# Infocommunications Journal

A PUBLICATION OF THE SCIENTIFIC ASSOCIATION FOR INFOCOMMUNICATIONS (HTE)

DECEMBER 2017 Volume IX Number 4 ISSN 2061-2079



Technically Co-Sponsored by

IEEE ComSoc™ IEEE Communications Society

hte

IEEE HUNGARY SECTION

## Indexing information

Infocommunications Journal is covered by Inspec, Compendex and Scopus.
**Infocommunications Journal is also included in the Thomson Reuters – Web of ScienceTM Core Collection, Emerging Sources Citation Index (ESCI)**

## Infocommunications Journal

Technically co-sponsored by IEEE Communications Society and IEEE Hungary Section

**Supporters**

FERENC VÁGUJHELYI – president, National Council for Telecommunications and Information Technology (NHIT)
GÁBOR MAGYAR – president, Scientic Association for Infocommunications (HTE)

**Editorial Office** (Subscription and Advertisements)**:**
Scientic Association for Infocommunications
H-1051 Budapest, Bajcsy-Zsilinszky str. 12, Room: 502
Phone: +36 1 353 1027
E-mail: info@hte.hu • Web: www.hte.hu

**Articles can be sent also to the following address:**
Budapest University of Technology and Economics
Department of Telecommunications and Media Informatics
Tel.: +36 1 463 1102, Fax: +36 1 463 1763
E-mail: vida@tmit.bme.hu

**Subscription rates for foreign subscribers:** 4 issues 10.000 HUF + postage

Publisher: PÉTER NAGY

HU ISSN 2061-2079 • Layout: PLAZMA DS • Printed by: FOM Media

# Retrospective Interference Neutralization for the Two-Cell MIMO Interfering Multiple Access Channel

Bowei Zhang, Xin Hu, Weinong Wu, Jie Huang, Jing Wang and Wenjiang Feng

***Abstract*—In this paper, we study the degrees of freedom (DoF) of the two-cell multiple-input multiple-output (MIMO) interfering multiple access channel (IMAC) with hybrid local channel state information (CSI) at the transmitter (CSIT). Hybrid CSIT means that part of the CSIT is delayed while part is instantaneous. The main contribution of this paper is that we propose a retrospective interference neutralization (RIN) method by using hybrid CSIT and obtain a new DoF result. The approach consists of two steps: 1) side-information learning and 2) retrospective interference transmission. In the first step, two cells of users are scheduled to transmit fresh symbols and each base station (BS) overhears the signals sent by the non-corresponding users. Each BS can learn the linear combination of interference symbols. In the second step, each user sends the previously transmitted symbols by using the beamforming matrix so that each BS can exploit the side-information to neutralize all interference symbols. Through theoretical analysis and numerical simulations, we first verify that the RIN method can obtain the theoretical DoF result. Furthermore, we show that with the help of hybrid CSIT, a higher DoF can be achieved compared with the existing retrospective interference alignment schemes based on completely delayed CSIT.**

***Index Terms*—degrees of freedom, delayed CSIT, hybrid CSIT, interference alignment, interference neutralization.**

## I. INTRODUCTION

Due to the broadcast nature of the wireless medium, simultaneous transmission from multiple nodes causes interference to each other. Interference is a major barrier to limit the capacity of multi-user wireless networks. Recently, interference alignment (IA) was shown to be able to achieve the information-theoretic maximum degrees of freedom (DoF) of some interference networks[1]-[2]. When relays are available in interference channel, it was found in [3-5] that another interference management technique, called interference neutralization (IN), can achieve a higher DoF than IA.

It is worth noting that the above DoF results were obtained under a requirement of instantaneous CSIT. In practical scenarios, it is unreasonable to ignore the impact of feedback delay. When the CSIT is delayed due to practical constraints, the conventional IA and IN is useless. To address this problem,

This work was supported by Chinese National Natural Science Foundation No. 61501066.

Bowei Zhang, Xin Hu, Weinong Wu, Jie Huang and Jing Wang are with State Grid Chongqing Information and Telecommunication Company, China (e-mail:boevyzbw@163.com, cq_hnx@163.com, wuweinong@sina.com. 2352386147@qq.com and 65314336@qq.com)

Wenjiang Feng is with College of Communication Engineering in Chongqing University, China(e-mail: Fengwj@cqu.edu.cn)

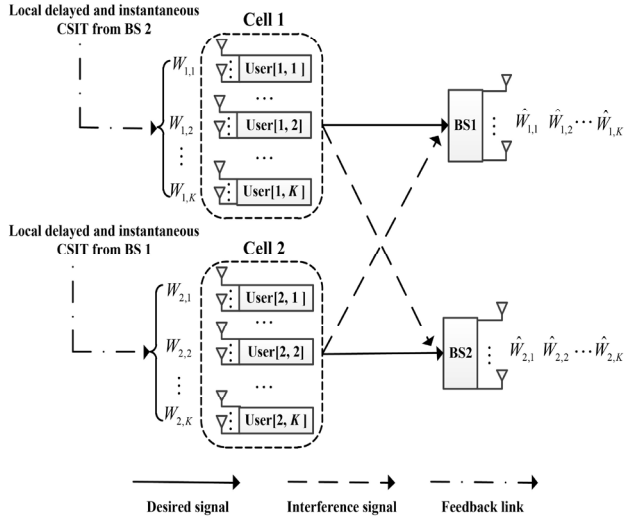



Fig.1 A two-cell MIMO IMAC with hybrid (delayed and local instantaneous) CSIT.

retrospective interference alignment (RIA) was proposed in [6-7] to show that RIA is still capable of increasing the DoF even when the CSIT is completely outdated. Subsequently, in [8], delayed CSIT was exploited to obtain a higher DoF than the time division multiple access (TDMA) scheme for the two-cell MIMO interfering multiple access channel (IMAC).

In this paper, we provide a new DoF result of the two-cell MIMO IMAC under an assumption that hybrid CSIT is available at transmitters. Here hybrid CSIT means that transmitters have a mix of delayed and limited amount of instantaneous CSI. The authors in [9] and [10], have shown that hybrid CSIT can obtain more DoF than the "delayed-only" CSIT for interference and X channels, respectively. By taking advantage of hybrid CSIT, we propose a retrospective interference neutralization (RIN) scheme for the two-cell MIMO IMAC. We show that the achievable sum-DoF is equal to $4M/3$ if $N < M \le KN/2$ and $M(1+1/(2\varphi+1))$ if $KN/2 < M < KN$, where $M$ and $N$ denote the number of antennas equipped at each base station (BS) and each user, respectively, and $K$ denotes the number of users at each cell; $\varphi = \lceil M/(KN-M) \rceil$ represents the number of time slots that each user transmit the fresh symbols. It is shown that the RIN scheme can achieve a higher DoF than the best-known "delayed-only" result in [8] by using fewer time slots.

Notations: For any matrix $\mathbf{A}$, $\mathbf{A}^T$, $\mathbf{A}^{-1}$, $\mathbf{A}^H$, $\mathbf{A}^\dagger$ and *rank* ($\mathbf{A}$) denote the transpose, inverse of $\mathbf{A}$, Hermitian transpose,

Moore–Penrose pseudo-inverse of **A** and the rank of **A**, respectively. $\mathbb{E}[\cdot]$ is the expectation operator. For convenience, we set $\mathcal{K} \triangleq \{1, 2, \ldots, K\}$.

## II. System Model

We consider a MIMO IMAC with two cells and $K$ ($K \geq 2$) users per cell as illustrated in Fig.1, where each BS and each user are equipped with $M$ and $N$ antennas, respectively. To be more realistic, we focus on a case with $M \geq N$, i.e., the number of antennas per BS is greater than or equal to the number of antennas per user. For convenience, we refer to the channel as a two-cell ($M$, $N$, $K$) MIMO IMAC. It is assumed that all nodes in the MIMO IMAC share the same frequency band. Owing to the simultaneous transmission, each BS receives the interference from the users in the other cell. For $k \in \mathcal{K}$ and $i \in \{1,2\}$, we denote the $k$-th user in the $i$-th cell as user [$i$, $k$]. User [$i$, $k$] intends to transmit the message $W^{[i,k]}$ to its corresponding BS.

When all users simultaneously send their signals at time slot $t$, the received signal $\mathbf{y}_l[t]$ at the $l$-th ($l \in \{1,2\}$) BS is given by

$$\mathbf{y}_l[t] = \sum_{i=1}^{2}\sum_{k=1}^{K} \mathbf{H}_l^{[i,k]}[t]\mathbf{x}^{[i,k]}[t] + \mathbf{n}_l[t] \tag{1}$$

where $\mathbf{H}_l^{[i,k]}[t]$ denotes the $M \times N$ channel matrix from user [$i$, $k$] to the $l$-th BS at time slot $t$; $\mathbf{x}^{[i,k]}[t]$ denotes the $N \times 1$ signal vector sent by the user [$i$, $k$] at time slot $t$ satisfying an average power constraint $\mathbb{E}\left[\left\|\mathbf{x}^{[i,k]}[t]\right\|^2\right] \leq P$. $\mathbf{n}_l[t]$ denotes the $M \times 1$ additive white Gaussian noise (AWGN) vector with zero mean and variance $\sigma^2$ per entry. We assume that the channels vary independently at different time slot.

Similar to [6], we normalize delay to one time slot because our achievable results hold for arbitrary delay. Specifically, at time slot $t$, user [$i$, $k$] has access to the delayed local CSIT $\left\{\mathbf{H}_j^{[i,k]}[t']\right\}_{t'=1}^{t-1}$. At a particular time slot $\bar{t} = \left\{\mathbf{H}_j^{[i,k]}[t']\right\}_{t'=1}^{\bar{t}}$, the instantaneous CSIT is available at each user, where $i, j \in \{1,2\}$, $k \in \mathcal{K}$, $i \neq j$. In addition, each BS only needs to know local instantaneous CSIR, which is more relaxed than that demanded in [8]. It is assumed that all CSI can be obtained through a noiseless feedback link.

The sum-DoF of the network is defined as the pre-log coefficient of network sum rate in high signal-to-noise ratio (SNR) regime. In each time slot, the achievable rate from user [$l$, $k$] to the $l$-th BS can be expressed as:

$$R^{[l,k]}(SNR) = d^{[l,k]}\log(SNR) + o\left(\log(SNR)\right) \tag{2}$$

Where $d^{[l,k]}$ is the achievable DoF of user [$l$, $k$], $l \in \{1,2\}$, $k \in \mathcal{K}$. Therefore, the sum rate of the $l$-th BS can be written as

$$R_l(SNR) = \sum_{k=1}^{K} R^{[l,k]}(SNR) \tag{3}$$

The sum-DoF of the two-cell MIMO IMAC can be given by

$$\mathrm{DoF}_{sum} = \sum_{l=1}^{2}\sum_{k=1}^{K} d^{[l,k]} = \sum_{l=1}^{2}\sum_{k=1}^{K} \lim_{SNR\to\infty} \frac{R^{[l,k]}(SNR)}{\log(SNR)} \tag{4}$$

Since noise does not affect DoF, we neglect noise terms in the following discussions.

## III. Proposed Method Based on Hybrid CSIT

In this section, we first use a simple example to illustrate the core idea behind our method. Then, we extend our approach into a general MIMO IMAC.

### *A. Example of (M, N, K)=(2, 2, 2)*

Throughout this example, we will show that the sum-DoF of 4 can be achieved for (2, 2, 2) MIMO-IMAC using hybrid CSIT. The RIN scheme involves two steps: side-information learning and retrospective interference transmission. Next we explain how each of these steps is performed to achieve the stated sum-DoF.

**1) Side-information learning:** This step consists two time slots. In the first time slot, user [1,1] and user [1,2] send signals $\mathbf{x}^{[1,1]}[1] = [a_1\ a_2]^T$ and $\mathbf{x}^{[1,2]}[1] = [b_1\ b_2]^T$ to their corresponding BS (i.e., BS 1), while users in cell 2 keep silence. The received signal at each BS is given by:

$$\begin{aligned}\mathbf{y}_l[1] &= \mathbf{H}_l^{[1,1]}[1]\mathbf{x}^{[1,1]}[1] + \mathbf{H}_l^{[1,2]}[1]\mathbf{x}^{[1,2]}[1] \\ &= \mathbf{L}_l[1](a_1, a_2, b_1, b_2)\end{aligned} \tag{5}$$

where $\mathbf{L}_l[t](\cdot)$ denotes the linear combination of symbols that is received by the $l$-th BS at time slot $t$. In the second time slot, user [2,1] and user [2,2] send signals $\mathbf{x}^{[2,1]}[1] = [d_1\ d_2]^T$ and $\mathbf{x}^{[2,2]}[1] = [e_1\ e_2]^T$, while users in cell 1 keep silence. The received signal at each BS is given by:

$$\begin{aligned}\mathbf{y}_l[2] &= \mathbf{H}_l^{[2,1]}[2]\mathbf{x}^{[2,1]}[2] + \mathbf{H}_l^{[2,2]}[2]\mathbf{x}^{[2,2]}[2] \\ &= \mathbf{L}_l[2](d_1, d_2, e_1, e_2)\end{aligned} \tag{6}$$

At each time slot of the side-information step, each BS obtains two linear independent equations (each BS has 2 antennas) with four variables and we need to provide two more linear independent equations to resolve the desired symbols. Note that if BS 1 and BS2 knows the linear combinations of $\mathbf{L}_2[1](a_1, a_2, b_1, b_2)$ and $\mathbf{L}_1[2](d_1, d_2, e_1, e_2)$, respectively, they will have enough equations to decode all desired symbols. The key idea of step 2 is to provide the interference-free linear independent equations $\mathbf{L}_2[1](a_1, a_2, b_1, b_2)$ to BS 1 and $\mathbf{L}_1[2](d_1, d_2, e_1, e_2)$ to BS 2.

**2) Retrospective interference transmission:** We use the third time slot for the retrospective interference transmission. In the third time slot, each user has knowledge of both current and outdated local CSIT thanks to feedback. User [1, $k$] and user [2, $k$] ($k \in \{1,2\}$) simultaneously send a linear combination of the previously transmitted symbols by using linear beamforming matrix as

$$\mathbf{x}^{[1,k]}[3] = \mathbf{V}^{[1,k]}[3]\mathbf{H}_2^{[1,k]}[1]\mathbf{x}^{[1,k]}[1],\ k \in \{1,2\} \tag{7}$$

$$\mathbf{x}^{[2,k]}[3] = \mathbf{V}^{[2,k]}[3]\mathbf{H}_1^{[2,k]}[2]\mathbf{x}^{[2,k]}[2],\ k \in \{1,2\} \tag{8}$$

Then, the received signal at each BS is given by:

$$\begin{aligned}\mathbf{y}_l[3] = &\sum_{k=1}^{2} \mathbf{H}_l^{[1,k]}[3]\mathbf{V}^{[1,k]}[3]\mathbf{H}_2^{[1,k]}[1]\mathbf{x}^{[1,k]}[1] \\ &+ \sum_{k=1}^{2} \mathbf{H}_l^{[2,k]}[3]\mathbf{V}^{[2,k]}[3]\mathbf{H}_1^{[2,k]}[2]\mathbf{x}^{[2,k]}[2]\end{aligned} \tag{9}$$

The key idea of the interference neutralization step is to exploit the hybrid CSIT at each user and the side-information at each BS so that the received interference symbols at time slot 3 and side-information of the previous time slots can be added up to

zero at each BS. It is noteworthy that the main difference between the traditional interference neutralization and the retrospective interference neutralization is that the former can eliminate interference in the air with the help of the relay that has global CSI, whereas the latter utilizes the received interference and does not need the aid of relay. For user [2, $k$] ($k \in \{1,2\}$), we construct the $2\times 2$ interference neutralization matrix $\mathbf{V}^{[2,k]}[3]$ so that the received interference symbols at time slot 3 can be neutralized by using the side-information at BS 1, i.e.,

$$\sum_{k=1}^{2}\mathbf{H}_1^{[2,k]}[3]\mathbf{V}^{[2,k]}[3]\mathbf{H}_1^{[2,k]}[2]\mathbf{x}^{[2,k]}[2] + \mathbf{L}_1[2](d_1,d_2,e_1,e_2)=\mathbf{0} \tag{10}$$

In order to satisfy (10), the following condition should be satisfied:

$$\mathbf{H}_1^{[2,k]}[3]\mathbf{V}^{[2,k]}[3]\mathbf{H}_1^{[2,k]}[2]+\mathbf{H}_1^{[2,k]}[2]=\mathbf{0} \tag{11}$$

Then, we design $\mathbf{V}^{[2,k]}[3]$ as:

$$\mathbf{V}^{[2,k]}[3]=-\left(\mathbf{H}_1^{[2,k]}[3]\right)^{-1},\ k\in\{1,2\} \tag{12}$$

Since $\mathbf{H}_1^{[2,k]}[t]$ is a $2\times 2$ channel matrix, $\left(\mathbf{H}_1^{[2,k]}[3]\right)^{-1}$ and $\mathbf{V}^{[2,k]}[3]$ exist and (10) holds.

For user[1, $k$], we construct the interference neutralization matrix $\mathbf{V}^{[1,k]}[3]$ to eliminate the interference symbols at BS 2 as follows:

$$\mathbf{H}_2^{[1,k]}[3]\mathbf{V}^{[1,k]}[3]\mathbf{H}_2^{[1,k]}[1]+\mathbf{H}_2^{[1,k]}[1]=\mathbf{0},\ k\in\{1,2\} \tag{13}$$

Similar to (10), (13) holds. As a result, BS 1 and BS 2 obtain interference-free linear independent combinations of $\mathbf{L}_2[1](a_1,a_2,b_1,b_2)$ and $\mathbf{L}_1[2](d_1,d_2,e_1,e_2)$. Without the retrospective interference transmission, step 2 will take two time slots to provide three interference-free linear independent equations for each BS, which yields DoF loss.

**3) Decoding:** We explain the decoding procedure for BS 1. Concatenating the received signals in three time slots, the equivalent input-output relationship at BS 1 is

$$\underbrace{\begin{bmatrix}\mathbf{y}_1[1]\\ \mathbf{y}_1[3]+\mathbf{y}_1[2]\end{bmatrix}}_{\bar{\mathbf{y}}_1}=\underbrace{\begin{pmatrix}\mathbf{H}_1^{[1,1]}[1] & \mathbf{H}_1^{[1,2]}[1]\\ \bar{\mathbf{H}}_2^{[1,1]}[1] & \bar{\mathbf{H}}_2^{[1,2]}[1]\end{pmatrix}}_{\bar{\mathbf{H}}_1}\begin{pmatrix}\mathbf{x}^{[1,1]}[1]\\ \mathbf{x}^{[1,2]}[1]\end{pmatrix} \tag{14}$$

where $\bar{\mathbf{H}}_2^{[1,k]}[1]=\mathbf{H}_1^{[1,k]}[3]\mathbf{V}^{[1,k]}[3]\mathbf{H}_2^{[1,k]}[1]$, $k\in\{1,2\}$; $\bar{\mathbf{y}}_1$ is a $4\times 1$ vector, $\bar{\mathbf{H}}_1$ is a $4\times 4$ matrix. Since all channel values are generic, $rank(\bar{\mathbf{H}}_1)=4$. Therefore, BS 1 can decode 4 desired symbols. Similarly, BS 2 can decode 4 desired symbols in three time slots by using the same method. Consequently, a total of 8 desired symbols have been transmitted in three time slots in the network, implying that the sum-DoF of 8/3 is achieved.

**4) Sum-DoF gains:** For the (2,2,2) MIMO IMAC, the RIA scheme in [8] can achieve the sum-DoF of 12/5 in 5 time slots. Since the RIN scheme attains the sum-DoF of 8 in 3 time slots, we utilizes fewer time slots to obtain 11.1% DoF gains over the RIA scheme in this network.

*B. The general configuration*

We now present the achievable DoF results for the general ($M$, $N$, $K$) MIMO IMAC in Theorem 1.

**Theorem 1:** For the two-cell ($M$, $N$, $K$) MIMO IMAC with hybrid CSIT where $N<M<KN$, the achievable sum-DoF is

$$\mathrm{DoF}_{\mathrm{sum}}=\begin{cases}\dfrac{4M}{3}, & N<M\le KN/2\\ M\left(1+\dfrac{1}{2\varphi+1}\right), & KN/2<M<KN\end{cases} \tag{15}$$

where $\varphi=\lceil M/(KN-M)\rceil$ represents the number of time slots used by each user to transmit the fresh symbols.

**Proof:** We prove Theorem 1 by using the proposed RIN scheme. In order to reach the maximal sum-DoF by using the RIN scheme, we should design different transmission strategies according to different system configurations. We first consider the case of $KN/2<M<KN$. In the following, we describe each transmissions in detail.

We first consider the side-information learning step. This step consists of two phases and each phase contains $\varphi$ time slots. The number of $\varphi$ time slots is determined on the values $M$, $N$, $K$. The $i$-th ($i\in\{1,2\}$) phase is dedicated to users in the $i$-th cell and user [$i$, $k$] ($k\in\mathcal{K}$) simultaneously sends $\mathbf{x}^{[i,k]}[t]=\left[x_1^{[i,k]},x_2^{[i,k]},\cdots,x_N^{[i,k]}\right]^T[t]$ over its own $N$ transmit antennas at time slot $t$. Note that the final transmission (phase 3) only contains one time slot, we only can provide $M$ new linear independent equations in phase 3. Due to the condition of $KN/2<M<KN$, we can find $\varphi$ so that the condition of $\varphi M+M\le KN\varphi$ can be satisfied. Since each BS can decode at most $\varphi M+M$ desired symbols in $2\varphi+1$ time slots, users in the $i$-th cell only transmit $\varphi M+M$ fresh symbols in the $i$-th phase.

During phase 1, all users in cell 1 transmit $\varphi M+M$ fresh symbols and $KN\varphi-\varphi M-M$ zero elements in total and users in cell 2 keep silence. The received signal at each BS is given by:

$$\begin{aligned}\mathbf{y}_l[t]&=\mathbf{H}_l^{[1,1]}[t]\mathbf{x}^{[1,1]}[t]+\cdots+\mathbf{H}_l^{[1,K]}[t]\mathbf{x}^{[1,K]}[t]\\ &\triangleq\mathbf{L}_l[t](\mathbf{x}^{[1,1]},\cdots,\mathbf{x}^{[1,K]})\end{aligned} \tag{16}$$

where $t\in\{1,2,\ldots,\varphi\}\triangleq\mathcal{T}_1$, $\mathbf{L}_l[t](\cdot)$ denotes the linear combination of symbols that is received by the $l$-th BS at time slot $t$. In phase 1, BS 1 receives the desired symbols and BS 2 overhears the interference symbols for use later (in phase 3). For each time slot, BS 1 has $M$ linearly independent equations with $KN$ variables. Due to the condition of $KN>M$, BS 1 cannot decode the desired symbols. In phase 1, if BS 1 knows $\mathbf{L}_2[1](\mathbf{x}^{[1,1]},\cdots,\mathbf{x}^{[1,K]})$ overheard by BS 2, it can decode all desired symbols.

