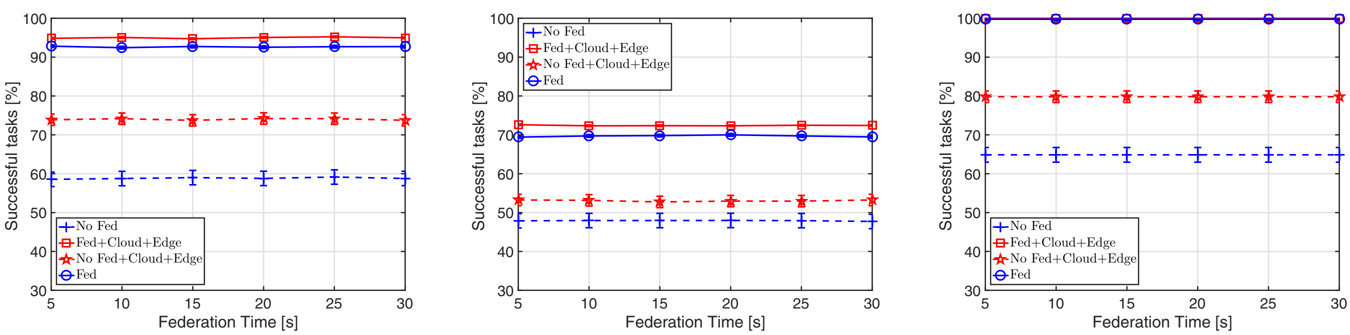
In the pedestrian scenario, the ICPs freely move across the given area according the to the Levy Flight mobility model with an α value equal to 1. Differently from the static case, here the ICPs are experiencing small changes in their channel quality conditions (they are moving at a low speed). As a consequence, in Fig. 6, it is possible to observe a decrease in the percentage of tasks successfully executed for all solutions with respect to the static case (a federation period of 30 s is considered). Moreover, due to mobility, ICPs can move across the coverage areas of different edge nodes during the federation period. Accounting for the RTT latency over X2 links that interconnect different edges, the delay accumulated by the exchanged messages between two cooperating ICPs can exceed the relevant task deadline, thus causing a failure in the task offloading. A further critical feature is the limited overall radio spectrum availability. As a consequence, the *MIFaaS+Cloud+Edge* approach allows to achieve a percentage of successful tasks around 85% in NB-IoT+LTE spectrum configuration. This is a higher value compared to the other spectrum configurations, that reach 81% and 67% (on average) respectively.

Another interesting result in Fig. 6 and Fig. 6(c) is that the approaches that rely on Cloud and edge support achieve the best



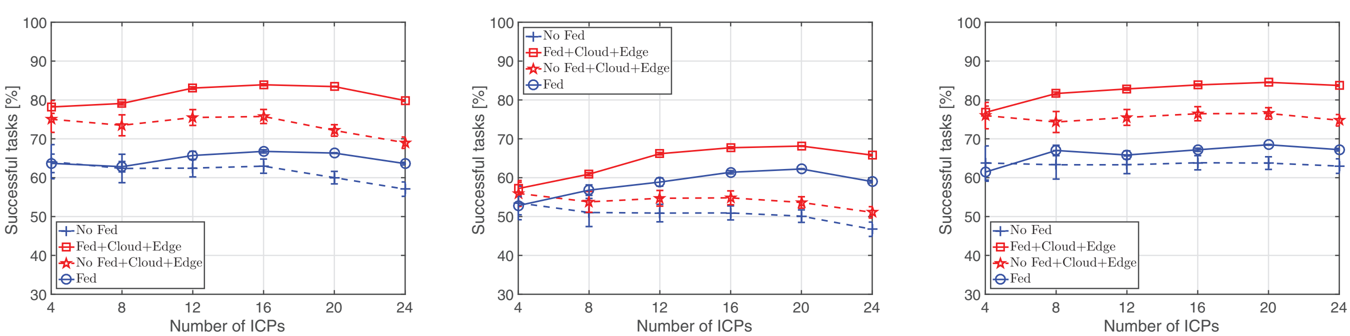
(a) LTE (b) NB-IoT (c) NB-IoT + LTE

Fig. 4. Percentage of successful tasks by varying the number of static ICPs.



(a) LTE (b) NB-IoT (c) NB-IoT + LTE

Fig. 5. Percentage of successful tasks by varying the federation period (24 static ICPs).



(a) LTE (b) NB-IoT (c) NB-IoT + LTE

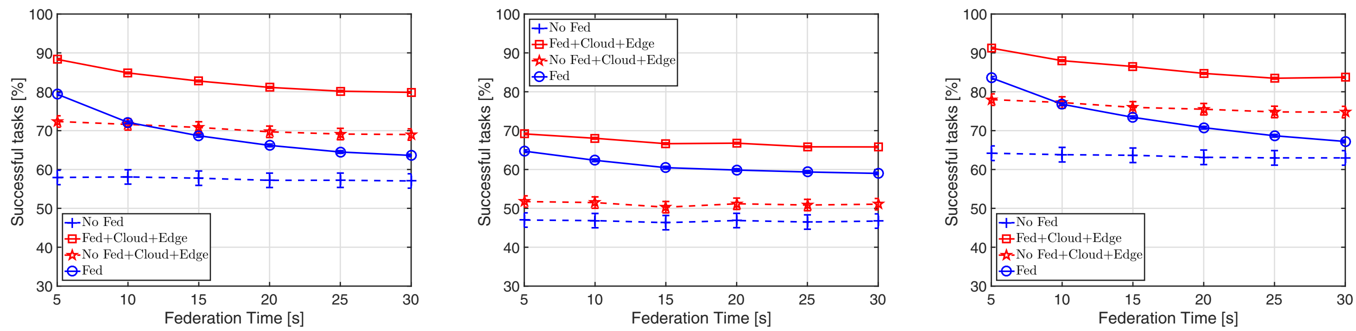
Fig. 6. Percentage of successful tasks by varying the number of ICPs with pedestrian mobility.

performance. This is because, according to our policy, tasks delegated to the Cloud have a typical delay-tolerant nature; therefore, mobility has no negative impact on the probability of their successful offloading. Differently, in the NB-IoT configuration, the *No Fed+Cloud+Edge* solution shows a lower percentage of successful tasks compared to the *MIFaaS-based* solution when increasing the number of involved ICPs. This is because the tasks offloaded to the Cloud are the most radio resource consuming and, therefore, the NB-IoT spectrum configuration is unable to properly support offloading of tasks towards the Cloud.

To characterize the impact of the federation period, in Fig. 7 we report the system-level performance for the sample case with 24 ICPs. As we can observe, the percentage of successfully delivered tasks decreases when the federation period increases. This is explained by considering that the longer the federation period is, the higher the probability that two cooperating ICPs move away from the same edge node, which causes an increased message exchange time and a higher risk to miss the task deadline. As shown in Fig. 7(b), the use of the NB-IoT bandwidth configuration leads to the worst performance. Instead, a combination of NB-IoT+LTE spectrum configuration guarantees a 5% gain (i.e., see Fig. 7(c)) with respect to the LTE configuration, which reaches a maximum of 80% of successful tasks (as reported in Fig. 7(a) for a federation period of 30 s).

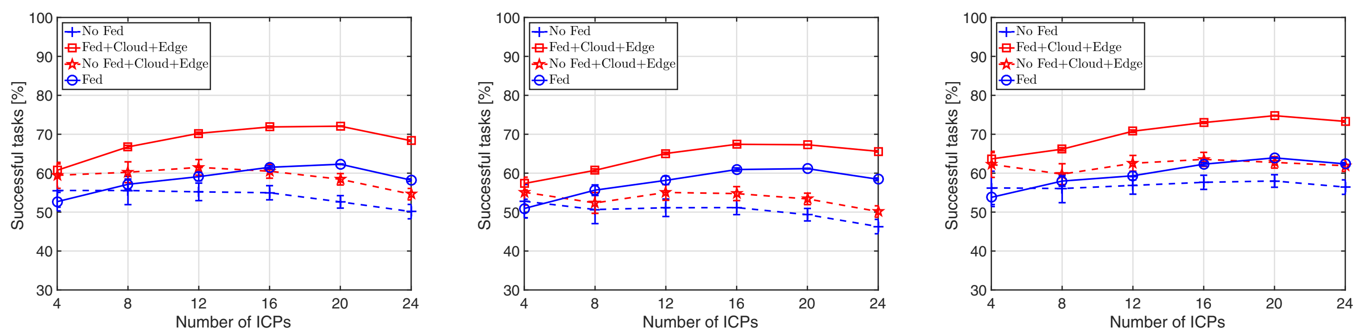
5.3.3. Vehicular scenario

In this last analysis we have assessed the performance of the proposed approach accounting for the impact of vehicular mobility. In particular, the deployed ICPs move according to the Random Direction mobility model with an average speed of 45 km/h. Since we are focusing on a urban scenario, the considered speed value is reasonable as, in most of the cities, cars have to adapt