During phase 2, all users in cell 2 transmit $\varphi M+M$ fresh symbols and the received signal at each BS is given by:

$$\begin{aligned}\mathbf{y}_l[2]&=\mathbf{H}_l^{[2,1]}[2]\mathbf{x}^{[2,1]}[2]+\cdots+\mathbf{H}_l^{[2,K]}[2]\mathbf{x}^{[2,K]}[2]\\ &\triangleq\mathbf{L}_l[2](\mathbf{x}^{[2,1]},\cdots,\mathbf{x}^{[2,K]})\end{aligned} \tag{17}$$

where $t\in\{\varphi+1,\varphi+2,\cdots,2\varphi\}\triangleq\mathcal{T}_2$. Similarly, if BS 2 knows $\mathbf{L}_1[t](\mathbf{x}^{[2,1]},\cdots,\mathbf{x}^{[2,K]})$ ($t\in\mathcal{T}_2$), it can decode all desired symbols.

For the retrospective interference transmission, we use one time slot. Without loss of generality, we assume time slot $2\varphi+1$ has instantaneous CSIT. For simplicity, let us denote

$\bar{\varphi} = 2\varphi + 1$ in subsequent equation expression. In phase 3, each user has knowledge of both current and outdated local CSIT. User [1, $k$] and user [2, $k$] ( $k \in \mathcal{K}$ ) simultaneously send a superposition of the linear combination of previously transmitted symbols by using linear beamforming matrix $\mathbf{V}^{[1,k]}[3]$ and $\mathbf{V}^{[2,k]}[3]$, respectively. Then, the received signal at each BS is given by:

$$\begin{aligned}\mathbf{y}_l[\bar{\varphi}] = &\sum_{k=1}^{K}\mathbf{H}_l^{[1,k]}[\bar{\varphi}]\mathbf{V}^{[1,k]}[\bar{\varphi}]\sum_{t=1}^{\varphi}\mathbf{H}_2^{[1,k]}[t]\mathbf{x}^{[1,k]}[t] \\ &+\sum_{k=1}^{K}\mathbf{H}_l^{[2,k]}[\bar{\varphi}]\mathbf{V}^{[2,k]}[\bar{\varphi}]\sum_{t=\varphi+1}^{2\varphi}\mathbf{H}_1^{[2,k]}[t]\mathbf{x}^{[2,k]}[t]\end{aligned} \tag{18}$$

For user [$i$, $k$] ( $i \in \{1,2\}$ , $k \in \mathcal{K}$ ), we construct the $N \times M$ interference neutralization matrix $\mathbf{V}^{[i,k]}[\bar{\varphi}]$ so that the received interference symbols at time slot $2\varphi+1$ can be neutralized by using the side-information at BS 1 and BS2, respectively, i.e.,

$$\begin{aligned}&\sum_{k=1}^{K}\mathbf{H}_1^{[2,k]}[\bar{\varphi}]\mathbf{V}^{[2,k]}[\bar{\varphi}]\sum_{t=\varphi+1}^{2\varphi}\mathbf{H}_1^{[2,k]}[t]\mathbf{x}^{[2,k]}[t] \\ &\quad+\sum_{t=\varphi+1}^{2\varphi}\mathbf{L}_1[t](\mathbf{x}^{[2,1]},\cdots,\mathbf{x}^{[2,K]}) = \mathbf{0}\end{aligned} \tag{19}$$

$$\begin{aligned}&\sum_{k=1}^{K}\mathbf{H}_2^{[1,k]}[\bar{\varphi}]\mathbf{V}^{[1,k]}[\bar{\varphi}]\sum_{t=1}^{\varphi}\mathbf{H}_2^{[1,k]}[t]\mathbf{x}^{[1,k]}[t] \\ &\quad+\sum_{t=1}^{\varphi}\mathbf{L}_2[t](\mathbf{x}^{[1,1]},\cdots,\mathbf{x}^{[1,K]}) = \mathbf{0}\end{aligned} \tag{20}$$

In order to satisfy (19) and (20), the following conditions should be satisfied:

$$\mathbf{H}_1^{[2,k]}[\bar{\varphi}]\mathbf{V}^{[2,k]}[\bar{\varphi}]\mathbf{H}_1^{[2,k]}[t] + \mathbf{H}_1^{[2,k]}[t] = \mathbf{0},\ k \in \mathcal{K}, t \in \mathcal{T}_2 \tag{21}$$

$$\mathbf{H}_2^{[1,k]}[\bar{\varphi}]\mathbf{V}^{[1,k]}[\bar{\varphi}]\mathbf{H}_2^{[1,k]}[t] + \mathbf{H}_2^{[1,k]}[t] = \mathbf{0}\ k \in \mathcal{K}, t \in \mathcal{T}_1 \tag{22}$$

Since $\mathbf{H}_l^{[i,k]}[t]$ is a $M \times N$ channel matrix and $M > N$, $\mathbf{H}_l^{[i,k]}[t]^{\dagger}$ exists, where $\mathbf{H}_l^{[i,k]}[t]^{\dagger} = \left(\mathbf{H}_l^{[i,k]}[t]^H \mathbf{H}_l^{[i,k]}[t]\right)^{-1}\mathbf{H}_l^{[i,k]}[t]^H$. We can design $\mathbf{V}^{[2,k]}[\bar{\varphi}] = -\mathbf{H}_1^{[2,k]}[\bar{\varphi}]^{\dagger}$ and $\mathbf{V}^{[1,k]}[\bar{\varphi}] = -\mathbf{H}_2^{[1,k]}[\bar{\varphi}]^{\dagger}$ so that (21) and (22) can be satisfied, As a result, BS 1 and BS 2 eliminate all interference symbols.

Next we show the decoding procedure for BS 1. Concatenating the received desired signals in $2\varphi+1$ time slots, the equivalent input-output relationship at BS 1 is shown in (23).

$$\underbrace{\begin{bmatrix}\mathbf{y}_1[1]\\ \mathbf{y}_1[2]\\ \vdots\\ \mathbf{y}_1[\varphi]\\ \mathbf{y}_1[\bar{\varphi}]+\sum_{t\in\mathcal{T}_2}\mathbf{y}_1[t]\end{bmatrix}}_{\tilde{\mathbf{y}}_1} = \underbrace{\begin{pmatrix}\mathbf{A}[1] & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{A}[2] & \mathbf{0} & \vdots\\ \vdots & \vdots & \ddots & \mathbf{0}\\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{A}[\varphi]\\ \mathbf{B}[1] & \mathbf{B}[2] & \cdots & \mathbf{B}[\varphi]\end{pmatrix}}_{\tilde{\mathbf{H}}_1}\bar{\mathbf{x}}_1 \tag{23}$$

where $\bar{\boldsymbol{H}}_2^{[1,k]}[t] = \boldsymbol{H}_1^{[1,k]}[\bar{\varphi}]\boldsymbol{V}^{[1,k]}[\bar{\varphi}]\boldsymbol{H}_2^{[1,k]}[t]$ , $\boldsymbol{A}[t] = \left[\boldsymbol{H}_1^{[1,1]}[t]\cdots\boldsymbol{H}_1^{[1,K]}[t]\right]$ , $\boldsymbol{B}[t] = \left[\bar{\boldsymbol{H}}_2^{[1,1]}[t]\cdots\bar{\boldsymbol{H}}_2^{[1,K]}[t]\right]$ , $\bar{\boldsymbol{x}}_1 = \left[\boldsymbol{x}^{[1,1]}[1]^T\cdots\boldsymbol{x}^{[1,K]}[\varphi]^T\right]^T$ , $k \in \mathcal{K}, t \in \mathcal{T}_1$ . $\bar{\mathbf{y}}_1$ is a $(\varphi+1)M \times 1$ vector, $\bar{\mathbf{H}}$ is a $(\varphi+1)M \times (\varphi+1)M$ matrix, $\bar{\boldsymbol{x}}$ is a $(\varphi+1)M \times 1$ vector. Since all channel values are generic, $rank(\bar{\mathbf{H}}) = (\varphi+1)M$ . Therefore, BS 1 can decode $(\varphi+1)M$ desired symbols. Similarly, BS 2 can decode $(\varphi+1)M$ desired symbols in $2\varphi+1$ time slots by using the same method. Consequently, a total of $2(\varphi+1)M$ desired symbols have been transmitted in $2\varphi+1$ time slots in the network, implying that the sum-DoF of $M(1+1/2\varphi+1)$ is achieved. Since the sum-DoF decreases as the parameter $\varphi$ increases and the condition $\varphi M + M \le KN\varphi$ should be satisfied, we select the minimum values of $\varphi$ , i.e., $\varphi = \lceil M/(KN-M) \rceil$.

For the case of $N < M \le KN/2$ , we can select $\varphi$ =1 to satisfy the condition of $\varphi M + M \le KN\varphi$, i.e., each phase only contains one time slot. The sum-DoF of $4M/3$ can be directly obtained by using the same method with the case of $KN/2 < M < KN$. We thus complete the proof of Theorem 1.

## IV. DoF Comparison and Simulation Results

### A. DoF Comparison

We first compares the DoF result obtained when there is only delayed CSIT and the one obtained when there is hybrid CSIT. The case with only delayed CSIT uses the RIA scheme in [8] and the achievable sum-DoF is $M\left(1+1/(2\eta+1)\right)$ when $K \le M \le KN - N/2$ , where $\eta = \lceil 2M/N \rceil$ . According to the results in the *Theorem 1*, we can compute the sum-DoF gain from the hybrid CSIT over the delayed CSIT, which is

$$\mathrm{DoF}_{gain} = \begin{cases} M\left(\dfrac{1}{3} - \dfrac{1}{2\eta+1}\right), & N < M \le KN/2 \\ M\left(\dfrac{1}{2\varphi+1} - \dfrac{1}{2\eta+1}\right), & KN/2 \le M \le KN - N/2 \end{cases} \tag{24}$$

where $\varphi = \lceil M/(KN-M) \rceil$. Since $\eta$ is larger than or equal to 2 in the case of $N < M \le KN/2$ and $\varphi$ is less than $\eta$ in the case of $KN/2 \le M < KN - N/2$, the RIN scheme can utilize fewer time slots to attain a higher DoF than the RIA scheme under the feasible condition of RIA scheme.

### B. Numerical simulation

We provide numerical results to evaluate the sum rate performance of the proposed RIN scheme. We demonstrate that the RIN scheme can exactly obtain the achievable DoF derived in Theorem 1. The channel is modeled as i.i.d. (independent identically distributed) complex Gaussian distribution with zero mean and unit variance. The numerical results are averaged over 10000 independent Monte Carlo runs.

We first explain how to compute the sum rate of the two-cell MIMO IMAC by applying the proposed RIN scheme. We calculate $R_1$ as a representative. According to (23), the received signal at BS 1 can be described as

$$\bar{\mathbf{y}}_1 = \bar{\mathbf{H}}\bar{\mathbf{x}} + \bar{\mathbf{n}}_1 \tag{25}$$

where $\bar{\mathbf{n}}_1 = \left[\mathbf{n}_1[1]\,\mathbf{n}_1[2]\cdots\mathbf{n}_1[\varphi]\right]^T$ . According to [11-12], we can calculate the sum rate $R_1$ by (26)

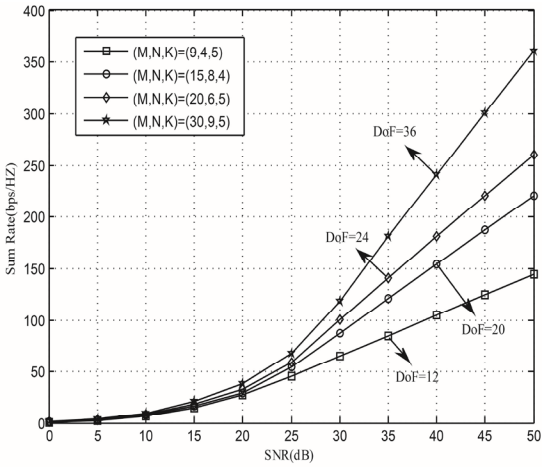



Fig.2 The sum rate and DoF of the proposed RIN scheme.

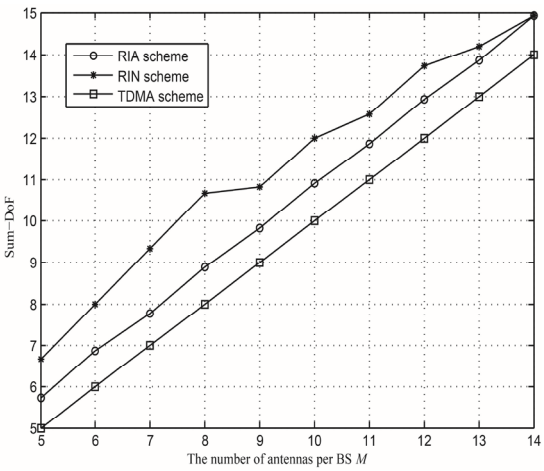



Fig.3 Comparison of sum-DoF among the proposed RIN scheme, the SNC scheme and the TDMA scheme

$$R_1 = \log_2\left[\det\left(\mathbf{I}+\overline{\mathbf{H}}\mathbb{E}\left(\overline{\mathbf{x}}\left(\overline{\mathbf{x}}\right)^H\right)\left(\overline{\mathbf{H}}\right)^H \mathbb{E}\left(\overline{\mathbf{n}}_1\left(\overline{\mathbf{n}}_1\right)^H\right)^{-1}\right)\right] \quad (26)$$

Fig.2 illustrates the sum rate and DoF results of the proposed RIN scheme with different system configurations. For the case of $N < M \le KN/2$, the curves of (9,4,5) and (15,8,4) are plotted. For the case of $KN/2 < M < KN$, the curves of (20,6,5) and (30,9,5) are plotted. The slopes of the sum rate curves at high SNR are equal to DoF/lg2 according to the Shannon capacity formula. Checking the case of $N < M \le KN/2$, we can observe that (9,4,5) and (15,8,4) obtain 12 and 20 DoF, respectively. The DoF results match with the theoretical DoF of $4M/3$ well at the high SNR regime. Similarly, For the curves of (20,6,5) and (30,9,5), the DoF results match with the theoretical DoF of $M\left(1+1/(2\varphi+1)\right)$ ($\varphi=\lceil M/(KN-M)\rceil=2$), which verifies the theoretic results shown in Theorem 1.

To show the superiority of the proposed RIN scheme, we plot the achievable DoF achieved by different schemes in Fig.3. The system configuration ($M$, $N$, $K$) is chosen based on Theorem 1, where $M \in [5,14]$, $N$=4 and $K$=4. We can observe that the RIN scheme attains a higher DoF than the RIA scheme and the TDMA scheme. The DoF gain comes from the fact that exploiting hybrid CSIT can transmit more interference-free symbols. Additionally, the RIN scheme can utilize the signal space at each BS more efficiently than the RIA scheme do so that the RIN scheme can eliminate more interference symbols than the RIA scheme. For the different number of $N$ and the same number of $M$, we can obtain the same result. It is worth note that the DoF gain will be reduced in the presence of channel estimation errors [13]. However, the sum DoF of the RIN scheme is still higher than the TDMA scheme when channel estimation errors exist.

## V. CONCLUSION

We proposed a RIN method to obtain the new achievable DoF of the two-cell MIMO IMAC with delayed and limited amount of instantaneous CSIT (hybrid CSIT). Through theoretical analysis and numerical simulations, we showed that the proposed RIN method can obtain the theoretical DoF results. Furthermore, it was shown that the proposed RIN method can provide significant performance gain over the conventional TDMA scheme and the RIA scheme in terms of DoF, which highlights the benefits brought by the hybrid CSIT for uplink cellular networks.

## REFERENCES

[1] V. R. Cadambe, and S. A. Jafar, "Interference alignment and degrees of freedom of the $K$ -user interference channel," *IEEE Trans. Inf. Theory,* vol. 54, no. 8, pp.3425-3441, 2008.

[2] T. Gou, and S. A. Jafar, "Degrees of freedom of the $K$-user $M \times N$ MIMO interference channel," *IEEE Trans. Inf. Theory,* vol. 56, no. 12, pp.6040-6057, 2010.

[3] T. Gou, S. A. Jafar, C. Wang, S.-W. Jeon, and S.-Y. Chung, "Aligned interference neutralization and the degrees of freedom of the 2×2×2 interference channel," *IEEE Trans. Inf. Theory,* vol. 58, no. 7, pp.4381-4395, 2012.

[4] N. Lee, and C. Wang, "Aligned interference neutralization and the degrees of freedom of the two-user wireless networks with an instantaneous relay," *IEEE Trans. Wireless Commun,* vol. 61, no. 9, pp.3611-3619, 2013.

[5] D. Wu, C. Yang, T. Liu, and Z. Xiong, "Feasibility conditions for interference neutralization in relay-aided interference channel," *IEEE Trans. Signal Process,* vol. 62, no. 6, pp.1408-1423, 2014.

[6] M. A. Maddah-Ali, and D. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Trans. Inf. Theory,* vol. 58, no. 7, pp.4418-4431, 2012.

[7] H. Maleki, S. A. Jafar, and S. Shamai, "Retrospective interference alignment over interference networks," *IEEE J. Selected Topics Signal Process,* vol. 6, no. 6, pp.228-240, 2012.

[8] W. Shin, and J. Lee, "Retrospective interference alignment for the two-cell MIMO interfering multiple cccess channel," *IEEE Trans. Wireless Commun,* vol. 14, no. 7, pp.3937-3947, 2015.

[9] K. Mohanty, C. S. Vaze, and M. K. Varanasi, "The degrees of freedom region of the MIMO interference channel with hybrid CSIT," *IEEE Trans. Wireless Commun,* vol. 14, no. 4, pp.1837-1848, 2015.

[10] K. Anand, E. Gunawan, and Y. L. Guan, "On degrees of freedom of three-user SISO-X networks under hybrid CSI at the transmitters," *IEEE Commun. Lett,* vol. 20, no. 5, pp.1-1, 2016.

[11] Telatar E, "Capacity of Multi-antenna Gaussian Channels," *Transactions on Emerging Telecommunications Technologies*, vol.10, no. 6, pp.585-595, 1999.

[12] Win M Z, Pinto P C, Shepp L A, "A Mathematical Theory of Network Interference and Its Applications," *IEEE Proceedings*, vol.92, no. 2, pp.205-230, 2009.

[13] Conti A, Panchenko D, Sidenko S, et al, "Log-Concavity Property of the Error Probability With Application to Local Bounds for Wireless Communications," *IEEE Trans. Inf. Theory,* vol. 55, no. 6, pp. 2766-2775, 2009.

**Bowei Zhang** received the M.S. and Ph.D. degree in the Department of communication engineering, Chongqing University in 2013 and 2017, respectively. He is currently a senior engineer in the State Grid Chongqing Information and Telecommunication Company. His research interests are in the design and analysis of future wireless communications such as cellular networks and interference-limited networks.

**Xin Hu** received the M.S.degree in the Department of communication engineering, Chongqing University in 2007. He is currently a senior engineer in the State Grid Chongqing Information and Telecommunication Company. His research interests include power communication and interference management.

**Weinong Wu** received the B.S.degree in the Department of communication engineering, Chongqing University in 2001. He is currently a senior engineer in the State Grid Chongqing Information and Telecommunication Company. His research interests include power communication and interference management.

**Jie Huang** received the B.S. and Ph.D. degree in the Department of communication engineering, Chongqing University in 2010 and 2017, respectively. He is currently a senior engineer in the State Grid Chongqing Information and Telecommunication Company. His research focuses on full-duplex communication and space-timing coding.

**Jing Wang** received the B.S. degree in the Department of communication engineering, Hainan University in 2017. She is currently an engineer in the State Grid Chongqing Information and Telecommunication Company. Her research focuses on signal processing and network coding.

**Wenjiang Feng** received his Ph.D. degree in electrical engineering from Chongqing University in 2000. Currently, he is a professor at the college of communication engineering in Chongqing University. His research interests fall into the broad areas of communication theory and wireless communication.

# On the Performance of Spectrum Handoff Framework for Next-generation 5G Networks

Pavel Masek, Zhaleh Sadreddini, Tugrul Cavdar and Jiri Hosek

***Abstract*—The lack of available radio spectrum and inefficiency in its usage necessitate a new communication paradigm requiring to exploit the existing spectrum opportunistically. One of the perspective spectrum sharing methods, which is currently under a heavy investigation by academia and industry as well across whole Europe, is called Licensed Shared Access (LSA). This novel technology allows for controlled sharing of spectrum between an original owner (primary user, incumbent) and a licensee (secondary user), such as the mobile network operators (MNOs), which coexist geographically. Despite certain benefits, there are still several issues to be solved before the LSA framework will be implemented in commercial infrastructure. One of them is the need to move secondary users (SUs) from the rented LSA band whenever the incumbent needs it. The potential solution for this problem is represented by spectrum handoff, which aims to help SUs to vacate the occupied licensed spectrum and find suitable network resources to resume the unfinished transmissions somewhere else in order to keep SUs satisfaction in terms of quality of experience (QoE) at negotiated quality of service (QoS) level. Inspired by this, we propose a decision making model considering several SUs attributes (RSSI, RSRP, RSRQ, SINR) in order to efficiently implement the handoff procedure and treat SUs to maximize total service time, spectrum utilization and SUs' satisfaction. As an input for our simulation model, we have used the set of measurements performed in real 3GPP LTE-A indoor cellular system located at Brno University of Technology, Czech republic. Our achieved simulation results evaluate the spectrum utilization of three 3GPP LTE-A cells and provide the total service time for each active SU, while different values of primary user's activity ratio are considered for each cell. Authors would like to recall that this paper represents extended version of their previously published work at TSP 2017 conference.**

***Index Terms*—Cellular communications, Decision making model, LSA framework, Spectrum handoff.**

## I. INTRODUCTION

Over the last decades, radio spectrum has transformed into a critical resource from the economic, cultural and societal points of view. Its scarcity has proven to be a major issue across particular frequency ranges, spanning 100 MHz to 6 GHz, with desired propagation characteristics for a wide range of non-mobile spectrum users (e.g., military, radar, TV broadcasting, medical and event production, etc) [2]. With respect to spectrum management, several concepts have been developed and deployed over the years [3], [4], [5].

With more dynamic spectrum management, the expensive frequency bands could be shared between different stakeholders flexibly, as opposed to exclusive use of licensed spectrum. This may go far beyond opening up unlicensed frequencies for collective uncontrolled use and promises to unlock the much needed additional bandwidth that is currently employed sparsely by its existing incumbents. It can also improve the utilization of presently allocated spectrum across its various dimensions (space, time, frequency), which is essential to support the throughput-hungry 5G applications. To this effect, powerful spectrum sharing technologies emerged recently, such as LTE in unlicensed spectrum (LTE-U), License Assisted Access (LAA), MulteFire, Citizens Broadband Radio Service (CBRS), and Licensed Shared Access (LSA). In [6], the joint optimization of user-to-base station (BS) association together with the time-frequency resource block (RB) and power allocation in heterogeneous networks (HetNets) are discussed. The proposed algorithm in the time-sharing and the non-time-sharing use-cases allows the RBs to be reused opportunistically. One of the most important functionality of the spectrum management is spectrum handoff (so-called spectrum mobility) [7], [8]. The spectrum handoff is implemented to make a facility for SUs in terms of reducing the denial-of-service (DoS).

Spectrum handoff is needed, when the primary users (PUs) request access to their own licensed band, which has been rented and occupied by the active secondary users (ASUs). In this case, the secondary users (SUs) are forced to vacate the spectrum. In such situation, to avoid a DoS state of SUs, the spectrum handoff procedures can help SUs to find idle frequency channels to resume their unfinished data transmission. Selected strategies of the handoff using a channel hopping technique for single-interface devices are mentioned in [9], [10]. At the same time, the protocol design needs to deal with new concerns that result from the use of multiple channels.

Going further, in [11], a variety of techniques for both single- and multiple-interface devices is considered. Further, the spectrum handoff mechanisms can be categorized as either the proactive-decision or reactive-decision spectrum handoff [11]. In the proactive-decision spectrum handoff, the suitable channels for future spectrum handoffs are determined before the data connection is established. As an input for the decision,

A preliminary version of this paper has been presented at the 2017 40th International Conference on Telecommunications and Signal Processing (TSP) [1].

P. Masek, and J. Hosek are with the Department of Telecommunications, Brno University of Technology, Czech Republic. They are also with Peoples' Friendship University of Russia (RUDN University), Miklukho-Maklaya St, Moscow, 117198, Russian Federation.

Z. Sadreddini is with Faculty of Communication, Giresun University, Giresun, Turkey.

T. Cavdar is with the Department of Computer Engineering, Karadeniz Technical University, Trabzon, Turkey.

Research described in this paper was financed by the National Sustainability Program under grant LO1401. For the research, infrastructure of the SIX Center was used. The research described in this paper was financed by the Ministry of Interior of Czech Republic project No. VI20162018007. The publication was prepared with the support of the "RUDN University Program 5-100".

Manuscript received October 27, 2017; revised December 18, 2017.

the connection parameters collected during long-term observations are utilized. Then, the secondary user can immediately change to the predetermined target channel, whenever it is required [12], [13].

In [14], various handoff strategies are classified into four models: (i) non-handoff, (ii) pure reactive, (iii) pure proactive, and (iv) hybrid handoff strategy. A qualitative comparison of these models is made with particular focus on the handoff latency performance metric. The analysis of handoff latency factors shows that pure proactive handoff strategy has the lowest handoff latency, while hybrid handoff and pure reactive handoff strategies have moderately long handoff latency. In [15], authors have proposed a PRP M/G/1 queuing network model to characterize the spectrum usage behavior with multiple handoffs. Another work [16] presents two approaches how to design the system parameters for the probability-based spectrum decision schemes and evaluate the latency performance of different spectrum handoff schemes. Further in [17], the adaptive handoff scheme based on Markov model is proposed. This model firstly allows an SU to decide whether to stay and wait on current channel or to perform handoff. Then, in case of handoff, an SU intelligently shifts between proactive or reactive handoff modes based on primary use (PU) arrival rate. Authors in [18] present a performance evaluation and comparative analysis of the proposed Feedback Fuzzy Analytical Hierarchical Process (FFAHP) together with six more algorithms of decision making used for the vertical handoff in cognitive radios. These algorithms utilized the multiple-criteria decision making (MCDM) method.

All above mentioned works are related to the spectrum handoff. However, none of them pay attention to Quality of Service (QoS) and Quality of Experience (QoE) in term of user satisfaction. Therefore, in this paper, we detail our proposed decision based model for the spectrum handoff in order to optimize the spectrum utilization and in addition to increase SUs' satisfaction while mitigating the DoS. We focus in detail on the modeling technique and performance analysis of the *proactive-decision spectrum handoff* scheme [11]. The obtained data from simulator is further analyzed against the real measurements performed in LTE-A indoor deployment [19]. In our proposed model, we consider 3GPP LTE-A deployment featuring the LSA concept as a new innovative framework, that enables more efficient usage of available frequency bands by allowing coordinated shared access to licensed spectrum [3], [5]. The whole mobile system is composed of three cells (band 17; 700 MHz; FDD), where the spectrum utilization is being continuously analyzed.

In this work, we investigate one specific stage of LSA operation, when the spectrum owner (PU) requests the spectrum back and therefore the handoff of SUs needs to be done in order to avoid ASUs' service interruption and so keep user satisfaction and cells' performance (data throughput) at negotiated QoS level. There are two key questions to be answered at the beginning of each handoff procedure: (i) are there any under-utilized spectrum resources (USRs), where the ASUs can be reallocated and (ii) which ASU should be handoff. To solve the first task, a regular analysis of available spectrum is required. To decide about the second, i.e., which ASUs needs to reallocated to substitute spectrum, we consider a variety of ASUs attributes (see Section III).

The main goal of our simulations was to evaluate spectrum utilization of examined 4G cells and calculate the total service time for each ASU, while considering spectrum handoff in case of different PU's activity ratio [20] for each cell.

The rest of the paper is organized as follows. Section II introduces the utilized 3GPP LTE-A testbed. Following Section III describes our proposed system model with respect to the spectrum handoff in next-generation (5G) mobile networks. Next, in Section IV, the obtained simulation results are revealed in detail. Finally, the conclusions are given in the last Section V.

## II. UTILIZED 3GPP LTE DEPLOYMENT

The prototype was developed within the 3GPP LTE deployment (see Fig. 1) which is located at BUT, Czech Republic. The LTE core is a fully-operational cellular infrastructure. The described LTE testbed is utilized for the purposes of research and education and its essential modules are listed in Table I.

TABLE I: Main components of our experimental 3GPP LTE deployment

| Core units | Components | Description |
|---|---|---|
| EPC | UGW (SGW, PGW) | Fully redundant 10 Gbps links. |
| | MME | Interface mirroring for probe-based analysis. |
| | HSS | |
| IMS | IMS-HSS | IMS core + RCS, Enables VoLTE, Public Safety Answering Point, Additional HSS, Full redundancy. |
| | ENUM/DNS | |
| | S-CSCF/MRFC | |
| | P-CSCF/A-SBC | |
| | MRFP | |

The installed experimental LTE-A (Release 10) network, see Fig. 1, stands for a complete commercial-grade implementation of all crucial subsystems including contemporary 4th generation mobile networks. Radio access network (RAN) comprises of two eNodeBs and one Home eNodeB. Indoor eNodeBs are divided into two elements: (i) Baseband unit (BBU), and (ii) Remote radio unit (RRU). As a eNodeB unit, the Huawei DBS3900 (fourth generation indoor macro base station) is utilized. It comprises of (i) BBU3900 (management logic for eNodeB e.g., mobility management of modulation techniques in base band), and (ii) RRU (modulation of used frequency). eNodeB 1 composes two cells (Cell A and Cell B) working in the band 17 (originally the AT&T band in US, 700 MHz). eNodeB 2 manages three cells (sectors) working at separated frequencies 700 MHz, 2600MHz (these two are indoor pico-cells), and outdoor 1800 MHz (not depicted in the figures). Next, the WiFi access points (APs) are incorporated (WiFi radio access network) together with the LTE cells in the network. APs, operating in ISM bands (2.4 GHz and 5 GHz), provide the possibility to use packet-switched data services as VoIP or VoLTE (Voice over LTE).

For testing the upcoming technologies, the evolved packet core (EPC) enables the possibility to access the RAN part of the LTE network up to 100,000 terminals (end-users). Following the recommendations given by Czech telecommunication

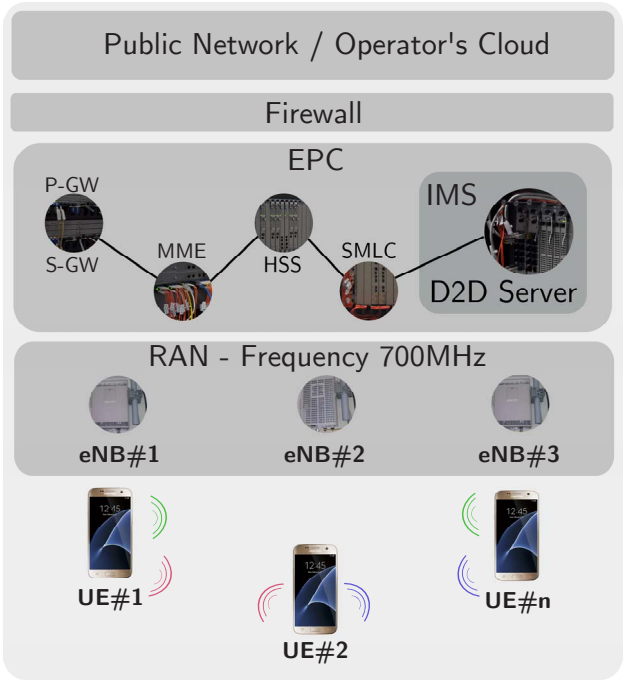



Fig. 1: Test 3GPP LTE-A deployment: structure and main modules of trialled mobile system.

office, the TX power for indoor cells was set to 1 W (30 dBm) and the provided data rates are up to 33 Mbps for downlink and up to 13 Mbps for uplink channel. On the top of that, the QoS and QoE techniques are implemented for served terminals.

This full-featured LTE testbed accommodates our research and educational purposes by allowing full access to the experimental 4th generation cellular network in order to obtain deeper understanding of its operation principles as well as to open the door to rapid and efficient prototyping of new wireless technology, as D2D communication or data offloading techniques.

While working on our intended trial setup, several modifications had to be done in the given LTE network – most frequently regarding the firewall policy and default configuration of the IP Multimedia System (IMS) part.

## III. System Model

The constant growth in mobile data traffic indicates that MNOs will no longer be able to meet this demand with their legacy approach on mostly fixed management of licensed spectrum [21]. Therefore, to increase the capacity of future cellular systems, the spectrum, as limited resource, needs to be handled more efficiently. In this paper, a decision-based spectrum handoff model enabling better spectrum utilization and in addition increasing SUs' satisfaction is proposed. In order to mimic real conditions as much as possible, during our work, we utilize a full-fledged mobile network at Brno University of Technology (BUT), Czech Republic, which offers an excellent example of a contemporary 3GPP LTE-A (Release 10) target system since this deployment has been recently complemented by the LSA framework [19], [22]. Setting up an LSA arrangement involves three parties: (i) incumbent, (ii) LSA licensees, and (iii) MNO[1].

- Incumbent (PU) has its rights to spectrum formalized or licensed during the process of establishing the sharing arrangement. The PU must adhere to the terms and conditions of its licensee / access rights.
- LSA licensee (SU) is granted to use spectrum resources that can be in long term period of time. While using the rented spectrum, the SU must adhere to the license / contract conditions with incumbent.

### A. Spectrum Handoff Functionality

The system model of proposed handoff procedure together with the interaction between PUs and SUs is shown in Fig. 2 and in Fig. 3.

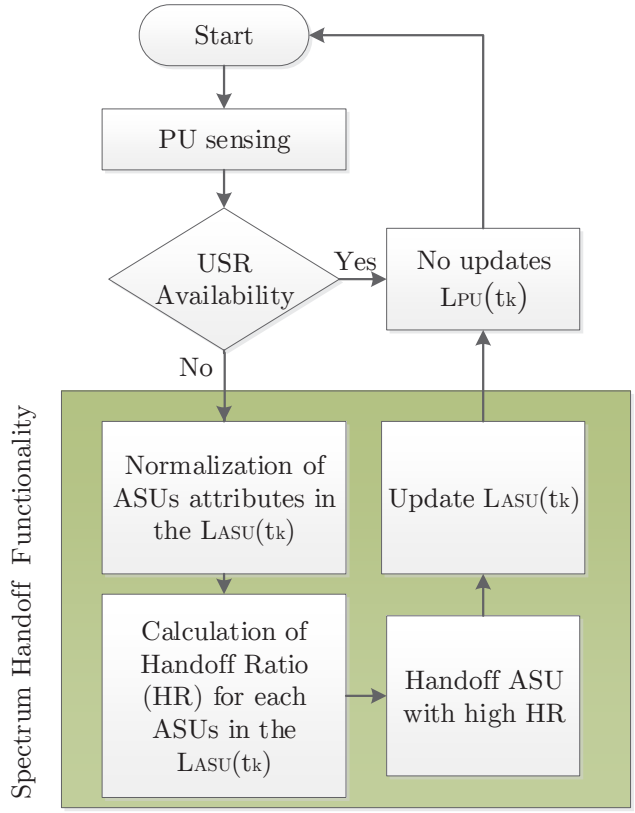



Fig. 2: Proposed system model of handoff procedure.

Firstly, the SU requesting the spectrum access (RSU) sends its request to MNO at $t_k$. Then, MNO checks the USRs' status at $t_k$, whether there is any USR or not [23]. If there is any available spectrum in the requested time slot, the RSU request will be accepted and its status will be changed to ASU. The ASU can get a service till there is enough USRs for both (i) sensed PU, and (ii) ASUs. If there is no available USR for the sensed PU, the case that is considered for the ASU spectrum handoff is described as follows

$$0 \leq USR(t_k) < |L_{PU}(t_k)|, \tag{1}$$

[1]LSA spectrum sharing framework enables spectrum sharing by allowing at least two users, the PU (i.e., current holder of spectrum rights (mobile operator)) and SU (i.e., temporary user of spectrum).

where $USR(t_k)$ shows the number of underutilized spectrum resources during the time slot $t_k$ and $|L_{PU}(t_k)|$ shows the number of sensed PUs at $t_k$ [23]. According to the $|L_{PU}(t_k)|$, MNO should handoff some ASUs following the list of ASUs

$$L_{ASU}(t_k) = \{ASU_j = (RSSI_j, RSRP_j, RSRQ_j, SINR_j) \\ |1 \leq j \leq \alpha(t_k), 1 \leq k \leq \Phi\}, \quad (2)$$

where $\Phi = T \times 60$, the interval between each time slot, is simulated for 60 seconds, where $T$ shows the latest time slot of the simulation process.

In the next step, according to Fig. 3, $RSU_1$ checks the USR status at $t_2$. If there is an USR at $t_2$ available, $RSU_1$ gets a service and his status is changed to $ASU_1$. This situation will continue till $t_3$, when PU requests the spectrum back and $ASU_1$ is handoffed to another USR (frequency band). After the PU finishes the transmission, another spectrum handoff occurs till $t_{13}$.

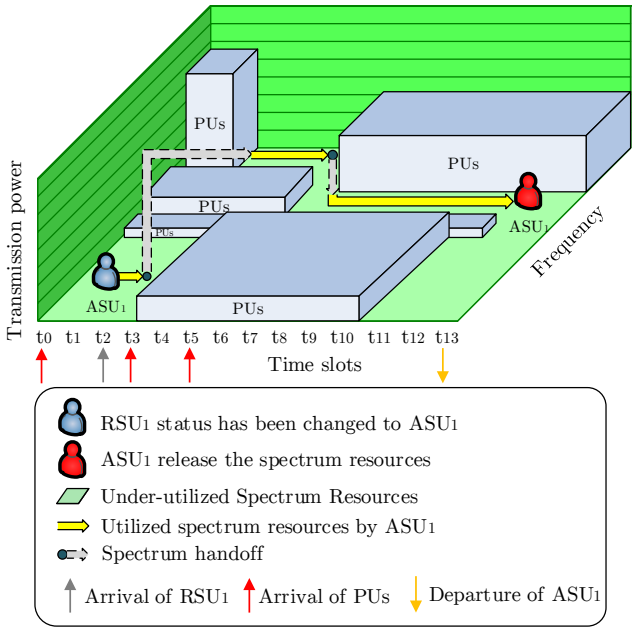



Fig. 3: Proposed system model of handoff procedure – detailed illustration.

### B. Decision-based Functionality

In this sub-section, the proposed decision-based spectrum handoff functionality is introduced. Following the condition given in Eq. (1), there are ASUs in the network, who are subjected at given time instance to state $L_{ASU}(t_k) > 0$. In this case, the spectrum handoff is conducted for existing ASUs in order to reconnect them to neighboring cells (with available spectrum resources) to avoid DoS.

At the very beginning of the whole handoff process, key ASUs attributes (RSSI, RSRP, RSRQ, and SINR) have to be normalized. To this end, we need to decide about benefit or cost (priority) of each attribute in handoff functionality. In case of our implementation, the RSSI, RSRP, and RSRQ represent the cost type (decreasing their value, the probability of ASU handoff increases); the detailed description of mentioned attributes is given below

- RSSI – Received Signal Strength Indicator: This parameter stands for a value that is crucial to determine if RSU has enough signal strength to get a demanded mobile connection.
- RSRP – Reference Signal Received Power: RSRP is a RSSI type of measurement, as follows there are some definition of it and some details as well.It is the power of the LTE Reference Signals spread over the full bandwidth and narrowband. A minimum of -20 dB SINR (of the S-Synch channel) is needed to detect RSRP/RSRQ.
- RSRQ – Reference Signal Received Quality: Quality considering also RSSI and the number of used Resource Blocks $(N)RSRQ = (N \times RSRP)/RSSI$ measured over the same bandwidth. RSRQ is a C/I type of measurement and it indicates the quality of the received reference signal. The RSRQ measurement provides additional information when RSRP is not sufficient to make a reliable handover or cell reselection decision.
- SINR – Signal-to-Noise Ratio: is a measure of signal strength relative to background noise.

In the procedure of handoff, the LTE specification provides the flexibility of using RSSI, RSRP, RSRQ, SINR or all of them. Regarding to the normalization of each attribute, it is needed to decide about benefit or cost (priority) type of attributes in handoff functionality. On the other hand, SINR stands for the benefit type i.e., whenever the SINR increases, the probability of the ASU handoff also rises.

After the normalization of the three cost types and one benefit type [24], the *handoff ratio (HR)* is calculated as follows

$$HR_{ASU}(t_k) = \sum_{i=1}^{|Cell|} \sum_{j=1}^{\sum_{L_{ASU}}^{Cell_i}(t_k)} \sum_{z=1}^{|Criterion|} r(j,z) \times \alpha, \quad (3)$$

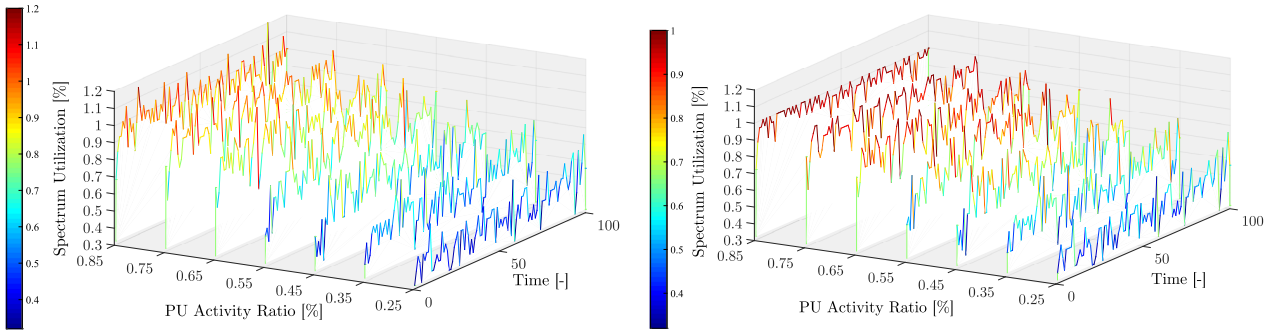
where $\sum_{L_{ASU}}^{Cell_i}(t_k)$ represents the list of ASUs in specific $Cell_i$ at time instance $t_k$ and the $|Criterion|$ is considered as 4 (equals to the number of selected attributes). The normalized value of the attributes for each ASU is given as $r(j,z)$; the decision matrix is shown in Table. II where

- $n$ shows the number of ASUs in the LASU,
- $m$ is the $|Criterion| = 4$; while the proposed scheme considers four criteria, the generality of the scheme permits the extension to any number of criteria as $m$.

Finally, $\alpha$ is calculated as $\frac{1}{|Criterion|}$; it proves same importance degree for SUs' attributes. After the HR is obtained, the ASUs with the highest HR are handoff to neighboring cell(s).

The normalized value of attributes for each ASU is given as $r(j,z)$ where in the decision matrix (Table. II), $r_{ij}$ shows the value of $j-th$ criterion for $i-th$ messages. In this case, the Eq. 4 is considered.

$$\sum_{j=1}^{m} \alpha_j = 1. \quad (4)$$




(a) (b)

Fig. 4: Spectrum utilization before (Fig. 4a) and after (Fig. 4b) the implementation of the proposed handoff mechanism.

TABLE II: Decision matrix

| List of ASUs | Criteria | | | | | |
|---|---|---|---|---|---|---|
| | $C_1$ | $C_1$ | $\cdots$ | $C_j$ | $\cdots$ | $C_m$ |
| $ASU_1$ | $r_{11}$ | $r_{12}$ | $\cdots$ | $r_{1j}$ | $\cdots$ | $r_{1m}$ |
| $ASU_2$ | $r_{21}$ | $r_{22}$ | $\cdots$ | $r_{2j}$ | $\cdots$ | $r_{2m}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $ASU_i$ | $r_{i1}$ | $r_{i2}$ | $\cdots$ | $r_{ij}$ | $\cdots$ | $r_{im}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $ASU_n$ | $r_{n1}$ | $r_{n2}$ | $\cdots$ | $r_{nj}$ | $\cdots$ | $r_{nm}$ |

## IV. SIMULATION RESULTS

To adequately evaluate the proposed system model (see Section III), as an input for the developed simulation framework, we used the data set obtained from the experimental 3GPP LTE-A system located at Department of Telecommunications, BUT, Czech Republic, which supports functionality of LTE Release 10 communications system [25]. The LTE-A network consists of three cells denoted as Cell 11 (ID1); Cell 12 (ID2); Cell 22 (ID3). The total number of available resource blocks (RBs) in each cell is 25 (devoted from the available 5 MHz bandwidth) – see Table III where all simulation parameters are listed.