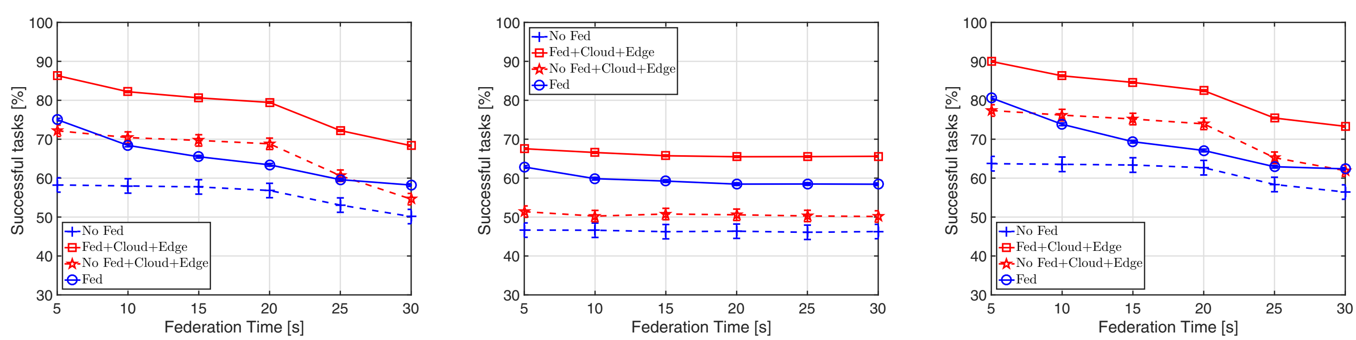
(a) LTE (b) NB-IoT (c) NB-IoT + LTE

Fig. 7. Percentage of successful tasks by varying the federation period in the pedestrian scenario.



(a) LTE (b) NB-IoT (c) NB-IoT + LTE

Fig. 8. Percentage of successful tasks by varying the number of ICPs with vehicular mobility.



(a) LTE (b) NB-IoT (c) NB-IoT + LTE

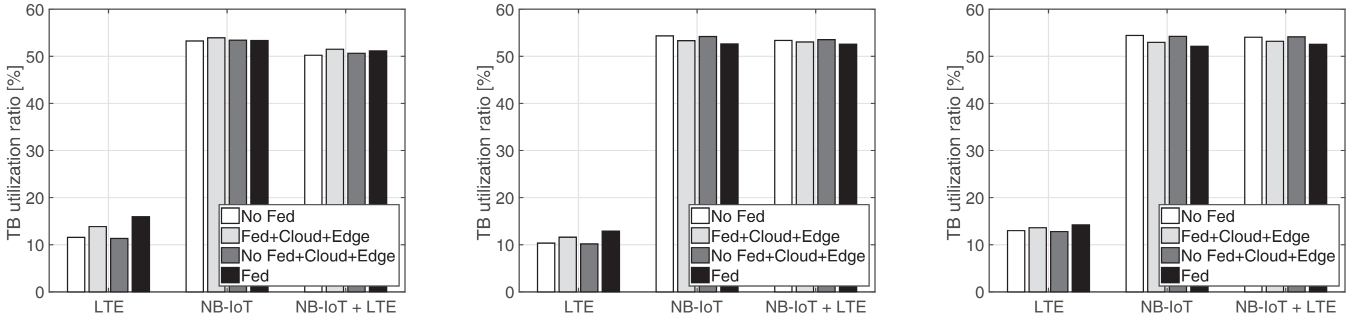
Fig. 9. Percentage of successful tasks by varying the federation period in a vehicular scenario.

to stringent speed limits (typically between 40 km/h and 60 km/h). Compared to previous results, in Fig. 8 we can observe that mobility has a considerable impact on the system-level performance. In particular, the three spectrum configurations obtain, on average, the same percentage of successful tasks with a slight gain observed for the NB-IoT+LTE configuration (namely 5% and 8% with respect to LTE and NB-IoT configurations, respectively). In addition, similar to the static and pedestrian scenarios, the percentage of successfully solved tasks begins to decrease when the ICPs are more than 16.

Interestingly, when varying the federation period (i.e., see Fig. 9) the decreasing performance trend is more evident for the LTE and NB-IoT+LTE solutions (i.e., see Fig. 9(a) and (b)) compared to the NB-IoT one. However, as shown in Fig. 9(b) the NB-IoT+LTE spectrum configuration outperforms the other configurations, allowing a higher percentage of successfully delivered tasks for all the federation period values considered. The obtained results provided useful insights for the system designer to select an efficient federation period. Indeed, when LTE cellular radio access is exploited (i.e., Fig. 9(a) and (b)), the percentage of successful tasks presents a clear degradation if the federation period increases over 20 seconds due to the ICPs mobility. This trend suggests that the federation period should be chosen shorter than 20 seconds, so to avoid the performance degradation in the vehicular scenario.

5.4. Impact of the considered spectrum configuration on the transport block size

The impact of the considered spectrum configuration on the transport block size utilization (with 24 ICPs and federation period equal to 30 s) is shown in Fig. 10. This metric points out how efficiently radio resources are exploited in managing the given tasks. It is worth noticing that for all the three cases considered (i.e., static, pedestrian, and vehicular) the results obtained are almost



(a) Static scenario (b) Pedestrian scenario (c) Vehicular scenario

Fig. 10. Transport block utilization.

the same. The explanation behind this trend is that the transport utilization size is not something that is strictly related to the mobility or the spectrum configuration used. In fact, the relevant metric that has a strong impact on the performance is represented by the considered packet size, that, in our analysis, differs among the investigated use cases (i.e., see columns "packet size request" and "packet size response" of Table 2).

Along these lines, as expected LTE shows very low efficiency as it utilizes less than 15% of the available radio spectrum, thus leading to a remarkable amount of wasted radio resources. Instead, the NB-IoT and NB-IoT+LTE configurations guarantee a more efficient spectrum usage, by reaching a value that is around 54% and 51%, respectively. According to the lack of radio resource that is affecting the current wireless/cellular systems, we may claim that this result can be extremely useful to the network operators. Indeed, an efficient exploitation of the available spectrum resources to meet subscribers' demands and maximize their profits plays a crucial role in meet the demanding requirements of most high-end IoT use cases.

6. Conclusion and future works

In this paper we investigated the Mobile-IoT-Federation-as-a-Service (MIFaaS) paradigm in 5G cellular systems under latency constraints. In particular, the performance assessment has been conducted in three different mobility scenarios, thus providing useful indications on the impact of mobile ICPs supported for different federation periods. Our first conclusion is that the proposed federation of ICPs enhances the number of solved tasks with respect to alternative solutions where cooperation is not implemented. Then, as a further achievement, the presented results showed that in handling high-end IoT data traffic, a combination between NB-IoT and LTE is essential in providing the needed high data rate and low latency.

As a future work analysis, we plan to investigate the definition of suitable ICP federation policies based on several refined criteria to optimize the cooperation at the edge of the network, such as: (i) the ability to *predict the mobility pattern* of the ICPs; (ii) the *power consumption* associated to each task aiming at improving energy efficiency of high-end IoT devices; (iii) *trustworthy interactions* among ICPs to build a reliable and secure system.

Acknowledgment

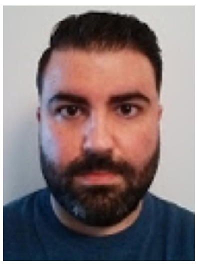
The publication was financially supported by the Ministry of Education and Science of the Russian Federation (Agreement number 02.a03.21.0008).