Fig. 4a shows the spectrum utilization of the Cell 12 before the proposed spectrum handoff mechanism is applied. Further, the Fig. 4b represents the situation, where our proposed spectrum handoff functionality takes place. In both cases, the spectrum utilization depends on PUs' activity ratio [0.25 – 0.85] during 100 time intervals. Also in both scenarios (Fig.4a and Fig. 4b), the percentage of USRs obviously depends on the PU's activity. If we consider higher PUs' activity (0.85), spectrum utilization over the 100 % of usable spectrum capacity is evident – i.e., in specific $t_k$ time slots, the case given by condition $0 \leq USR(t_k) < |L_{PU}(t_k)|$ occurred and some of the ASUs have to be disconnected from the network.

During the simulation, the mentioned condition has occurred at $t_{60}$. At $t_{60}$, $USR(t_{60}) = 5$ and $|L_{PU}(t_{60})|$ is sensed as 6. Thus, one of the ASU in $L_{ASU}$ should be handoffed. Table IV shows the ASUs list at $(t_{60})$ $(L_{ASU_{t_{60}}})$. There are 8 ASUs in total and the values of the all four attributes for each ASU

TABLE III: Simulation parameters

| Parameter | | Value |
|---|---|---|
| 3GPP LTE system baseline | | Release 10 |
| Division multiplexing | | FDD |
| Number of cells | | 3 |
| Frequency band | | 17 (700 MHz) |
| Bandwidth | | 5 MHz |
| Max. eNB power level | | 0 dB |
| Min. eNB power level | | -30 dB |
| Interference threshold | | -85 dBm |
| Simulation time ($\theta$) | | [100] seconds |
| Total number of RBs (USR) | | [25] RB |
| $PU_{ActivityRatio}$ | | [0.25; 0.85] |
| Base price ($P_b$) | | 100 mu/SB |
| Criterion | RSSI | [-79; -63] dBm |
| | RSRP | [-100; -30] dBm |
| | RSRQ | [-10; -6] dBm |
| | SINR | [1; 30] dB |

are exist in the $L_{ASU}(t_{60})$. According the $HR_{ASU}$ $(t_{60})$, the suitable ASU will be selected; in this specific case, the $ASU_{74}$ is selected – where RSRP value is minimum among all ASUs where RSRP $\leq -100$ proves the 0% quality. If the $|LPU(t60)|$ was more than 6, the ASUs with High HR as $ASU_{56}$, $ASU_{92}$ would be handed-off, respectively.

TABLE IV: History Ratio $(HR_{ASU}(t_{60}))$ of ASUs $(L_{ASU}(t_{60}))$

| $ASU_{ID}$ | $L_{ASU}(t_{60})$ | | | | $HR_{ASU}(t_{60})$ |
|---|---|---|---|---|---|
| | RSRP [dBm] | RSRQ [dB] | SINR [dB] | RSSI [dBm] | |
| ASU8 | -74 | -7 | 30.00 | -64 | 0.5156 |
| ASU38 | -75 | -7 | 17.80 | -64 | 0.4089 |
| ASU30 | -56 | -7 | 30.00 | -63 | 0.3875 |
| ASU26 | -44 | -6 | 14.40 | -63 | 0.1055 |
| ASU74 | -84 | -8 | 19.80 | -70 | 0.6399 |
| ASU56 | -63 | -10 | 3.0 | -79 | 0.6187 |
| ASU92 | -63 | -10 | 3.0 | -79 | 0.6187 |
| ASU115 | -62 | -6 | 30.00 | -63 | 0.3625 |

Following the proposed implementation of the spectrum handoff (see Fig. 4b), some ASUs (originally selected to be

rejected) can be handoffed to the neighboring cell instead of facing the DoS issue. If we take a look at the spectrum utilization with lower PU's activity ratio (0.25), we can see similar progression in both cases (Fig.4a and Fig. 4b). Finally, the Fig. 5 details the spectrum handoff functionality for each of the ASU in all utilized cells. In our work, we categorize the ASUs as follows:

- ASUs with ID 7, ID 8, ID 9 (marked by blue color in Fig.5) continue their service in Cell 11. Initial conditions: $ASU_7$ and $ASU_9$ started their operation in Cell 22, whereas $ASU_8$ in Cell 12.
- ASUs (green color) with ID 4 and ID 5 have been handoffed to the Cell 12 after some ping pong handoff and finally they have been continued in Cell 12.
- ASUs with ID 1 (originally located in Cell 11), ID 2 (located in Cell 12), ID 3 (originally located in Cell 12) and ID 6 (originally located in Cell 12) continue their operations in Cell 22 as shown by gray color in Fig. 5.

Each of the ASU has started service operation in different time slot and based on the USRs situation, they had to reallocated to another (neighboring) cell(s) e.g., $ASU_2$ requests the service connection at time instance $t_{20}$ and gets the service till $t_{23}$ within the Cell 12. After $t_{23}$, $ASU_2$ has been handoffed to Cell 11. The $ASU_2$ performs service in Cell 11 till $t_{28}$ and then has been handoffed to the Cell 22 till $t_{37}$.

We have also investigated the reason for inconsistencies in spectrum handoff for several ASUs, which depend on PU's activity ratio. For instance, $ASU_8$ does several handoffs at the beginning of the service operation – therefore, we should focus also on the overall system efficiency and complexity, which leads us to the tradeoff between spectrum handoff and system efficiency.

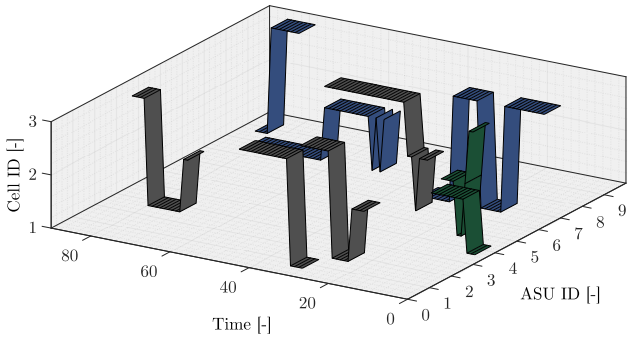



Fig. 5: Spectrum handoff situation for each ASU in each of the considered cells.

For the sake of clarity in terms of user satisfaction, stability of spectrum utilization in each cell is illustrated in Fig. 6, Fig. 7 and Fig. 8 for Cell 11, Cell 12 and Cell 22, respectively. The point is that the the PU activity ratio has the direct effect to the spectrum utilization and handoff times. Showing the stability of the getting service in the same spectrum resources, PU activity ratio is 0.25 in the all above mentioned figures (Fig. 6, Fig. 7 and Fig. 8). Thus, handoff procedure does not occurred continuously.

As it can be clearly seen in the in Fig. 6, Fig. 7 and Fig. 8, most of the ASUs continue the data transmissions in the same cell they started the service or they have handoffed just for

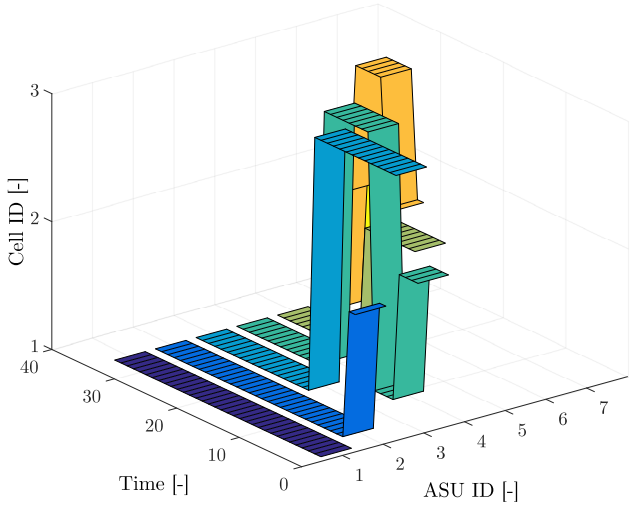



Fig. 6: Spectrum handoff situation for each ASU in coverage of Cell 11.

one time. Only in case of ASUs with ID 4, 7, 9, 11, 15, 16 and 17, they have handoffed more than once. Although, this is occurred for almost 40% of ASUs, we should take into account that the PU activity is low. However, fluctuation can be occurred continuously in peak-hours which is occurred for $ASU_8$ in Fig. 5. In this case, tradeoff between handoff and DoS for ASUs should be considered. Finally, network operator can handoff or deny ASU based on its attributes and number of handoff procedures.

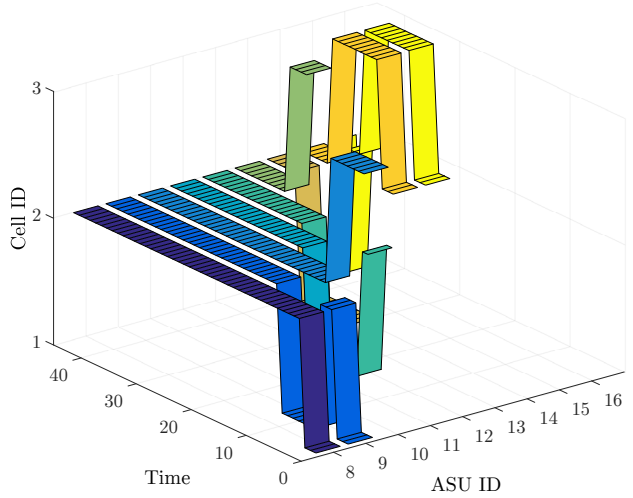



Fig. 7: Spectrum handoff situation for each ASU in coverage of Cell 12.

## V. CONCLUSION

This work accentuates the importance of the decision-based spectrum handoff mechanism in real 3GPP LTE-A mobile system under LSA framework to leverage additional bandwidth that may be lightly used by its original incumbents. To further improve upon spectrum utilization in demanding 5G systems, we focus on the emerging LSA framework for vertical sharing, where the incumbent(s) and the licensee(s) operate over the same geographical area by utilizing common frequencies in a carefully controlled manner.

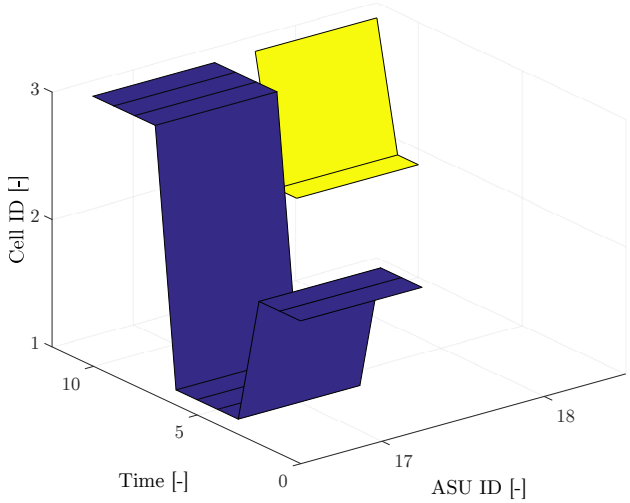



Fig. 8: Spectrum handoff situation for each ASU in coverage of Cell 13.

In this work, we investigate one specific stage of LSA operation, when the spectrum owner (PU) requests the spectrum back handoff procedure is occured. Therefore the handoff of SUs needs to be done in order to avoid ASUs service interruption and so keep user satisfaction (QoE) and cells' performance (data throughput) at negotiated QoS level. There are two key questions to be answered at the beginning of each handoff procedure: (i) are there any under-utilized spectrum resources (USRs), where the ASUs can be reallocated and (ii) which ASU should be handed-off. To solve the first task, a regular analysis of available spectrum is required. To decide about the second, i.e., which ASUs needs to reallocated to substitute USR, we consider a variety of ASUs attributes.

Based on the raised questions, we have created simulation tool that was calibrated by data-set provided by 3GPP and also data-set from our 3GPP LTE-A mobile network located at BUT, Czech Republic. Obtained simulation results prove that USRs are available, when the activity of the PU stagnates. Therefore, spectrum handoff mechanism can be applied instead of ASUs' denial-of-service. We have also carefully analyzed the obtained results with respect to the spectrum utilization for each of the examined cells within the 3GPP testbed in off peak hours. In our future work, we will pay attention to the implementation of the *dynamic* handoff system and tradeoff between handoff and rejection of ASUs specially in high PU activity ratio.

## REFERENCES

[1] P. Masek, Z. Sadreddini, T. Cavdar, and J. Hosek, "Modeling of spectrum handoff in 3gpp lte-a indoor deployment," in *Telecommunications and Signal Processing (TSP), 2017 40th International Conference on*. IEEE, 2017, pp. 204–207.

[2] Ericsson, "The Spectrum Crunch - busting the solutions myth," 2016. [Online]. Available: https://www.ericsson.com/openarticle/the-spectrum-crunch_1473414408_c

[3] K. Buckwitz, J. Engelberg, and G. Rausch, "Licensed shared access (LSA)Regulatory background and view of administrations," in *Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), 2014 9th International Conference on*. IEEE, 2014, pp. 413–416.

[4] N. U. Hasan, W. Ejaz, N. Ejaz, H. S. Kim, A. Anpalagan, and M. Jo, "Network selection and channel allocation for spectrum sharing in 5G heterogeneous networks," *IEEE Access*, vol. 4, pp. 980–992, 2016.

[5] M. Mustonen, M. Matinmikko, M. Palola, S. Yrjölä, and K. Horneman, "An evolution toward cognitive cellular systems: Licensed shared access for network optimization," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 68–74, 2015.

[6] H. U. Sokun, R. H. Gohary, and H. Yanikomeroglu, "A novel approach for qos-aware joint user association, resource block and discrete power allocation in hetnets," *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7603–7618, 2017.

[7] T. A. Weiss and F. K. Jondral, "Spectrum pooling: an innovative strategy for the enhancement of spectrum efficiency," *IEEE communications Magazine*, vol. 42, no. 3, pp. S8–14, 2004.

[8] S. Yrjölä, V. Hartikainen, L. Tudose, J. Ojaniemi, A. Kivinen, and T. Kippola, "Field Trial of Licensed Shared Access with Enhanced Spectrum Controller Power Control Algorithms and LTE Enablers," *Journal of Signal Processing Systems*, vol. 89, no. 1, pp. 119–132, 2017.

[9] L. C. Wang, C. W. Wang, and C. J. Chang, "Modeling and Analysis for Spectrum Handoffs in Cognitive Radio Networks," *IEEE Transactions on Mobile Computing*, vol. 11, no. 9, pp. 1499–1513, Sept 2012.

[10] J. Crichigno, M.-Y. Wu, and W. Shu, "Protocols and architectures for channel assignment in wireless mesh networks," *Ad Hoc Networks*, vol. 6, no. 7, pp. 1051–1077, 2008.

[11] L. C. Wang and C. W. Wang, "Spectrum Handoff for Cognitive Radio Networks: Reactive-Sensing or Proactive-Sensins?" in *2008 IEEE International Performance, Computing and Communications Conference*, Dec 2008, pp. 343–348.

[12] S. Srinivasa and S. A. Jafar, "The Throughput Potential of Cognitive Radio: A Theoretical Perspective," in *2006 Fortieth Asilomar Conference on Signals, Systems and Computers*, Oct 2006, pp. 221–225.

[13] H. Su and X. Zhang, "Channel-hopping based single transceiver MAC for cognitive radio networks," in *2008 42nd Annual Conference on Information Sciences and Systems*, March 2008, pp. 197–202.

[14] I. Christian, S. Moh, I. Chung, and J. Lee, "Spectrum mobility in cognitive radio networks," *IEEE Communications Magazine*, vol. 50, no. 6, 2012.

[15] L.-C. Wang, C.-W. Wang, and C.-J. Chang, "Modeling and analysis for spectrum handoffs in cognitive radio networks," *IEEE Transactions on Mobile Computing*, vol. 11, no. 9, pp. 1499–1513, 2012.

[16] L.-C. Wang, C.-W. Wang, and K.-T. Feng, "A queueing-theoretical framework for QoS-enhanced spectrum management in cognitive radio networks," *IEEE Wireless Communications*, vol. 18, no. 6, 2011.

[17] U. Mir and A. Munir, "An adaptive handoff strategy for cognitive radio networks," *Wireless Networks*, pp. 1–16.

[18] C. A. H. Suárez, L. F. P. Martínez, and E. R. de la Colina, "Fuzzy feedback algorithm for the spectral handoff in cognitive radio networks," *Revista Facultad de Ingeniería*, no. 81, pp. 47–62, 2016.

[19] P. Masek, E. Mokrov, A. Pyattaev, K. Zeman, A. Ponomarenko-Timofeev, A. Samuylov, E. Sopin, J. Hosek, I. A. Gudkova, S. Andreev, V. Novotny, Y. Koucheryavy, and K. Samouylov, "Experimental Evaluation of Dynamic Licensed Shared Access Operation in Live 3GPP LTE System," in *2016 IEEE Global Communications Conference (GLOBECOM)*, Dec 2016, pp. 1–6.

[20] Y. Saleem and M. H. Rehmani, "Primary radio user activity models for cognitive radio networks: A survey," *Journal of Network and Computer Applications*, vol. 43, pp. 1–16, 2014.

[21] M. Mustonen, T. Chen, H. Saarnisaari, M. Matinmikko, S. Yrjola, and M. Palola, "Cellular architecture enhancement for supporting the european licensed shared access concept," *IEEE Wireless Communications*, vol. 21, no. 3, pp. 37–43, 2014.

[22] A. Ponomarenko-Timofeev, A. Pyattaev, S. Andreev, Y. Koucheryavy, M. Mueck, and I. Karls, "Highly dynamic spectrum management within licensed shared access regulatory framework," *IEEE Communications Magazine*, vol. 54, no. 3, pp. 100–109, 2016.

[23] T. Çavdar, E. Güler, and Z. Sadreddini, "Instant overbooking framework for cognitive radio networks," *Computer Networks*, vol. 76, pp. 227–241, 2015.

[24] R. Ginevičius, "Normalization of quantities of various dimensions," *Journal of business economics and management*, vol. 9, no. 1, pp. 79–86, 2008.

[25] P. Masek, M. Slabicki, J. Hosek, and K. Grochla, "Transmission power optimization in live 3GPP LTE-A indoor deployment," in *Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2016 8th International Congress on*. IEEE, 2016, pp. 164–170.

**Pavel Masek** received his M.S. and Ph.D. degrees in electrical engineering from the Faculty of Electrical Engineering and Communication at Brno University of Technology (BUT), Czech Republic, in 2013 and 2017, respectively. He has publications on a variety of networking-related topics in internationally recognized venues including those published in the IEEE Communications Magazine, as well as several technology products. His primary research interest lies in the area of M2M / H2H / D2D communication, next-generation cellular networks, heterogeneous networking, and data offloading techniques.

**Zhaleh Sadreddini** received the B.S. degree in Computer Science from Nabi-Akram University, Tabriz, Iran, and the MsC degree in Computer Engineering from Islamic Azad University, Shabestar, Iran, in 2012. She is currently pursuing a Ph.D. degree in Computer Engineering under the supervision of Assist. Prof. Dr. Tugrul Cavdar at Karadeniz Technical University, Trabzon, Turkey. In 2014, she started to work at Avrasya University as a teaching assistant. Now, she is working as a lecturer and teaching various computer courses in Faculty of Communication at Giresun University. Her primary research interest lies in the area of cognitive radio networks, next-generation cellular networks, heterogeneous networking, spectrum management and multi criteria decision making techniques.

**Tugrul Cavdar** was born in Trabzon, Turkey, in 1973. He received the B.S., M.Sc., and Ph.D. degrees in Electronics Engineering from Karadeniz Technical University, Trabzon, Turkey in 1994, 1997 and 2003, respectively. He joined the Department of Computer Engineering of the same university as an Assistant Professor in 2007. He studied on Cognitive Radio Networks as a visiting professor in Northeastern University, Boston, USA in 2012 - 2013. Currently, he is the Chair of the Department of Computer Hardware, and the Director of Wireless Communications & Networks Laboratory at the Department of Computer Engineering in Karadeniz Technical University. His research interests are Wireless Communications, Wireless Networks, CRN, WSN, IoT, 5G, Channel Equalization and Robotics.

**Jiri Hosek** received his M.S. and Ph.D. degrees in electrical engineering from the Faculty of Electrical Engineering and Communication at Brno University of Technology (BUT), Czech Republic, in 2007 and 2011, respectively. He is currently working as an associate professor and senior researcher at the Department of Telecommunications, BUT. His research work has concentrated on industry-oriented R&D projects in the area of future mobile networks, the Internet of Things, and home automation services. He has (co-)authored more than 70 research works on networking technologies, wireless communications, quality of service, user experience, and IoT applications.

# On Sensitive and Weighted Routing and Placement Schemes for Network Function Virtualization

Diogo Oliveira, Jorge Crichigno and Nasir Ghani

***Abstract*—Virtualization is a fast-growing technology that is being widely adopted to help improve network and datacenter resource manageability and usage optimization. However, given increasing deployments, new challenges are starting to arise, e.g., such as management complexity. Hence in order to deliver a higher degree of service provisioning flexibility, two key technologies have attracted attention, namely network function virtualization (NFV) and software defined networking (SDN). The former enables the implementation of network functions (NFs) via top-of-the-shelf commodity servers in datacenters. The latter decouples the data and control planes, centralizing flow rules definitions in a controller system to facilitate management and routing. Although recent NFV studies have focused on minimizing resource usage to satisfy a set of requested NFs, they do not consider scenarios with limited resources. Hence this paper presents an optimization-based solution for the joint routing and placement of virtual NFs. In particular, the scheme tries to maximize the number of satisfied requests as well as well minimize routing and deployment costs. The model also introduces weighting factors to allow operators to select cost preferences. However findings indicate that the proposed optimization solution can only be solved for smaller networks. Hence a more scalable greedy heuristic scheme is also developed.**

***Index Terms*—Greedy algorithm; integer linear programming (ILP); network function virtualization (NFV).**

## I. INTRODUCTION

NETWORK function virtualization (NFV) is a paradigm that allows network functions (NFs) to be executed on commercial-of-the-shelf (COTS) commodity servers, e.g., such as firewalls, load balancers, address translation boxes, etc. Namely, it decouples physical network devices from the functions that run on them, allowing such functions to be executed as software instances within datacenters. As such, NFV reduces capital and operational expenses by replacing embedded devices and facilitating the deployment and management of networking services [1].

Meanwhile, software defined networking (SDN) is being widely adopted due to its dynamic traffic flow management capabilities. This methodology decouples the data and control planes, relegating decision-making capacity to a controlling system called the SDN controller. As such, network control becomes much more flexible, and the underlying infrastructure can be reduced to a layer of programmable packet forwarding devices. Namely, SDN controllers can build global network views and directly compute traffic flow rules for routing packets from source to destination nodes [2].

In general, NFV and SDN concepts have an orthogonal relationship. However, even though these two paradigms do not depend upon each other, they are still very complementary. Namely, the former dynamically deploys NFs as software instances, whereas the latter steers traffic to datacenters hosting the desired NFs associated with a request. Now each service request is comprised of a source and destination pair, a set of requested NFs and a load rate. Hence a service provider instantiates the pertinent NFs (associated with the network services that are offered) in the datacenters and defines the proper flow rules in the SDN controller(s). This procedure helps establish the traffic flow between the source and the destination. Now clearly each traffic flow must pass through the datacenters hosting the requested NFs at the minimum data rate requested. However, the *sequence* in which desired functions are applied to a traffic flow is not considered here, i.e., service function chaining (SFC).

To date, existing studies have focused on the minimization of resources required for orchestration and placement of NFs under the assumption that datacenters have infinite resources to satisfy all requests [3], [4], [5], [6]. However, in resource constrained environments, the underlying physical infrastrurcture may not be able to satisfy all requests, i.e., during periods of heavy load or after large failure events. To the best of the authors' knowledge, the only known work on NF placement (to maximize the number of satisfied requests) in resource-constrained environments is presented in [7]. However this initial study does not consider link capacity constraints, link load balancing, or weighting factor selection/adjustment.

In light of the above, this paper presents a novel integer linear programming (ILP) solution for weighted joint routing and placement of NFs, i.e., termed as the *Minimized Link Load* ILP (MLL-ILP) scheme. The proposed scheme tries to maximize the number of requested NFs, and also minimize the deployment cost, routing cost and link load. Additionally, the scheme implements weighting factors for all three minimization costs to allow operators to fine tune NF placement and routing according to their operational needs. However, the MLL-ILP formulation is still intractable for larger network sizes. As a result, a greedy heuristic scheme is also developed to improve scalability, i.e., termed as the *Minimized Link Load Greedy* (MLL-GR) scheme.

This paper is organized as follows. First, Section II discusses some related work on NF placement, and then Section III

A preliminary version of this paper has been presented at the 2017 40th International Conference on Telecommunications and Signal Processing (TSP) [7].

D. Oliveira and N. Ghani are with the Department of Electrical Engineering, University of South Florida, Tampa, FL, 33620.
E-mail: diogoo@mail.usf.edu, nghani@usf.edu

J. Crichigno is with the College of Engineering and Computing, University of South Carolina, Columbia, SC, 29208.
E-mail: jcrichigno@sc.edu

formulates the proposed ILP and heuristic algorithms. Next, Section IV presents some detailed performance evaluation results, and Section V highlights the overall conclusions.

## II. Related Work

As noted earlier, NFV is a very recent technology, the first white paper on which was published by the European Telecommunications Standards Institute (ETSI) in October 2012. However since then many researchers have studied the NF placement problem. For example, Coen et al. [4] present an ILP to minimize NF deployment cost, making use of multiple approximation techniques. However, this work does not consider routing costs. Following the minimization model, Addis et al. [3] define an ILP scheme to reduce the number of processors used to satisfy NF requests. However, this solution can only be applied to a small number of instances due to its complexity, i.e., high numbers of variables/constraints. Focusing on heterogeneous Layer 1 networks, Xia et al. [5] propose a method to minimize NF placement cost in hybrid networks comprised of optical and electronic network elements, i.e., optical-to-electronic and electronic-to-optical conversion costs. As a complementary solution, the authors also present a heuristic algorithm. Finally, Gebert et al. [6] present a heuristic NF placement scheme to allow service providers to optimize cellular coverage for large events.

However, to the best of the authors' knowledge, the only work which addresses the problem of maximizing the number of requested NFs while minimizing infrastructure (node, link) load and routing costs is presented in [7]. This technique is particularly germane in scenarios where datacenters may not have enough resources to satisfy all requests. Extending upon this, the work herein proposes additional schemes that incorporate link capacity constraints as well as load balancing. Furthermore, the proposed solution can also be adapted to suit provider needs. These flexible characteristics allows providers to fine tune NF placement according to their requirements, i.e., both in terms of routing cost and NF deployment cost. Additionally, link capacity constraints and load balancing features deliver a more realistic solution, i.e., since both physical node and transmission link resources can be limited in real-world settings. These solutions are now presented.

## III. Weighted Joint Routing and Placement of NFs

As mentioned above, most NF placement schemes focus on minimizing placement cost by assuming unlimited network resources to satisfy all requests. Conversely, this paper looks at resource-limited scenarios and proposes two schemes (MLL-ILP and MLL-GR) to maximize the number of satisfied requests and reduce infrastructure and routing costs. Additionally, each cost has appropriately-defined weighting factors to create an adaptive model. For example if a provider has limited physical resources but sufficient link and routing capacity, then deployment cost can be given higher preference. On the other hand, if link capacities are lower and/or hops and delay times are higher, then reduced routing costs may be more favorable. Hence the proposed schemes allow providers to achieve a proper tradeoff between deployment and routing costs. Consider the details now.

### A. Optimization Model

The network topology is represented as a graph $G = (V, E)$, where $V$ is the set of nodes and $E$ the set of links. Each link $(i, j) \in E$ also has an associated cost $c^{ij}$ and capacity $b^{ij}$, which quantifies the cost of using that link and its link capacity, respectively. Meanwhile, a subset of nodes $D \subseteq V$ represents the set of datacenters where NFs are implemented, and the set of NFs is denoted by $F$. A given datacenter $d \in D$ implements a subset of functions $F_d \subseteq F$. Furthermore, $R$ represents the set of requests. Namely, each request $r \in R$ is characterized by a 4-tuple ($src_r$, $dst_r$, $F_r$, $b_r$), which denotes the source and destination nodes of the flow, the set of requested functions $F_r \subseteq F$, and the minimum available link capacity, respectively.

The customizable number of resource types is also denoted by an integer $m$. For example, $m$=3 can refer to processor, storage and memory. It is also assumed that a datacenter $d \in D$ has a finite amount of resources $W_d = \{w_{d,1}, w_{d,2}, ..., w_{d,m}\}$. Hence to implement a function $i \in F_d$, datacenter $d \in D$ employs $w^i_{d,1}, w^i_{d,2}, ..., w^i_{d,m}$ resources. This resource requirement is datacenter dependent, which reflects the fact that some datacenters may specialize in supporting certain functions. Also, the setup cost of locating an instance of a function $i \in F_d$ at datacenter $d$ is $c^i_d$, and an instance of function $i$ at datacenter $d$ can serve $\lambda^i_d$ requests. In order to accommodate more requests, multiple instances of function $i$ can also be deployed at datacenter $d$. However each instance will consume additional resources and entails added setup cost.

Overall, the MLL-ILP objective function tries to maximize the total number of satisfied NFs and is composed of four terms weighted by their factors, i.e., $w_1, w_2, w_3$ and $w_4$:

$$\begin{aligned}\max \text{ F} \quad = \quad & w_1 \sum_{r \in R} \sum_{i \in F_r} \sum_{d \in D | i \in F_d} x^i_{r,d} - w_2 \sum_{d \in D} \sum_{i \in F_d} c^i_d y^i_d \\ & - w_3 \sum_{r \in R} \sum_{(i,j) \in E} c^{ij} l^{ij}_r - \alpha w_4 \end{aligned} \tag{1}$$

The first term represents the total number of requested NFs. Meanwhile the second term is the total cost to setup/deploy NFs at the various datacenters. The third term is the total routing cost, and the fourth term represents the maximum overall link load. Note that the second, third and fourth terms are negative, since maximizing a negative term is equivalent to minimizing it [8], [9]. Also, the setup cost is directly related to the number of satisfied functions and number of instances of each function. Hence the second term (representing setup or deployment cost) depends upon the number of NFs per request and on how many datacenters are used to host these NFs. On the other hand, the third and fourth terms (representing routing and maximum link load costs, respectively) depend upon the number of requests and number of links used. These latter two terms are independent of the number of NFs and NF instances.

Next, Eqs. 2 to 12 list the complete set of model constraints:

$$x^i_{r,d} \quad \in \quad \{0, 1\} \quad r \in R, i \in F_r, d \in D | i \in F_d \tag{2}$$

$$y^i_d \quad \in \quad Z^+ \quad d \in D, i \in F_d \tag{3}$$

$$l_r^{i,j} \in \{0,1\} \quad r \in R, (i,j) \in E \tag{4}$$

$$0 \geq \alpha \leq 1 \tag{5}$$

$$\sum_{d \in D} x_{r,d}^i \leq 1 \quad r \in R, i \in F_r \tag{6}$$

$$x_{r,d}^i \leq y_d^i \quad r \in R, i \in F_r, d \in D | i \in F_d \tag{7}$$

$$\sum_{i \in F_d} w_{d,j}^i y_d^i \leq w_{d,j} \quad d \in D, r \in R, j \in \{1,2,...,m\} \tag{8}$$

$$\sum_{r \in R} x_{r,d}^i \leq \lambda_d^i y_d^i \quad d \in D, i \in F_d \tag{9}$$

$$\sum_{j:(i,j) \in E} l_r^{ij} - \sum_{j:(j,i) \in E} l_r^{ji} = \begin{cases} -1; i = dst_r, src_r \neq dst_r \\ 1; i = src_r, src_r \neq dst_r \\ 0; \text{ otherwise. } i \in V, r \in R \end{cases} \tag{10}$$

$$\sum_{(d,j) \in E} l_r^{dj} \geq x_{r,d}^i \quad r \in R, i \in F_r, d \in D | i \in F_d \tag{11}$$

$$\sum_{r \in R} l_r^{i,j} b_r \leq \alpha b_{i,j} \quad \{i,j\} \in E \tag{12}$$

The above constraints are now detailed further. Foremost, variable $x_{r,d}^i$ in Constraint 2 indicates whether function $i \in F_r$ requested by request $r \in R$ is implemented at datacenter $d \in D$. Meanwhile, variable $y_d^i$ in Constraint 3 represents the number of instances of function $i \in F$ at node $d \in D$. Variable $l_r^{i,j}$ in Constraint 4 indicates whether link $(i,j) \in E$ is used to route traffic flow for request $r \in R$, i.e., it is binary. Also, Constraint 5 limits $\alpha$ between 0 and 1 since this variable represents the highest link usage ratio, i.e., sum of all $b_r$ using a link $(i,j) \in E$ divided by link capacity $b_{i,j}$. Constraint 6 indicates that a function $i$ requested by request $r$ is serviced by at most one datacenter $d$. Since the objective of the ILP is to maximize the sum of all variables $x_{r,d}^i$ (first term in Eq. 1), the optimal solution will drive Constraint 6 to equality. Meanwhile Constraint 7 states that if request $r$ is assigned to function $i$ at datacenter $d$, then function $i$ is located at $d$. Constraint 8 also states that the aggregate amount of type $j$ resources used by all functions instantiated at datacenter $d$ is limited by the total amount of resources $w_{d,j} \in W_d, j \in \{1,2,...,m\}$. Similarly, Constraint 9 indicates that the total number of requests for function $i$ served at datacenter $d$ is at most the number of instances of $i$ at $d$ times the capacity $\lambda_d^i$ of an instance $i$. Meanwhile, Constraint 10 ensures necessary flow conservation. Finally, Constraint 11 guarantees that if a function $i$ requested by request $r$ is placed at datacenter $d$, then the traffic flow will be routed through that datacenter.

Now one of the key goals of MLL-ILP is also to avoid link overload or at least minimize link load. Therefore the fourth term in Eq. 1 introduces a link load minimization function variable $\alpha$, which is linked to Constraint 12. Namely, this constraint checks if the sum of all links $(i,j)$ used by a traffic flow associated with request $r$ times the load $b_r$ demanded by request $r$ is bounded by the product of the link capacity $b_{i,j}$ times $\alpha$. As a result, traffic flows are associated with links that have little/no load profile (higher available capacity).

Note that a non-link load balancing optimization scheme can also be defined by eliminating $\alpha$, i.e., $\alpha$=0. This instance is the same as the scheme presented in [7] and is termed as the *standard routing and placement ILP* (STD-ILP) scheme.

### B. MLL-ILP Complexity

Since the objective function and all constraints are linear, and all variables are either integer or binary, the MLL-ILP problem is NP-hard [10]. Furthermore, the complexity of MLL-ILP can be determined by the number of variables that the scheme utilizes. Namely, the number of variables $x_{r,d}^i$, $y_d^i$ and $l_r^{i,j}$, given by Constraints 2, 3 and 4 is upper-bounded by $|R||F||D|$, $|F||D|$ and $|R||E|$, respectively. Also, Constraint 5 refers to a single variable, $\alpha$. Hence the upper-bounds for the number of variables in Constraints 6, 7, 8 and 9 are $|R||F|$, $|R||F||D|$, $|R||D||W_d|$ and $|F||D|$, respectively. Meanwhile, the upper-bounds on variable counts for Constraints 10 and 11 are $|V||R|$ and $|R||F||D|$. Also, the upper-bound for Constraint 12 is $|V||R|+1$. Hence the upper-bound for the total number of variables is dominated by the product $|R||F||D|$. Now in general it is very difficult to pre-define limits on the number of requests or functions to ensure ILP convergence. However for small to medium network topologies, the proposed ILP can still be solved in a reasonable amount of time, i.e., low hundreds of nodes.

### C. MLL-GR Heuristic

Although the MLL-ILP approach can provide an optimal placement solution, its scalability is limited for large/complex scenarios with many variables. It is here that heuristic schemes can be developed to provide non-optimal solutions. Now a broad range of strategies are available here, i.e., such as genetic algorithms, particle swarm optimization, simulated annealing, greedy heuristics, etc. Accordingly, the latter approach is chosen here to solve the NF placement problem for larger network sizes, as shown in Fig. 1.

Overall, the proposed heuristic scheme implements a two-stage solution. Namely, given a graph and input parameters (akin to MLL-ILP; source, destination, set of NFs and minimum link capacity), the algorithm returns the values for variables $x^i, r, d, y_d^i, l_r^{i,j}$ and $\alpha$. Foremost, the first stage of the algorithm in Fig. 1 (lines 4-15) places the NFs at datacenters to reduce deployment cost. For each function $i$ requested by $r \in R$, the scheme selects the datacenter $d_k$ that implements $i$ with the lowest setup cost (and at the same time has sufficient resources to host upcoming requests, line 8). Once a NF is placed, the resources at datacenter $d_k$ are updated accordingly as follows: if there is an instance of NF $i$ at datacenter $d_k$ with enough instance capacity to satisfy request $r$, the remaining capacity is decremented by 1. Otherwise, if there is no instance

of $i$ to serve request $r$, another instance is created at cost $c^i_{d_k}$. Either way, the available resources $w_{d_k,j}$ ($j$=1, 2, ..., $m$, where $m$ is the number of resources) at $d_k$ are reduced by $w^i_{d_k,j}$ (line 9). Each newly-created instance can serve $\lambda^i_{d_k}$-1 requests. Carefully note that serving a request does not incur any cost, i.e., only new instantiations of $i$ do. Additionally, for each new instance of $i$, the total number of instances $y^i_{d_k}$ at $d_k$ are incremented by 1 (line 10) and for each deployed NF, $x^i_{r,d_k}$ is set to 1 (line 11). To conclude NF placement (first stage), datacenter $d_k$ is also added to the subset of datacenters that serve request $r$, $D(r)$.

Meanwhile, the second stage in Fig. 1 (lines 16-32) computes the shortest path between the source and destination nodes that passes through all datacenters $d_k \in D(r)$. Initially, the algorithm connects the source node $src_r$ to the first datacenter $d_1$ of $D(r)$. In particular the constrained Dijkstra's shortest path algorithm is used here to compute the connection route (line 22). This approach first verifies whether each link has enough capacity to support the requested $b_r$. All variables $l^{i,j}_r$ along the path are then set to 1 (line 23), and datacenter $d_1$ is also added to the subset $C(r)$ (line 24). This subset defines the datacenters that serve request $r$ and are already connected to their respective neighbors within the $src_r$ and $dst_r$ path. Now if there is another datacenter $j \in D$ in the path between $src$ and $dst$, it is also added to $C(r)$ (line 25). This method avoids duplicate path computation for datacenters yet to be analyzed in upcoming iterations. Before returning to the beginning of the loop, the $src$ variable is replaced by the destination datacenter $dst$ (line 27), which was initially set to $d_k$ (line 20). At the top of the loop, $dst$ is reset to the next datacenter $d_k \in D(r)$. Hence a shortest path is computed between the previous destination and the following datacenter, i.e., in the second iteration a path is computed between $d_1$ and $d_2$. In the two final steps, a shortest path is computed between the last datacenter and the destination $dst$ (line 30), and all links $l^{i,j}_r$ within that traffic flow are set to 1 (line 31).

Akin to the STD-ILP scheme detailed at the end of Section III-A, a non-load balancing greedy heuristic scheme is also defined, i.e., $\alpha$=0. Again this solution is similar to the approach presented in [7] and is termed as the *standard routing and placement greedy* (STD-GR) heuristic.

### D. MLL-GR Complexity

The run-time complexity of the heuristic scheme is now analyzed. Conceptually, the greedy algorithm tries to lower complexity by finding the first acceptable solution. Therefore in the first stage it selects the datacenter $d$ that can implement function $i \in F_r$ at the lowest cost $c^i_d$ (line 8). Note that NF placement here is done in a serialized manner, i.e., each NF placement is independent of the following NFs requirements. Meanwhile, the second stage of the heuristic has higher complexity, i.e., since link-weighted shortest paths are computed between the datacenters and endpoints. To achieve this, the routing cost is defined as the sum of the setup cost of all links associated with a traffic flow and each respective link capacity/request load ratio, as shown in Eq. 13.

1: INPUT: G(V, E), $c^{ij} \forall (i,j) \in E, R, F, D$
2: OUTPUT: $x^i_{r,d}, y^i_d, l^{ij}_r$ values
3: set $x^i_{r,d} = 0, y^i_d = 0, l^{ij}_r = 0$ for all $r \in R, i \in F_r, d \in D, (i,j) \in E$
{BEGIN FIRST STAGE}
4: **for** all $r \in R$ **do**
5: $D(r) = \{\}$
6: $k = 1$
7: **for** all $i \in F_r$ **do**
8: $dk =$ datacenter that implements $i$ at minimum cost and has enough resources to serve an additional request
9: update resources of $d_k$
10: update $y^i_{d_k}$
11: set $x^i_{r,d_k} = 1$
12: $D(r) = D(r) \cup d_k$
13: $k = k + 1$
14: **end for**
15: **end for**
{END FIRST STAGE}
{BEGIN SECOND STAGE}
16: **for** all $r \in R$ **do**
17: $src = src_r$
18: $C(r) = \{src\}$
19: **for** $k = 1$ to $|D(r)|$ **do**
20: $dst = d_k$
21: **if** $d_k \notin C(r)$ **then**
22: $SP = constrained\_Dijkstra(src, dst)$
23: set $l^{ij}_r = 1$ for all link $(i,j) \in SP$
24: $C(r) = C(r) \cup d_k$
25: $C(r) \cup j$, for all datacenter $j \in SP, j \in D(r)$
26: **end if**
27: $src = dst$
28: **end for**
29: $dst = dst_r$
30: $SP = constrained\_Dijkstra(src, dst)$
31: set $l^{ij}_r = 1$ for all $(i,j) \in SP$
32: **end for**
{END SECOND STAGE}
33: **return** $x^i_{r,d}, y^i_d, l^{ij}_r$

Fig. 1: MLL-GR Algorithm

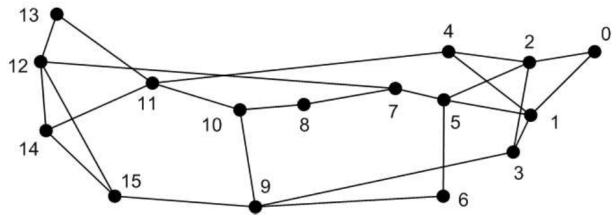
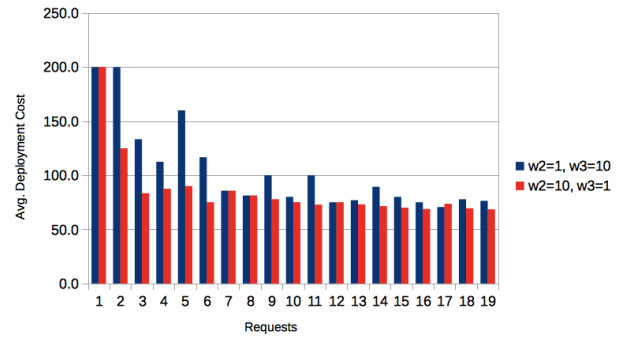


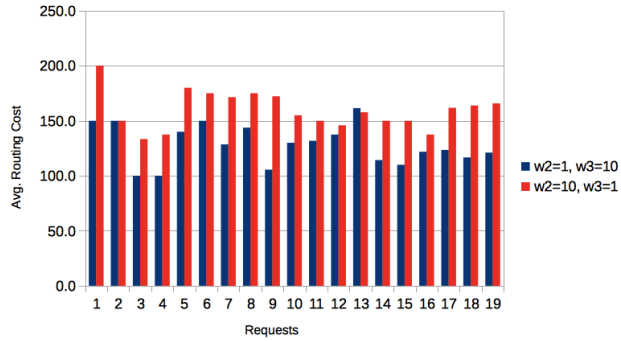

Fig. 2: NSF network topology

$$\sum_{(i,j)\in E} c^{ij} + b_r/b_{ij} \tag{13}$$

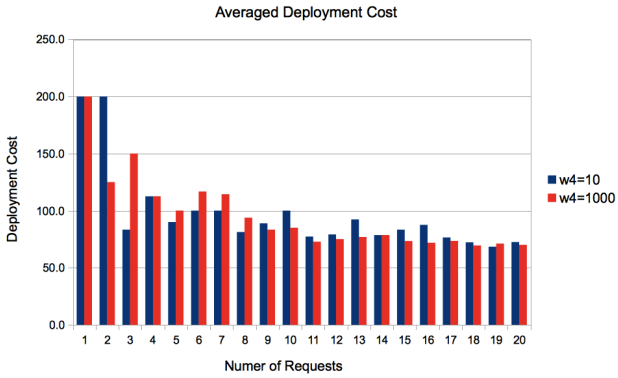
Since the constrained Dijkstra algorithm is executed within a double loop in Fig. 1, the second stage poses higher complexity than the first stage. Hence the MLL-GR run-time complexity is dominated by the second stage. In particular, this stage is invoked $|R||D|$ times, $D(r) \leq D$. Assuming a binary heap Dijkstra implementation complexity of $O(|E|log|V|)$ [10], the overall run time is of order $O(|R||D||E|log|V|)$.