References

[1] LTE evolution for IoT connectivity white paper, Nokia White Paper, 2016.
[2] O. Hahm, E. Baccelli, H. Petersen, N. Tsiftes, Operating systems for low-end devices in the internet of things: a survey, IEEE Internet Things J. 3 (5) (2016) 720–734, doi:10.1109/JIOT.2015.2505901.
[3] Manufacturing Reengineered: Robots, 5G and the Industrial IoT, Technical Report, Ericsson Business Review, 2015.
[4] A. Orsino, A. Ometov, G. Fodor, D. Moltchanov, L. Militano, S. Andreev, O.N.C. Yilmaz, T. Tirronen, J. Torsner, G. Araniti, A. Iera, M. Dohler, Y. Koucheryavy, Effects of heterogeneous mobility on D2D- and drone-assisted mission-critical MTC in 5G, IEEE Commun. Mag. 55 (2) (2017) 79–87, doi:10.1109/MCOM.2017.1600443CM.
[5] Y.C. Hu, M. Patel, D. Sabella, N. Sprecher, V. Young, Mobile edge computing a key technology towards 5G, ETSI White Paper, 11, 2015.
[6] I. Farris, L. Militano, M. Nitti, L. Atzori, A. Iera, Mifaas: a mobile-iot-federation-as-a-service model for dynamic cooperation of iot cloud providers, Future Gener. Comput. Syst. 70 (2017) 126–137. https://doi.org/10.1016/j.future.2016.06.028.
[7] Mobile Edge Computing (MEC): Framework and Reference Architecture, Technical Report, ETSI GS MEC 003 V1.1.1, 2016.
[8] N. Fernando, S.W. Loke, W. Rahayu, Mobile cloud computing: a survey, Future Gener. Comput. Syst. 29 (1) (2013).
[9] J. Gozalvez, New 3GPP standard for IoT [mobile radio], IEEE Veh. Technol. Mag. 11 (1) (2016) 14–20, doi:10.1109/MVT.2015.2512358.
[10] A. Gyrard, C. Bonnet, K. Boudaoud, M. Serrano, Assisting IoT projects and developers in designing interoperable semantic web of things applications, in: 2015 IEEE International Conference on Data Science and Data Intensive Systems, 2015, pp. 659–666, doi:10.1109/DSDIS.2015.60.
[11] F. Bonomi, R. Milito, J. Zhu, S. Addepalli, Fog computing and its role in the internet of things, in: Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing, ACM, 2012, pp. 13–16.
[12] M. Chiang, Fog networking: an overview on research opportunities, arXiv preprint arXiv:1601.00835 (2016).
[13] U. Shaukat, E. Ahmed, Z. Anwar, F. Xia, Cloudlet architectures, applications, and open challenges to deployment in local area wireless networks, J. Netw. Comput. Appl. (2015). Elsevier.
[14] C. Shi, V. Lakafosis, M.H. Ammar, E.W. Zegura, Serendipity: enabling remote computing among intermittently connected mobile devices, in: Proceedings of the 13-th ACM international symposium on Mobile Ad Hoc Networking and Computing, ACM, 2012, pp. 145–154.
[15] E. Miluzzo, R. Cáceres, Y.-F. Chen, Vision: mClouds-computing on clouds of mobile devices, in: Proceedings of the Third ACM Workshop on Mobile Cloud Computing and Services, ACM, 2012, pp. 9–14.
[16] T. Nishio, R. Shinkuma, T. Takahashi, N.B. Mandayam, Service-oriented heterogeneous resource sharing for optimizing service latency in mobile cloud, in: Proceedings of the First International Workshop on Mobile Cloud Computing & Networking, ACM, 2013, pp. 19–26.
[17] K. Habak, M. Ammar, K.A. Harras, E. Zegura, FemtoClouds: leveraging mobile devices to provide cloud service at the edge, IEEE Cloud, 2015.
[18] A. Gupta, R.K. Jha, A survey of 5g network: architecture and emerging technologies, Access IEEE 3 (2015) 1206–1232.
[19] J. Sachs, N. Beijar, P. Elmdahl, J. Melen, F. Militano, P. Salmela, Capillary Networks - A Smart Way to Get Things Connected, vol. 8, Ericsson, 2014, pp. 1–8.
[20] S. Andreev, O. Galinina, A. Pyattaev, M. Gerasimenko, T. Tirronen, J. Torsner, J. Sachs, M. Dohler, Y. Koucheryavy, Understanding the iot connectivity landscape: a contemporary m2m radio technology roadmap, Commun. Mag. IEEE 53 (9) (2015).
[21] 3GPP, TS 22.368, Service requirements for machine-type communications (MTC), V13.1.0, Technical Report, 2014.
[22] Y.-P. E. Wang, X. Lin, A. Adhikary, A. Grövlen, Y. Sui, Y. Blankenship, J. Bergman, H.S. Razaghi, A primer on 3gpp narrowband internet of things (nb-iot), arXiv preprint arXiv:1606.04171 (2016).
[23] 3GPP, TSG RAN Meeting #69, Narrowband IoT, Technical Report, 2015.