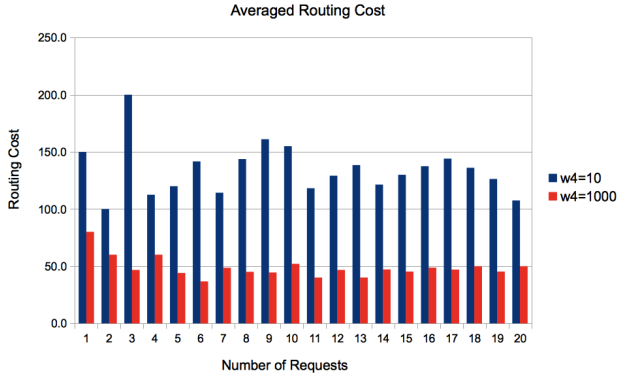

(a) Average deployment cost for differing $w_1$ and $w_2$




(b) Average routing cost for differing $w_1$ and $w_2$




(c) Average deployment cost for varying $w_4$




(d) Average routing cost for varying $w_4$

Fig. 3: First testcase (weight selection)

## IV. Performance Evaluation

In order to evaluate the proposed NF placement schemes, a realistic network topology and a set of metrics must be defined. Hence the NSF topology is chosen here, as shown in Fig. 2. This topology has 16 nodes, and it is assumed that each node is also a datacenter. Three overall testcases are also defined and tested here. In particular, the first testcase is used to perform sensitivity analysis and fine tune the weighting factors to be used in the second and third testcases. Meanwhile, the second and third testcases are used to evaluate the proposed schemes for varying physical and links resource levels. Specifically, the second testcase, termed as an *under-resourced* scenario, uses lower values for both datacenters and connectivity resources. Meanwhile, the third testcase, termed as an *highly-resourced* scenario, emulates settings with higher resources levels.

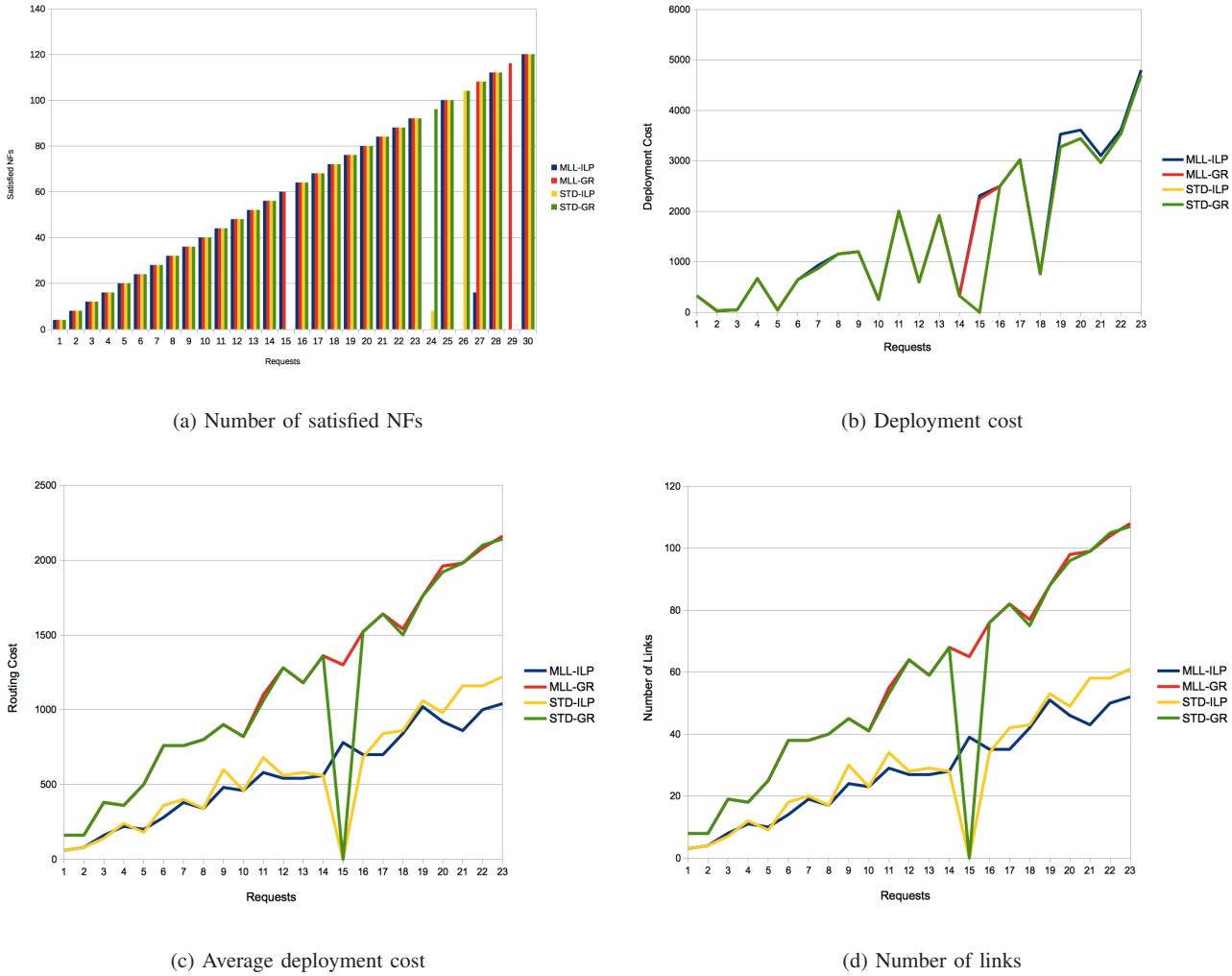
All testcases assume three different datacenter resources, i.e., $W_d$=$\{w_{d,1}, w_{d,2}, w_{d,3}\}$, representing processor, memory and storage. Namely, the first and second testcases set these values to $w_{d,1}$=$w_{d,2}$=$w_{d,3}$=500 units, respectively. Meanwhile, the third testcase sets all resource levels to 5,000 units. Similarly, in the first and second testcases, all links capacities are set to 1,000 units, for all $l_d^{i,j}$, where $(i,j)\in E$. Meanwhile, in the third testcase, these levels are increased to 10,000 units.

Finally, all other testcase values are defined as follows. First, the function set $F$ has five types, $F$=$\{f_0, f_1, ..., f_4\}$. The amount of resources $w_{d,j}^i$ required to implement a function $i\in F_d$ is also uniformly distributed between $30 \leq w_{d,j}^i \leq 70$ units. Meanwhile the setup cost of placing an instance of function $i\in F_d$ at datacenter $d\in D$ is set to $c_d^i$=50. The instance capacity $\lambda_d^i$ of a function $i$ at datacenter $d$ is also set to 2. It is also assumed that the set of functions $F_d$ implemented by a datacenter $d$ is comprised of all NFs $i\in F$. In other words, $F_d$=$F$. Finally, each request $r$ requests four NFs, namely $F_r$ is randomly selected from $F$ ($F_r\subseteq F$).

### A. Weighting Factors Selection

Proper selection of the weighting factors for each term in the objective function, Eq. 1, is crucial for effective NF placement. Therefore the first challenge is to evaluate the impact (sensitivity) of these factors in order to maximize the number of satisfied requests, $w_1$, and minimize the deployment, routing and link load costs, i.e., $w_2, w_3$ and $w_4$, respectively.

Now recall that the first term in Eq. 1 represents the number of satisfied NFs, which is typically a small value.




(a) Number of satisfied NFs

(b) Deployment cost

(c) Average deployment cost

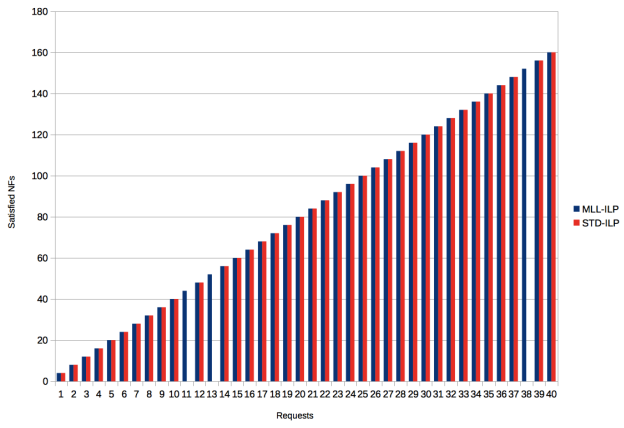
(d) Number of links

Fig. 4: Second testcase (under-resourced): request batch size 30

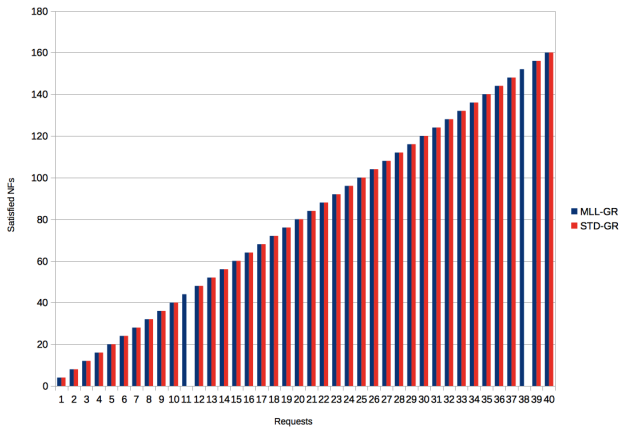
Meanwhile, the other terms represent costs with negative values (minimization). Hence it is imperative to assign a higher value to $w_1$. Therefore sensitivity analysis has to be done for the other weight values, i.e., $w_2, w_3$ and $w_4$. Now the deployment cost (second term) relates to the number of satisfied NFs $i$ and their respective costs. Hence deploying multiple NFs in the same datacenter reduces the deployment cost of each instance capacity, $\lambda_d^i$, at a datacenter. On the other hand, the routing cost (third term) and the link load cost (fourth term) are associated with the number of links determined for each traffic flow. Although both of these costs are related to the number of hops, minimizing link load cost can result in longer paths. Clearly this trade-off needs to be studied further.

In general, deployment and routing costs are associated with a number of variables, i.e., up to $|D||F|$ and $|R||E|$, respectively. Hence increasing one of these minimization term factors minimizes the associated overall cost. However this approach may also increase the other costs. Accordingly, Figs. 3a and 3b show the trade-off between deployment and routing costs. Namely, for the first testcase, Fig. 3a compares the deployment cost for different values of $w_2$ and $w_3$. Overall, a larger routing cost weight ($w_3$=10) gives increased deployment costs, whereas a larger deployment cost weight ($w_2$=10) results in lower deployment cost. Next, Fig. 3b shows that higher routing cost weights result in lower routing cost and higher deployment cost. Finally, the link load minimization term (fourth term) only has a single variable $\alpha$ (ranging from 0 to 1). Therefore $w_4$ can be used as a balancing factor to select between one of the other two minimization terms, i.e., deployment or routing. Additionally, $w_4$ introduces increased sensitivity due to its decimal value (greater range). Note that this fourth term is also directly related to the third term, i.e., routing cost, since both use variables $l_d^{ij}$ and $b_r$. Therefore minimizing $\alpha$ reduces routing cost and increases deployment cost. Furthermore, increasing $w_4$ decreases routing cost and increases deployment cost.

Next, to demonstrate the effect of varying $w_4$, an empirical methodology is deployed where two different values are used,




(a) MLL-ILP and STD-ILP




(b) MLL-GR and STD-GR

Fig. 5: Third testcase (highly-resourced): Number of satisfied NFs for request batch sizes from 1-40

i.e., $w_4$=10 and $w_4$=1,000. Tests are then performed in order to satisfy a number of requests, ranging from 1 to 20, i.e., 20 placement solutions with the values for $w_1$, $w_2$ and $w_3$ set to 1,000, 1 and 1, respectively. Here, Fig. 3c compares the average deployment cost for both values of $w_4$, i.e., computed as deployment cost/number of requests. Overall, these results show that the two costs are very similar. However Fig. 3d also shows that the average routing cost for $w$=10 is 2 to 4 times larger than the routing cost for $w_4$=1,000, i.e., computed as routing cost/number of requests.

Overall, the maximum link load variable $\alpha$ can be used to define the placement solution according to service provider needs. For example, some may prefer placing NFs to reduce deployment costs due to physical datacenter resource limitations. Meanwhile others may prefer to reduce routing cost due to network link transmission constraints, i.e., low link capacity, delay and management complexity, etc.

### *B. Under-Resourced Scenarios (Second Testcase)*

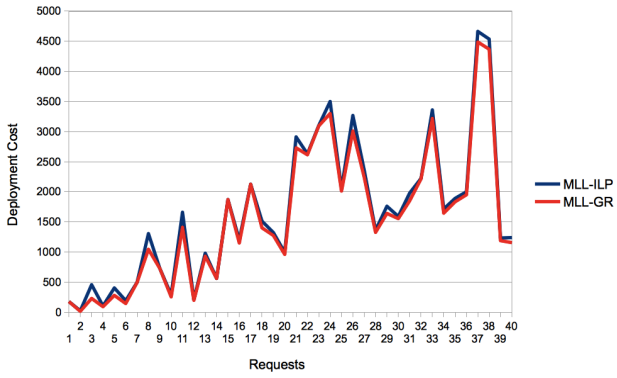
Based upon the above sensitivity analysis, the first three weighting factors are set to $w_1$=1,000 and $w_2$=$w_3$=1. All other parameters are defined as per the start of Section IV. Now the second (under-resourced) testcase assumes a total of 30 arriving requests. Fig. 4a plots the number of satisfied NFs for this scenario and shows that all four schemes begin to drop demands after 24 requests, i.e., due to resource limitations. However, note that both the STD-ILP and STD-GR schemes (yellow and green lines) fail to satisfy the $15^{th}$ request due to their inability to minimize link loads. Specifically, these non-MLL schemes do not perform load balancing. Hence for subsequent plots, results are only shown for up to 23 requests.

Now the next plot in Fig. 4b compares the deployment cost. Clearly all schemes here yield very similar values (note STD-ILP and STD-GR drop to zero for request 15 since they cannot satisfy the requested NFs). It is also noted that from request 19 and onwards, the MLL-ILP scheme (blue) and STD-ILP scheme (yellow) introduce negligibly higher deployment cost (about 5%). However in terms of routing cost there is a huge gap between the ILP (blue and yellow) and greedy heuristic schemes (red and green), as shown in Fig. 4c. Namely, the greedy heuristics yield about twice as much routing cost than the ILP solutions. Clearly the latter are more efficient.

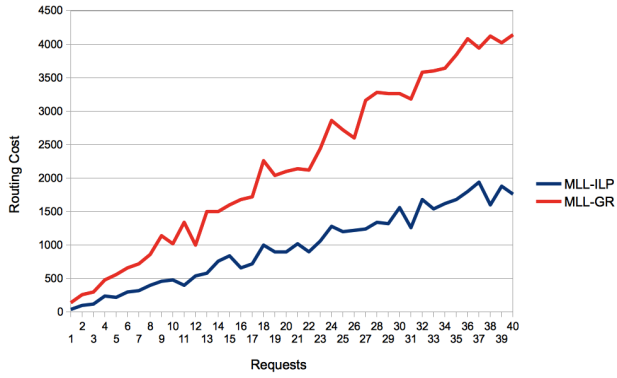
Note that the routing cost is directly related to the number of links used. Along these lines, Fig. 4d also plots the number of links associated with the deployed traffic flows. Again, there is notable similarity between the routing costs (about 100%). Additionally, for over 19 requests the MLL-ILP scheme starts to stabilize, whereas the STD-ILP scheme maintains its linear growth. Initially one may postulate that the MLL-ILP scheme should use more links than the STD-ILP scheme since it implements link load minimization. However, the MLL-ILP solution tries to minimize routing cost, which is directly related to link setup cost. Hence this scheme is more capable of minimizing routing cost and the number of links as well.

### *C. Highly-Resourced Scenarios (Third Testcase)*

Most service providers deploy highly-resourced nodes in order to achieve rapid demand scalability. Along these lines the third testcase is designed to evaluate these scenarios. Foremost, initial tests are done to determine if all four schemes are appropriate for this type of network scenario, e.g., to differentiate between the MLL and STD schemes. Namely, Fig. 5a compares the two ILP schemes with regards to the number of satisfied NFs (for request batch sizes ranging from 1-40 requests). Here it is seen that the STD-ILP scheme fails to satisfy requests 11, 13 and 38 (unlike the MLL-ILP scheme). Equivalent results for the greedy heuristics in Fig. 5b also show that the STD-GR scheme fails to satisfy the above-noted requests as well (unlike the MLL-GR scheme). These findings clearly indicate that schemes which do not minimize link load exhibit higher blocking rates (by approximately 5 to 7.5%),




(a) MLL schemes deployment costs




(b) MLL schemes routing costs

Fig. 6: Third testcase (highly-resourced): MLL costs for request batch sizes from 1-40

even in such high-resourced testcases. In light of this, further tests are only done for the MLL-based schemes.

Fig. 6a plots overall deployment costs and confirms similar performance between the MLL-ILP and MLL-GR schemes. Furthermore both schemes can satisfy up to 160 requested NFs across 40 demands, i.e., 4 NFs per request. For example, the average overall deployment costs are 1,649.3 and 1,569.2 for the MLL-ILP and MLL-GR schemes, respectively (MLL-ILP deployment cost is 5% higher than MLL-GR). However, as Fig. 6b shows, the MLL-GR scheme gives much higher routing costs, i.e., approximately 128% higher (average overall routing costs are 982 and 2,221.5 for MLL-ILP and MLL-GR, respectively). Similarly, the MLL-GR heursitic also uses an average of 111 links to establish all traffic flows, whereas the MLL-ILP scheme only uses 48.6 links, as shown in Fig. 7. Finally, additional tests (not shown) are also done with larger batch sizes of up to 60 requests. Overall findings confirm successful placement of all NFs and the same relative performance between the MLL-ILP and MLL-GR schemes.

## V. CONCLUSIONS

Network function virtualzation (NFV) is a new technology that is helping reduce network management cost and complexity. Namely, this approach decouples network services from embedded hardware devices, enabling the execution of network functions as software instances on commercial-of-the-shelf (COTS) systems. However, NFV-based deployments pose new challenges of their own, including the network function (NF) placement problem. As a result, various placement schemes have been proposed recently. However, most of these studies focus on minimizing cost and assume that there are adequate resources to satisfy all requests. Although some efforts have looked at maximizing the number of satisfied NFs, they have not considered link capacity constraints and link load balancing. To address these concerns, this paper presentes a novel integer linear programming (ILP) solution to jointly route and place NFs, termed as the *minimized link load ILP* (MLL-ILP) scheme. This strategy guarantees that a traffic flow

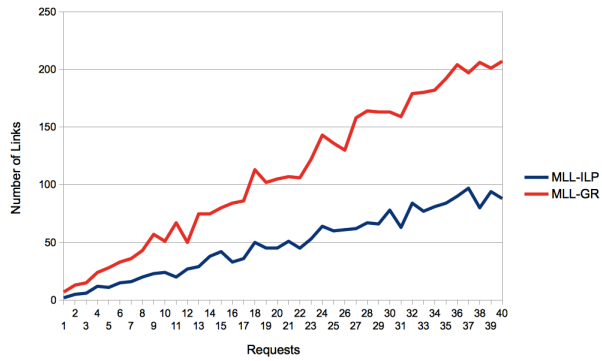



Fig. 7: Third testcase (highly-resourced): Link usage

designated to satisfy a request has adequate bandwidth capacity, and it also minimizes link load. The MLL-ILP solution also includes further provisions to allow service providers to select between deployment and routing costs. However, owing to ILP scalability concerns for larger networks, a greedy heuristic algorithm is also proposed. Both strategies are evaluated for a range of network testcases, including under-resourced and highly-resourced scenarios. Overall findings confirm that the optimization-based MLL-ILP approach gives notable routing improvements over its heuristic counterpart for both scenarios. However the deployment costs are very similar.

## REFERENCES

[1] Mijumbi, R., Serrat, J., Gorricho, J., Bouten, N., De Turck, F., Boutaba, R., "Network Function Virtualization: State-of-the-Art and Research Challenges," *IEEE Communications Surveys & Tutorials*, Vol. 8 (1), 2016.

[2] Kreutz, D., Ramos, F., Verissimo, P., Rothenberg, C., Azodolmolky, S., Uhlig, S., "Software-Defined Networking: A Comprehensive Survey," *Proceedings of the IEEE*, Vol. 103 (1), January 2015, pp. 14-76.

[3] Addis, B., Belabed, D., Bouet, M., Secci, S., "Virtual Network Functions Placement and Routing Optimization," *IEEE International Conference on Cloud Networking (CloudNet) 2015*, Niagara Falls, Canada, Oct. 2015.

[4] Cohen, R., Lewin-Eytan, L., Naor, J., Raz, D., "Near Optimal Placement of Virtual Network Functions," *IEEE International Conference on Computer Communications (INFOCOM) 2015*, Hong Kong, April 2015.

[5] Xia, M., Shirazipour, M., Zhang, Y., Green, H., Takacs, A., "Network Function Placement for NFV Chaining in Packet/Optimal Datacenters," *IEEE/OSA Journal of Lightwave Technology*, Vol. 33, Issue 8, April 2015, pp. 1565-1570.
[6] Bouet, M., Leguay, J., Conan, V., "Cost-Based Placement of Virtualized Deep Packet Inspection Functions in SDN," *IEEE Military Communications Conference (MILCOM) 2013*, San Diego, CA, November 2013.
[7] Crichigno, J., Oliveira, D., Pourvali, M., Ghani, N., Torres, D., "A Routing and Placement Scheme for Network Function Virtualization," *International Conference on Telecom. and Signal Processing (TSP) 2017*, Barcelona, Spain, July 2017.
[8] Crichigno, J., Ghani, N., Khoury, J., Shu, W., Wu, M., "Dynamic Routing Optimization in WDM networks," *IEEE Global Communications Conference (GLOBECOM) 2010*, Miami, FL, December 2010.
[9] Crichigno, J., Shu, W., Wu, M., "Throughput Optimization and Traffic Engineering in WDM Networks Considering Multiple Metrics," *IEEE International Conference on Communications (ICC) 2010*, Cape Town, South Africa, May 2010.
[10] Cormen, T., Leiserson, C., Rivest, R., Stein, C., *Introduction to Algorithms, 2nd Edition*, McGraw Hill, 2001.

**Diogo Oliveira** is a Research Assistant in the Dept. of Electrical Engineering at the University of South Florida (USF), where he is working towards his Ph.D. degree. He received his M.Sc. degree from the Federal University of Goias (UFG), Brazil, in 2009. His current research interests include sofware defined networking (SDN), network virtualization and services, disaster recovery, and design and application of optimization and meta-heuristic algorithms.

**Jorge Crichigno** received his Ph.D. in Electrical and Computer Engineering from the University of New Mexico, Albuquerque, NM, in 2009. Prior to that, he received his M.Sc. and B.Sc. in Electrical Engineering from the University of New Mexico and from the Catholic University of Asuncion respectively, in 2008 and 2004. Dr. Crichigno is currently an Associate Professor in the Department of Integrated Information Technology in the College of Engineering and Computing at the University of South Carolina, Columbia, SC. In 2016, he was a visiting professor in the Florida Center for Cybersecurity, Tampa, FL. His research interests include wireless and high-speed networks, network security and Science DMZs, and STEM education. He has served as reviewer and TPC member for journals and conferences such as IEEE Transactions on Mobile Computing and IEEE Globecom, and as panelist for NSF STEM education initiatives. He is a member of the IEEE Computer Society.

**Nasir Ghani** is a Professor in the Electrical Engineering Department at the University of South Florida and Research Liaison for the Florida Center for Cybersecurity (FC2). Earlier he was Associate Chair of the ECE Department at the University of New Mexico. He has also spent several years working at large Blue Chip organizations (IBM, Motorola, Nokia) and hi-tech startups. His research interests include cyberinfrastructure networks, cybersecurity, cloud computing, and cyber-physical systems. He has published over 200 articles, and his research has been supported by many organizations. He is also the recipient of the NSF CAREER Award and is an Associate Editor for IEEE/OSA Journal of Optical and Communications and Networking. He has also served on the editorial boards of IEEE Systems and IEEE Communications Letters. He has co-chaired many symposia for IEEE GLOBECOM, IEEE ICC, and IEEE ICCCN. He received the Ph.D. from the University of Waterloo.editorial boards of IEEE Systems and IEEE Communications Letters. He has co-chaired many symposia for IEEE GLOBECOM, IEEE ICC, and IEEE ICCCN. He received the Ph.D. from the University of Waterloo.

# Industrial IoT techniques and solutions in wood industrial manufactures

Zoltán Pödör, Attila Gludovátz, László Bacsárdi, Imre Erdei and Ferenc Nandor Janky

***Abstract*—The world of Industrial IoT (Internet of Things) generates different challenges for every company. For the small and medium-size enterprises these challenges involve material and special human resource constraints. For these enterprises, the complex solutions may not be feasible due to financial reasons, so individual solutions and individual IT support must be provided for them. Still, generally accepted architectural principles, protocols, data analysis solutions are available as building blocks for solutions tailored for smaller enterprises.**

**In order to offer good examples for such individual solutions, this paper presents various infocommunication technologies and environment as a complex infrastructure background. This setup has already been used in two different wood industrial companies to improve the effectiveness of their productivity.**

***Index Terms*—IoT, Industrial IoT techniques, sensors, energy management**

## I. INTRODUCTION

IoT, Industrial IoT, Industry 4.0, Big Data; we often meet these expressions if we look into the processes and technical implementations of the modern producing companies. In our accelerated and digitized world, the pretension is getting bigger to reduce the costs and make the production more efficient. Sensors, various IT solutions and applications can help us in this process. Today, these devices can measure everything and digitally transmit the measured data, even without the construction of a wired network. The different solutions process the collected data and the results can be used to control machines or to support managerial decisions.

Nowadays, the conception of Industry 4.0 is becoming increasingly popular [1], which is a determining phase of the industry's development; this is why the present age is labeled as the fourth industrial revolution. The information technology and the automation domains closely interweave, and the result appears in the most various production processes.

With the advent of the Industrial Internet of Things (IIoT) and the Industry 4.0 concept [2], the usage of advanced automation within various devices created the need for all sorts of sensors in significant numbers. The data provided by sensors do not only serve the above mentioned, efficiency and optimization goals, but also facilitate the decision support.

Besides its on-the-fly processing, this data have to be stored properly in order to reduce the required post processing effort. This, however leads to a big challenge in several use-cases, since it is not uncommon that sensors provide data under a few milliseconds interval that has to be transmitted, stored, analyzed and displayed appropriately.

In order to tackle such challenges, there is a need for complex physical and software infrastructures, that are capable of receiving and storing great deal of online data, and providing the opportunity to access and process data, besides generating reports and results, as well. There are many relational or non-relational database based solutions that offer efficient data handling and data storage [3]. However the SensorHUB has been developed – by BME AUT – specifically for treating different kind of sensor data while providing the afore mentioned services [4]. Clearly, there are other, commercially available solutions. Among others, these include AWS IoT by Amazon [5], Azure IoT Suite by Microsoft [6], Google Cloud IoT Core [7], ThingSpeak by MathWorks [8], ThingWorx by PTC [9], Watson Internet of Things by IBM [10], and the Arrowhead Framework [11]. The main reason we have chosen SensorHUB is its appealing flexibility, beside the close regional support.

The current paper covers industrial examples from two different Hungarian wood factories referred to as *Company 1* and *Company 2*. Furthermore, the different Industrial IoT solutions of these companies are introduced. These are already used in their live production systems.

The paper is organized as follows. In Section II, we introduce relevant and modern IIoT and Industry4.0 solutions and technologies that address the small and medium sized enterprises. These can be used for taking steps towards the main goals of Industrial IoT and Industry 4.0: create, retrieve and use important digital data connected to production and the utilization of machines). These are the basic steps to create an IoT based, digital industrial environment. Furthermore, an overview is given about the most important questions and problems related to Industrial IoT. Moreover, the main tasks, components and expectations of the Industrial Internet are introduced through a wood industrial company.

Other IIoT and Industry 4.0 related questions are discussed in Section III, through an industrial example at another company. We demonstrate a hardware and software environment, and the manufacturing process at a wood industrial company. The aim is to reduce the cost of energy used in the industrial processes. We have implemented an integrated system, through which one can collect, store, analyze data, and create executive summaries. The *mFi mPorts* and *mFi Current Sensors* of UBIQUITI [12] were used in a furniture factory to handle and process the collected current measurement data of six different machines in this factory.

In Section IV, we describe the infrastructure and technologies used for the implemented data collector system that begins with the Current sensors and ends with the SensorHUB structure and a mobile application. The focus of the

expanded structure is to determine the electricity consumption and utilization ratio of the examined machines. These results allow us to improve utilization, since the increase in electric energy consumption of a machine may indicate a mistake or a potential problem in the future.

## II. BRIEF OVERVIEW OF SOME MODERN TECHNOLOGIES IN THE INDUSTRY 4.0 CONCEPT

In this section, we demonstrate several approaches and techniques, which are related to the Industry 4.0 subject.

### *A. The main objects of Industrial IoT and Industry 4.0 concept*

Industry 4.0 means integration between individual systems in the industrial manufacturing environment and the company's other IT systems in order to manage entire value chains along with product lifecycles. An Industry 4.0 system uses modern information technologies to achieve cooperation between the integrated data-collection, data-processing and information visualization sub-systems. Industry 4.0 is about *making* things smartly, over the whole value chain – from planning through creation to the actual logistics of the manufactured product. The definitive reference model of this concept is RAMI4.0 [13].

On the other hand, Industrial IoT is about making things *work* smartly [2]. It is a cross-domain initiative, influencing various sectors, including energy, healthcare, transportation, the public sector, as well as manufacturing. It is more about industrial applications of the IoT and CPS (Cyber-Physical Systems) concepts; and does not include those value chain- and lifecycle-models that are core elements of the Industry 4.0 concept.

When IoT systems have purposeful, direct effect on the physical world – especially through modeling the world through their digital twin – and use the control information in a feedback-loop, the system is often referred to as CPS. Such systems and their physical environments are closely connected to each other.

In a manufacturing environment, the Cyber-Physical System's tasks are among the followings [14]: to find the bottleneck of the production, to determine the causes of the manufacturing faults, to change the configuration of the machines, to discover the utilization of the resources, to give information to the participants of the production, and to make a forecast of failures and optimize the working hours of the operators.

IIoT CPS systems are much more effective compared to traditional systems due to the following reasons: distributed operation, integrated software and hardware solution, which collects data and gives information in a feedback loop. By using this information, the operation of the system needs less intervention.

The tangible results delivered by the system are the following: increased productivity, better quality, quick readjustment of the parameters of the manufacturing machines, individual, special production, better resource utilization, transparency, maintainability, and less production outage.

The components of the system and their characteristics are the following. (1) Sensors and actuators: there are lots of devices and manufacturers (suppliers) and these actuators make local calculations; (2) Communication network: it contains many solutions, for example TCP/IP Ethernet, Wi-Fi, or another wireless sensor-network; (3) Information processing infrastructure: the cloud systems (mainly private solutions) are really important and popular. The earlier "dedicated server" conception is not scalable to this information infrastructure level; (4) Security sensitive environment: there are some important expectations, for example real time and fault-tolerant working.

### *B. PTP, IEEE 1588 Standard [15]*

The clocks of the computers are not accurate, and they skew even when comparing to each other. The clocks can be synchronized via the network, but it is a major challenge to set always the same and exact time on all of the clocks. The monitoring of the relatively slow processes requires accurate clock synchronization [16]. In case of energy systems, the convention is within the 1 μs error range. The special GPS-based systems stay very closely to the realization of clock synchronization with such accuracy. Unfortunately, they have some disadvantages. The solution cannot be applied in every device because of their technology and price or cost.

The typical industrial example – as one of our partner companies use as well – is periodic clock synchronization within a long time-range. This means that the machines and the computers connect to the common network, and their clocks synchronize every day.

### *C. Real-time Ethernet*

The Audio-Video Bridging standard (part of the IEEE 802.1 standard) was created in 2011. It guarantees the stable and high bandwidth, low-delay and time-synchronized voice and image transmission through the Ethernet layer. This protocol focuses on the clock synchronization and on the image and voice recording and replaying [17]. Otherwise, every client machine would record the contents based on their own clock.

At Company 1, this workflow can be adapted into the production process, as a part of the quality assurance. Now, they apply such solution, which determines the colors and faults of timber boards after the shaving process. This application evaluates the quality parameters before the subsequent manufacturing processes. We have developed and implemented this system, and this standard supports a part (real time monitoring) of our workflow in the video analysis task at the factory.

### *D. System management*

The Industrial Internet of Things concept cannot be used without an up-to-date system management solution. From the viewpoint of classical IT, the general system management tasks include the registration of machines, the monitoring of the usage of resources, capacity planning, diagnostics etc. These can be extended with the configuration and monitoring

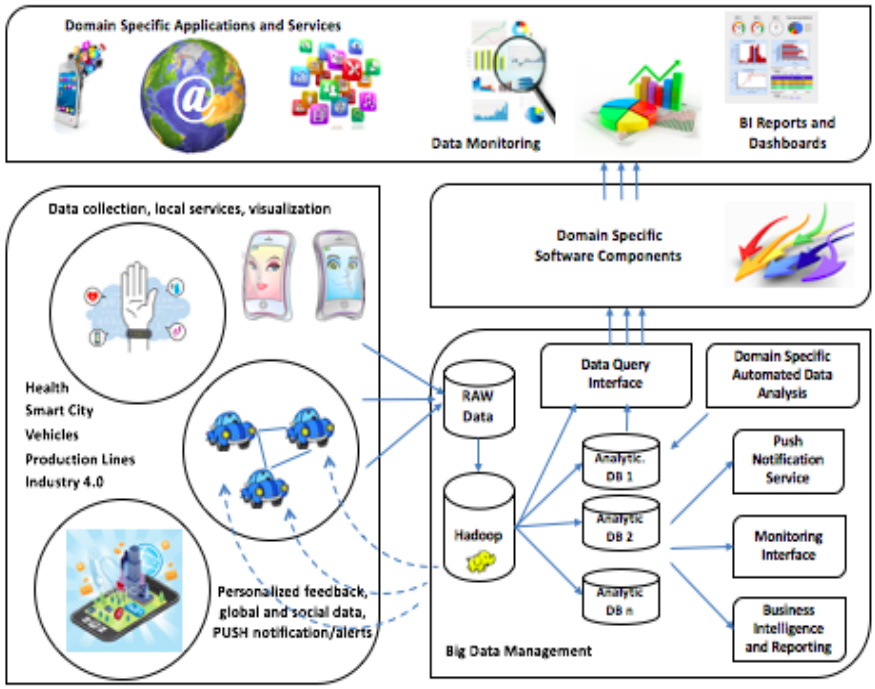



Fig. 1. The architecture of SensorHUB framework [4]

of the real-time processes. The usage of earlier solutions, such as CLI (Command Line Interface), or SNMP [18] is still necessary because of compatibility reasons, but these are not sufficient, because we want to know everything about the manufacturing. The new technological solutions' installations are imperative (Netconf [19] and Yang [22]). There are several open source solutions on the market (NeDi [21], Cacti [22], Nagios [23], etc.), which, on the other hand require wide and deep competences to use effectively.

At Company 1, the used server solutions' management applications control the system management processes (see Section III.A).

### *E. Security*

Security is a major topic in the field of Internet of Things; challenges arise at various levels [24]. Sensors have different kind of security issues than the communication architecture has, not to mention challenges at the data centers or the feedback loop. Covering these have to be taken care of, which also require various competences and continuous upgrade of procedures and tools.

The qualification conditions regarding authorization and authentication may also cause problems: the security protocols are changing and getting better, therefore, the re-qualification is necessary, because of the mentioned expectation.

IIoT systems do not require the availability of the complete system over the Internet. Moreover, the main system can be decomposed into virtual networks, which have different security-level.

At Company 1, the security-critical systems are defended with firewalls, protocols and restrictive actions. The last level machines can be reached from the outside world via the Internet. The company's intranets are working side by side in separate ways. The cloud solutions are not typical compared to other companies, especially in the area of high security expectations [25].

### *F. A special example: the SensorHUB*

The SensorHUB [4] is a distributed IoT framework developed by BME-AUT. It provides a unified architecture platform for secure storage and processing of various data. The architecture of the system is shown in Figure 1. The secure data management associated with flexible expandability creates a robust implementation that can be used in different areas of economy (automotive industry, vehicle manufacturing, healthcare, etc.).

The basis of SensorHUB is a distributed data management technology, Hadoop [26]. Contrary to the fact that traditional relational databases often result in slow processing for large data sets, distributed file system databases provide a significant improvement in processing time.

Hadoop is an open source framework that helps to store and process large amount of data in a distributed environment. The distributed environment means that in a client-server communication process the client requests need at least two servers to process them. The system provides scalable, fault-tolerant storage. The HDFS (Hadoop Distributed File System) detects and compensates the different hardware anomalies and files are decomposed into blocks, and each block is written to more than one server. Hadoop uses the MapReduce

programming model for accessing its distributed file system.

The SensorHUB contains several domain-specific applications. These applications are ready to be used, they are just waiting the necessary data from the SensorHUB. At the same time, the system provides an opportunity to the development of the user's application.

## III. THE CREATION AND INTRODUCTION OF THE PROTOTYPE SYSTEM

In this section, we describe how we applied IIoT and Industry 4.0 techniques at Company 1, which is a wood production company based in Hungary.

The 4th industrial revolution motivated the preparation of the prototype system. Machines, production devices and workpieces integrate in a huge information network, which will be the network of the other networks. Every machine and workpiece knows their own role in the system, and their needs in the given processes.

Our prototype system is an extensive data collecting, data processing, display and controlling system, which is integrated with Internet technologies into the customer company's industrial environment. This system is closely related to the physical embedded environment, furthermore, it provides information and also optimizes the processes.

### A. *The system's environment*

In order to create an effective model system, we had to know the possibilities of the company, and also the potential of the hardware and software components. Furthermore, we had to cooperate with the elements during the data collecting, data transmission, analysis and displaying.

*1) Storage: MySQL and Microsoft SQL (MSSQL) Server*

Features of data storage are: (1) various equipment's function on various internal intranet networks, (2) the data are stored on different physical servers, (3) the corporate governance system and the supervisory system automatically cooperate with different types of database management systems. The MSSQL server, which serves ERP (Enterprise Resource Planning), is situated in the international IT center (headquarter). The whole process is shown by Figure 2.

*2) Data collection and storage: Movex ERP and Movex Operator*

The services of the standard system are not always enough, if an exotic problem or task has to be solved. The company, which develops the Movex ERP, made a sub-system, which supports the production of Company 1. The Movex operator watches the performance of the production machines and archives the related data to the central database (MSSQL) with the help of data collecting sensors.

*3) Analyzing and representation: QlikView*

QlikView is a business intelligence application used to generate professional queries. The resulting business reports of the queries aid the decision making process.

*4) Data collection, representation, supervision: Vision*

During real-time monitoring, the system makes incremental data backups every 10 minutes.

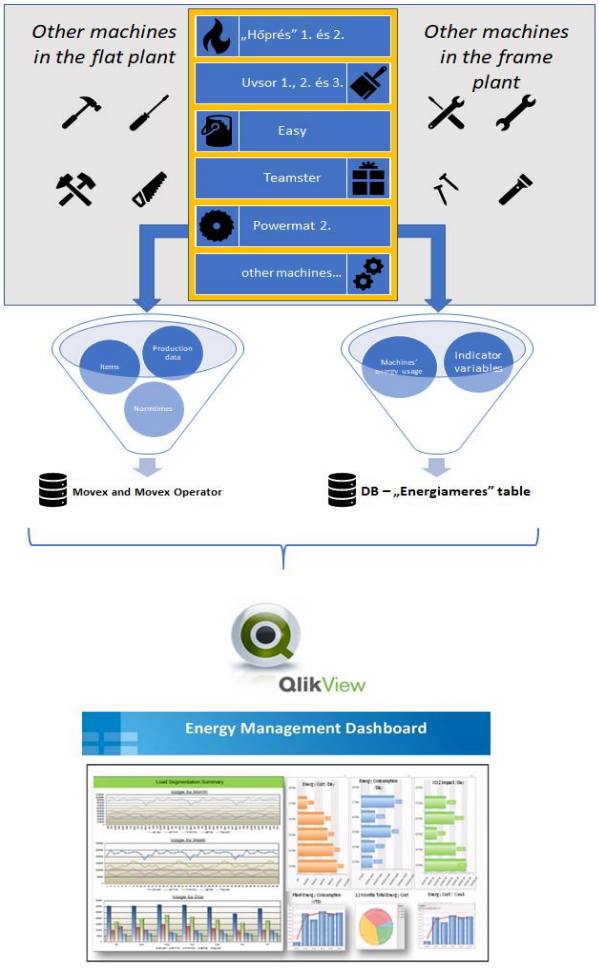


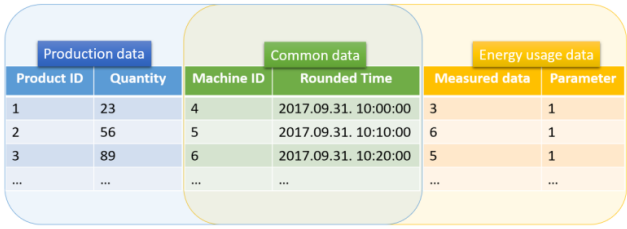

Fig. 2. The data stream from the sources to the analyzing and reporting tool

### B. *Architecture and processes: a summary*

Our main task was the interconnection of data sources (energy consumption and production data). The interconnection is based on the identity of production machines. It is important to compare historical data (e.g., data of the past shifts, hours or weeks), and based on the result of comparison, better decisions can be made. Furthermore, it is also important that we can try to predict and prevent future downtimes, problems and higher energy consumption. This concept is illustrated in Figure 2. In this figure, one can see, that we are collecting the data from different sources (manufacturing machines, environmental sensors etc.). The data is going to the company's ERP system (middle-left side of the Figure 2.) and to the database management system. Then, the initialized procedure is measuring the data from both sides.

The connection between the data is based on the common parameters: machine ID and the rounded time (see Figure 3.).

Then, based on the data analysis, a lot of questions can be answered about the manufacturing process. After the analysis, result reports are created (dashboards, tables, diagrams); this process is operating in real time. For example, we can show to the managers how much was the energy waste of a machine in the last shift. Such indication of wasting energy could mean that the machine is working without manufacturing. Thus, we can make optimizations in the operation of the factory.



Production data

| Product ID | Quantity |
|---|---|
| 1 | 23 |
| 2 | 56 |
| 3 | 89 |
| ... | ... |

Common data

| Machine ID | Rounded Time |
|---|---|
| 4 | 2017.09.31. 10:00:00 |
| 5 | 2017.09.31. 10:10:00 |
| 6 | 2017.09.31. 10:20:00 |
| ... | ... |

Energy usage data

| Measured data | Parameter |
|---|---|
| 3 | 1 |
| 6 | 1 |
| 5 | 1 |
| ... | ... |

Fig. 3. The connection between the production and the energy usage data (with several sample rows)

## IV. SENSOR NETWORK AND THE RELATED INFOCOMMUNICATION INFRASTRUCTURE

In this section, we show a detailed example about the data collection infrastructure and system which was placed at a furniture company referred as Company 2. The aim of this system was to introduce a basic Industrial IoT structure and reduce energy costs using the collected data.

There are six different analog machines in this factory. They are not suitable to generate digital data about the production or the condition of the machines. We need sensors which can be adapted into these machines to be able to provide important data and information about the operation of the machines.

Our aim was to establish a cheap, but trustworthy structure as shown in Figure 4., in order to measure the electricity consumption of the machines and to store this data for a later processing.

The UBIQUITI Networks Company produces the mFi Current sensors, which were attached to these machines, the sensors are directly connected to the mFi mPort with cable [12]. MFi mPort is an interface and monitoring function that connects to the mFi sensors. It uses special, proprietary software, the so called Controller Software to handle the sensors. These sensors can measure the current intensity in Amper and can send them to mPorts. The mPorts identify the sensors based on the ports, where the sensor is connected, because the sensors do not have their own MAC address. All sensors can have an ID, which is defined by the user. An mPort can treat only two Current Sensors, so three mPorts were utilized, because at Company 2, six mFi Current sensors were used in six different machines.

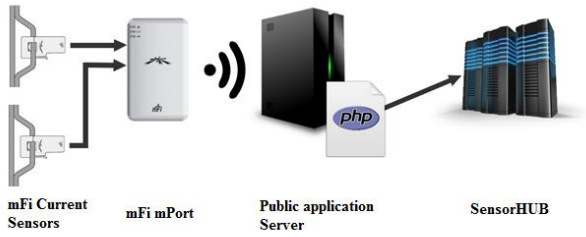



Fig. 4. The developed data collector structure for the given IIoT environment

The sensors and mPorts are placed next to the machines in the factory, but the application server is in Sopron at our University. The mPorts and the server communicate with each other through a wireless network. The mPorts have special, proprietary software (the so called Controller Software) to collect and visualize the data.