[24] J.O. Fajardo, F. Liberal, I. Giannoulakis, E. Kafetzakis, V. Pii, I. Trajkovska, T.M. Bohnert, L. Goratti, R. Riggio, J.G. Lloreda, et al., Introducing mobile edge computing capabilities through distributed 5g cloud enabled small cells, Mobile Netw. Appl. 21 (4) (2016) 564–574.
[25] W. Saad, Z. Han, M. Debbah, A. Hjørungnes, A distributed coalition formation framework for fair user cooperation in wireless networks, Wireless Commun. IEEE Trans. 8 (9) (2009) 4580–4593.
[26] L. Luo, N. Chakraborty, K. Sycara, Distributed algorithms for multirobot task assignment with task deadline constraints, IEEE Trans. Autom. Sci. Eng. 12 (3) (2015) 876–888, doi:10.1109/TASE.2015.2438032.
[27] 3GPP, TS 36.423, Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 Application Protocol (X2AP), Technical Report, 2014.
[28] K. Zhao, M. Musolesi, P. Hui, W. Rao, S. Tarkoma, Explaining the power-law distribution of human mobility through transportation modality decomposition, Sci. Rep. 5 (2015). 9136 EP –
[29] R.R. Roy, Realistic Random Direction Mobility, Springer US, Boston, MA, pp. 223–244. 10.1007/978-1-4419-6050-4_7.

Ivan Farris is pursuing his Ph.D. degree in Information Technology Engineering at the University of Reggio Calabria, Italy. He received a B.Sc. degree in Telecommunications Engineering and a M.Sc. degree in Computer and Telecommunications Systems Engineering from the University of Reggio Calabria in 2011 and 2013, respectively. His research interests include Internet of Things, Edge computing, and network softwarization.



Antonino Orsino is currently an Experienced Research at Ericsson Research, Finland. He is also a research assistant in the Department of Electronics and Communications Engineering at Tampere University of Technology, Finland. He received the B.Sc. degrees in Telecommunications Engineering from University Mediterranea of Reggio Calabria, Italy, in 2009 and the M.Sc. from University of Padova, Italy, in 2012. He also received his Ph.D. from University Mediterranea of Reggio Calabria, Italy, in 2017. He is actively working in 5G standardization activities and his current research interests include Device-to-Device and Machine-to-Machine communications in 4G/5G cellular systems, and Internet of Things.

Leonardo Militano is an Assistant Professor at the Mediterranea University of Reggio Calabria, Italy. He received a M.Sc. degree in Telecommunications Engineering with full marks in 2006 and his Ph.D. in Telecommunications Engineering in 2010 from the University of Reggio Calabria. His main areas of research are 5G systems, Internet of Things, wireless networks optimization and game theory applications to networking.

Antonio Iera graduated in Computer Engineering at the University of Calabria, Italy, in 1991 and received a Master Diploma in Information Technology from CEFRIEL/Politecnico di Milano, Italy, in 1992 and a Ph.D. degree from the University of Calabria in 1996. Since 1997 he has been with the University of Reggio Calabria and currently holds the position of Full Professor of Telecommunications and Director of the Laboratory for Advanced Research into Telecommunication Systems. He is IEEE Senior Member since 2007. His research interests include, next generation mobile and wireless systems, RFID systems, and Internet of Things.

Giuseppe Araniti is an Assistant Professor of Telecommunications at the University Mediterranea of Reggio Calabria, Italy. From the same university, he received the Laurea (2000) and Ph.D. degree (2004) in Electronic Engineering. His major area of research includes Personal Communications Systems, Enhanced Wireless and Satellite Systems, Traffic and Radio Resource Management, Multicast and Broadcast Services, Device-to-Device and Machine Type Communications over 4G/5G cellular networks.