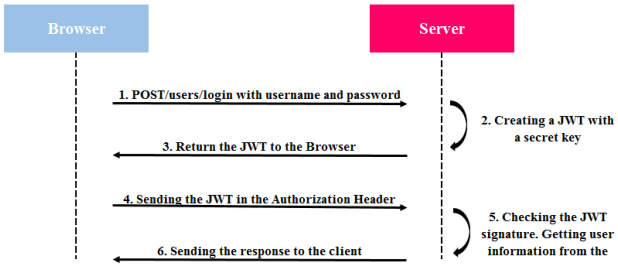
The Controller Software has limited possibilities to store the collected data; it has only an 8 Mbyte memory card, and the data cannot be directly retrieved. Still, on the Management area, we can give rules and triggers to define what the software has to do in case of an occurrence of a given event. The Controller Software's POST-URL method was used to transmit the collected data, when the sensor measures an Amper value above -1. This rule makes it possible to convey all measured Amper data to a given URL address. The POST-URL function has the following rules:

- The Controller Software automatically sends the name of the actual rule, the sensor ID, the measured data.
- The Controller Software creates automatically the HTTP header, and cannot treat cookies. This is due to the fact that the direct data sending to SensorHUB is not possible.
- The measured data cannot be directly converted to another unit, so we can only send in Amper.

Since the Controller software is not capable to directly connect to the SensorHUB, it was necessary to develop a data writer script between the Controller Software and SensorHUB. This is a PHP script, which receives the above mentioned data from the Controller Software using the POST-URL function to a given HTTP address. Then it transforms the data to JSON format, because the SensorHUB has a NoSQL based database.

Finally, the script sends the adequately prepared JSON data to the SensorHUB Upload Service based on the JSON Web Token (JWT) as it is illustrated in Figure 5 [27].

To realize the data upload using the given appID and secret key (given by SensorHUB), we need a token from SensorHUB "auth service" function, which can be used during a 24-hour time interval to send data to the SensorHUB.




Fig, 5. The JSON Web Token to send the collected data to the SensorHUB

The collected data is visualized on an Android mobile application. Earlier, personal computers were the primary tools to display data and results, but this situation has changed [28]. Due to the rapid development and spread of the smart phones and the costs of the decreasing data traffic, the smart phones overtook the desktop computers and notebooks among the average users whose primary purpose is to obtain the information.

To get the right data from SensorHUB infrastructure, it is important to give the time interval that can be selected using date-, and timepickers. Presently, the current data of 6 machines are available directly from the SensorHUB. A sample of the result is illustrated in Figure 6.

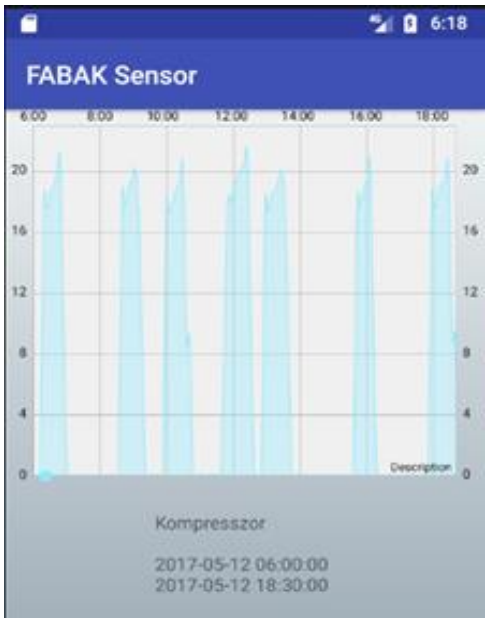



Fig. 6. Displaying IoT-related Amper data for a compressor in a Hungarian mobile application

The current form of the application offers easy and simple integration of further developments. As an example, the monitoring of values and sending threshold limit alerts in SMS, or to the given e-mail addresses, if needed.

## V. CONCLUSION

With the help of Industrial IoT applications, we can rethink the known production methods and help to achieve the main goal of the market: having bigger income with lower cost i.e. maximize the profit.

In this paper, we presented two different examples for Industrial IoT applications-related systems from the wood industry.

The first example shows us a prototype system for managing the data (collecting, storing, analyzing and representing processes) at a wood industrial company. We can continuously analyze the data with help of the prototype system. Information can be retrieved about past events, as well as about estimations on possible future events, which are provided by using forecasting procedures. Furthermore, we can focus on real-time events by defining different alarm limits, in order to prevent unexpected outages in production, or excessive energy use in the plants.

The second example introduces a simple and cheap infrastructure based on sensors and SensorHUB system, which is capable of collecting, sending, storing and processing the sensor data. The developed mobile application makes it possible to filter and visualize the collected data, and it can serve as a basis of a decision support system.

## VI. REFERENCES

[1] A. Gilchrist, Industry 4.0 - The Industrial Internet of Things. Springer, 2016, pp 250.

[2] Industrial Internet Consortium, Plattform Industrie 4.0, Architecture Alignment and Interoperability, IIC:WHT:IN3:V1.0:PB:20171205, Joint Whitepaper, 2017.

[3] Hector Garcia-Molina, Jeffrey D. Ullman, Jennifer Widom, DATABASE SYSTEMS The Complete Book, Pearson, 2009.

[4] P. Ekler, L. Lengyel, SensorHUB - An IoT Driver Framework for Supporting Sensor Networks and Data Analysis, [Online, last accessed on Sep 1, 2017] https://www.aut.bme.hu/SensorHUB/

[5] Amazon Web Services, AWS IoT, [Online, last accessed on Dec 19, 2017] https://aws.amazon.com/iot/

[6] Microsoft, Azure IoT Suite, [Online, last accessed on Dec 19, 2017] https://www.microsoft.com/en-us/internet-of-things/azure-iot-suite

[7] Google, Cloud IoT Core, [Online, last accessed on Dec 19, 2017] https://cloud.google.com/iot-core/

[8] The MathWorks, ThingSpeak - The open IoT platform with MATLAB analytics, [Online, last accessed on Dec 19, 2017] https://thingspeak.com/

[9] PTC, The ThingWorx IoT Technology Platform, [Online, last accessed on Dec 19, 2017] https://www.thingworx.com/

[10] IBM, Watson Internet of Things, [Online, last accessed on Dec 19, 2017] https://www.ibm.com/internet-of-things/spotlight/watson-iot-platform

[11] Jerker Delsing, Pal Varga, Luis Lino Ferreira, Michele Albano, Pablo Punal Pereira, Jens Eliasson, Oscar Carlsson, Hasan Derhamy, The Arrowhead Framework architecture, IoT Automation: Arrowhead Framework, ch. 3, CRC Press Publisher, 2017.

[12] UBIQUITI, mFi mPort Interface and Sensors, [Online, last accessed on Dec 19, 2017] https://www.ubnt.com/mfi/mport/

[13] Plattform Industrie 4.0, Reference Architectural Model Industrie 4.0, RAMI 4.0 - An Introduction, [Online, last accessed on Dec 19, 2017] https://www.plattform-i40.de/I40/Redaktion/EN/Downloads/Publikation/rami40-an-introduction.html

[14] J. Lee, B. Bagheri, Hung-An Kao, A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems, Manufacturing Letters, vol. 3, pp. 18-23, 2015.

[15] IEEE Standard for a Precision Clock Synchronization Protocol for for Networked Measurement and Control Systems https://standards.ieee.org/findstds/standard/1588-2008.html [Online, last accessed on Dec 18, 2017]

[16] T. Cooklev, J. Eidson, and A. Pakdaman, An implementation of IEEE 1588 over IEEE 802.11b for synchronization of wireless local area network nodes, IEEE Trans. on Instrumentation and Measurement, vol. 56, no. 5, pp. 1632-1639, October 2007.

[17] G. M. Garner, H. Ryu, Synchronization of audio/video bridging networks using IEEE 802.1AS, IEEE Communications Magazine, vol. 49, issue: 2, February 2011.

[18] D. Harrington, R. Presuhn, B. Wijnen, An Architecture for Describing Simple Network Management Protocol (SNMP) Management Frameworks, IETF RFC 3411, Dec. 2002.

[19] R. Enns, NETCONF Configuration Protocol, IETF RFC 4741, Dec. 2006.

[20] M. Björklund, YANG - A Data Modeling Language for NETCONF, IETF RFC 6020, Oct. 2010.

[21] NeDi, Network Discovery, management and monitoring, [Online, last accessed on Dec 19, 2017] http://www.nedi.ch/

[22] Cacti, The Complete RRDTool-based Graphing Solution, [Online, last accessed on Dec 19, 2017] https://www.cacti.net/

[23] Nagios, The Industry Standard In IT Infrastructure Monitoring, [Online, last accessed on Dec 19, 2017] https://www.nagios.org/

[24] Pal Varga, Sandor Plosz, Gabor Soos, Csaba Hegedus, Security Threats and Issues in Automation IoT, on Factory Communication Systems - a Conference (WFCS), 2017, 13th IEEE International Workshop on

[25] G. Kulkarni, N. Chavan, R. Chandorkar, R. Waghmare, R. Palwe, Cloud security challenges, Telecommunication Systems, Services, and Applications (TSSA), 2012, 7th International Conference on.

[26] Apache, Hadoop, [Online, last accessed on Dec 19, 2017] http://hadoop.apache.org/

[27] Introduction to JSON Web Tokens [Online, last accessed on Sep 1, 2017] / Auth0® Inc. - https://jwt.io/introduction/

[28] Mobile web browsing overtakes desktop for the first time, [Online, last accessed on Sep 1, 2017] https://www.theguardian.com/technology/2016/nov/02/mobile-web-browsing-desktop-smartphones-tablets

**Zoltán Pödör** received the M.Sc. degree in Mathematics and Computer Science from the University of Szeged in 1999. He wrote his PhD thesis on the extension opportunities of time series analysis based on special method at University of West Hungary in 2014. His research interests cover the time series analysis and data mining techniques in practice. Ha has more than 50 publications. Currently he is working as associate professor at University of Sopron.

**Attila Gludovátz** obtained M.Sc. degree in business informatics at the University of West-Hungary (NymE) in 2008. He is an assistant lecturer at the University of Sopron. He won the ÚNKP-16-3-3 New National Excellence Program in 2016. He writes his dissertation about the data collecting and its analysis. These subjects connect to the manufacturing.

**László Bacsárdi** obtained M.Sc. degree in computer engineering at Budapest University of Technology and Economics (BME) in 2006. He holds an associate professor position at the University of Sopron, where he is the Head of the Institute of Informatics and Economics. He wrote his PhD thesis on the possible connection between space communications and quantum communications at the BME Department of Telecommunications in 2012. His current research interests are in mobile ad hoc communication, quantum computing and quantum communications. He is the Secretary General of the Hungarian Astronautical Society (MANT), which is the oldest Hungarian non-profit space association founded in 1956. He is member of the board of a Hungarian scientific newspaper ('World of Nature') and he is the publisher of a non-profit Hungarian space news portal ('Space World'). Furthermore he is member of IEEE, AIAA and the HTE. He has joined the Space Generation Advisory Council (SGAC) where currently he is member of the Executive Office.

**Imre Erdei** has been studying Business Informatics at University Of Sopron since 2015. He won the ÚNKP-16-1-1 New National Excellence Program in 2016. He is part of the IoT work team which was founded in the in information technology institute where he his studies. His research interests cover the application developing and data analysis.

**Ferenc Nandor Janky** received the MSc degree in Electrical Engineering from BME, Budapest, Hungary, in 2013. He gained experienced while working for various telecommunication companies including Vodafone,AITIA International Inc. and Ericsson. His main areas of interest are network protocols, FPGA programming and software development. Ferenc is currently working as a C++ software developer for an international corporate bank.

## CALL FOR PAPERS


# INES 2018

**22nd IEEE International Conference on**

## Intelligent Engineering Systems 2018

## Las Palmas de Gran Canaria

## June 21-23, 2018

**Organizers:** *Óbuda University, Budapest, Hungary*
*University of Las Palmas de Gran Canaria, Spain*
*Elder Museum*

**Sponsored by:** *IEEE Hungary Section*
*IEEE Joint Chapter of IES and RAS, Hungary*
*IEEE SMC Chapter, Hungary*
*IEEE Computational Intelligence Chapter, Hungary*

**Technical Co-Sponsor:** *IEEE Industrial Electronics Society*

**Topics** include but not limited to:

**Artificial Intelligence in Engineering**: Reasoning, Learning, Decision Making, Knowledge Based Systems, Expert Systems

**CAD/CAM/CAE Systems**: Product Modeling, Shape Modeling, Manufacturing Process Planning

**Communications Software and Systems in Engineering**: Design Methodologies and Tools, Object-oriented, UML, Software Engineering

**Computational Intelligence in Engineering**: Machine Learning, Genetic Algorithms, Neural Nets, Fuzzy Systems, Fuzzy and Neuro-fuzzy Control

**Intelligent Manufacturing Systems**: Production Planning and Scheduling, Rapid Prototyping, Flexible Manufacturing Systems, Collaborative Engineering, Concurrent Engineering

**Intelligent Mechatronics and Robotics Systems**: Control, Perception and Recognition, Sensing and Sensor Data Fusion, Intelligent Sensors, Intelligent Motion Control, Service Robots

**Intelligent Signal Processing**

**Intelligent Transportation Systems**: Navigation Systems, On-board Systems, Real-time Traffic Control

**Man-Machine Systems**: Human Computer Interaction, Multimedia Communications, Advanced Computer Graphics, Virtual Reality

**Ontologies and Semantic Engineering**: Ontology, Thesaurus, Disambiguation, Semantic Inference, Natural Language Interaction

**Systems Engineering**: Systems Analysis, Systems Methodology, Self-Organizing Systems, Systems Integration, Large Scale Systems, Systems Simulation, Diagnosis and Performance Monitoring

**Web Engineering**: Intelligent Information Retrieval and Recommendation, Textual and Visual Interaction and Interfaces, User Behavior Analysis and Modeling, Data Modeling, Information Visualization, Knowledge Acquisition

**Submission of Papers**
The working language of the conference is **English.** Prospective authors are invited to submit full papers.

**Paper Acceptance**
Each accepted paper has to be presented by one of the authors at the conference and must be accompanied with a registration fee.

**http://www.ines-conf.org**

Full paper submission: **March 21, 2018**
Notification deadline: **April 20, 2018**
Final paper submission: **May 23, 2018**

**INES 2018 Honorary Chair**
*Lotfi A. Zadeh*, USA

**INES 2018 Honorary Committee**
*Bogdan M. Wilamowski,* USA
*Mihály Réger,* Óbuda Universiy, Budapest, Hungary
*Levente Kovács,* IEEE Hungary Section chair

**INES 2018 General Chairs**
*Imre J. Rudas,* Óbuda Universiy, Budapest, Hungary
*Alexis Quesada-Arencibia,* IUCTC, ULPGC, Spain

**INES 2018 Technical Program Committee Chairs**
*Levente Kovács*, Óbuda University, Budapest, Hungary
*Rudolf Andoga,* Technical University of Košice, Slovakia

**INES 2018 Technical Program Committee**
*Monika Bakošová,* Slovak University of Technology, Slovakia
*Valentina Balas*, "Aurel Vlaicu" University, Arad, Romania
*Ildar Batyrshin*, Mexican Petroleum Institute, Mexico
*Barnabás Bede*, DigiPen, Seattle, USA
*Attila L. Bencsik*, Óbuda University, Budapest, Hungary
*Balázs Benyó*, BME, Hungary
*György Eigner,* Óbuda University, Budapest, Hungary
*Ladislav Főző*, Technical University of Košice, Slovakia
*Gábor Hegedűs,* Óbuda University, Budapest, Hungary
*László Horváth,* Óbuda University, Budapest, Hungary
*Karel Jezernik*, University of Maribor, Slovenia
*Zsolt Csaba Johanyák*, Kecskemét College, Hungary
*George Kovács*, CAI of HAS, Hungary
*Szilveszter Kovács*, University of Miskolc, Hungary
*Krzysztof Kozlowski*, University of Poznan, Poland
*Krisztián Lamár,* Óbuda University, Budapest, Hungary
*Alajos Mészáros,* Slovak University of Technology, Bratislava, Slovakia
*Endre Pap*, University of Singidunum, Belgrade, Serbia
*Péter Pausits*, Óbuda University, Budapest, Hungary
*Béla Pátkai*, Tampere University of Technology, Finland
*Emil M. Petriu*, University of Ottawa, Canada
*Radu-Emil Precup*, „Politehnica" University of Timişoara, Romania
*Stefan Preitl*, „Politehnica" University of Timişoara, Romania
*Octavian Prostean*, „Politehnica" University of Timişoara, Romania
*Tamás Révai,* University National of Public Service, Budapest, Hungary
*Andrea Serester,* TÜV Rheiland, Hungary
*János Somló*, BME, Hungary
*Sándor Szénási,* Óbuda University, Budapest, Hungary
*Gábor Szögi,* Óbuda University, Budapest, Hungary
*Árpád Takács,* Óbuda University, Budapest, Hungary
*Pavol Tanuška,* Slovak University of Technology, Slovakia
*József K. Tar*, Óbuda University, Budapest, Hungary
*Jose Antonio Tenreiro Machado*, Institute of Engineering of Porto, Portugal
*Annamária R. Várkonyi-Kóczy*, Óbuda University, Budapest, Hungary
*Jozef Živčák,* Technical University of Košice, Slovakia

**INES Financial Chair**
*Anikó Szakál,* Óbuda University, Budapest, Hungary

**INES 2018 Organizing Committee Chair**
*Tamás Haidegger*, Óbuda University, Budapest, Hungary

**INES 2018 Organizing Committee**
*József Gáti,* Óbuda University, Budapest, Hungary
*Franciska Hegyesi,* Óbuda University, Budapest, Hungary
*Gyula Kártyás*, Óbuda University, Budapest, Hungary
*Krisztina Némethy,* Óbuda University, Budapest, Hungary

**INES 2018 Local Organizing Committee**
*Gabriel De Blasio*, IUCTC, ULPGC, Spain
*Carmelo Ruben García-Rodriguez*, IUCTC, ULPGC, Spain
*Roberto Moreno-Díaz Jr.*, IUCTC, ULPGC, Spain
*José Carlos Rodriguez-Rodriguez*, IUCTC, ULPGC, Spain

**INES Series Life Secretary General**
*Anikó Szakál*
Óbuda University, Budapest, Hungary
E-mail: szakal@uni-obuda.hu

# Guidelines for our Authors

## Format of the manuscripts

Original manuscripts and final versions of papers should be submitted in IEEE format according to the formatting instructions available on
*http://www.ieee.org/publications_standards/publications/authors/authors_journals.html#sect2,*
"Template and Instructions on How to Create Your Paper".

## Length of the manuscripts

The length of papers in the aforementioned format should be 6-8 journal pages.
Wherever appropriate, include 1-2 figures or tables per journal page.

## Paper structure

Papers should follow the standard structure, consisting of *Introduction* (the part of paper numbered by "1"), and *Conclusion* (the last numbered part) and several *Sections* in between.
The Introduction should introduce the topic, tell why the subject of the paper is important, summarize the state of the art with references to existing works and underline the main innovative results of the paper. The Introduction should conclude with outlining the structure of the paper.

## Accompanying parts

Papers should be accompanied by an *Abstract* and a few *index terms (Keywords)*. For the final version of accepted papers, please send the *short cvs* and *photos* of the authors as well.

## Authors

In the title of the paper, authors are listed in the order given in the submitted manuscript. Their full affiliations and e-mail addresses will be given in a footnote on the first page as shown in the template. No degrees or other titles of the authors are given. Memberships of IEEE, HTE and other professional societies will be indicated so please supply this information. When submitting the manuscript, one of the authors should be indicated as corresponding author providing his/her postal address, fax number and telephone number for eventual correspondence and communication with the Editorial Board.

## References

References should be listed at the end of the paper in the IEEE format, see below:

a) Last name of author or authors and first name or initials, or name of organization
b) Title of article in quotation marks
c) Title of periodical in full and set in italics
d) Volume, number, and, if available, part
e) First and last pages of article
f) Date of issue

*[11] Boggs, S.A. and Fujimoto, N., "Techniques and instrumentation for measurement of transients in gas-insulated switchgear," IEEE Transactions on Electrical Installation, vol. ET-19, no. 2, pp.87–92, April 1984.*

Format of a book reference:

*[26] Peck, R.B., Hanson, W.E., and Thornburn, T.H., Foundation Engineering, 2nd ed. New York: McGraw-Hill, 1972, pp.230–292.*

All references should be referred by the corresponding numbers in the text.

## Figures

Figures should be black-and-white, clear, and drawn by the authors. Do not use figures or pictures downloaded from the Internet. Figures and pictures should be submitted also as separate files. Captions are obligatory. Within the text, references should be made by figure numbers, e.g. "see Fig. 2."
When using figures from other printed materials, exact references and note on copyright should be included. Obtaining the copyright is the responsibility of authors.

## Contact address

Authors are requested to submit their papers electronically via the EasyChair system. The link for submission can be found on the journal's website:
www.infocommunications.hu/for-our-authors

If you have any question about the journal or the submission process, please do not hesitate to contact us via e-mail:

Rolland Vida – Editor-in-Chief:
vida@tmit.bme.hu

Árpád Huszák – Associate Editor-in-Chief:
huszak@hit.bme.hu

# IEEE GLOBECOM™ 2018

## IEEE GLOBAL COMMUNICATIONS CONFERENCE

**9 - 13 December 2018 // Abu Dhabi, United Arab Emirates**

***GATEWAY TO A CONNECTED WORLD***

# CALL FOR PAPERS

The 2018 IEEE Global Communications Conference (GLOBECOM) will feature a comprehensive technical program including numerous symposia, tutorials, workshops and an industrial program featuring prominent keynote speakers, technology and industry forums and vendor exhibits.

## TECHNICAL SYMPOSIA

We invite you to submit original technical papers in the following areas:

- Selected Areas in Communications
  - *Access Networks and Systems*
  - *Big Data*
  - *Cloud Computing, Networking and Storage*
  - *e-Health*
  - *Internet of Things*
  - *Molecular, Biological and Multi-Scale Communications*
  - *Satellite & Space Communications*
  - *Smart Grid and Power Line Communications*
  - *Social Networks*
  - *Tactile Internet*
- Ad Hoc & Sensor Networks
- Communication & Information System Security
- Cognitive Radio & Networks
- Communication QoS, Reliability & Modeling
- Communication Theory
- Communications Software, Services & Multimedia Applications
- Green Communication Systems & Networks
- Mobile & Wireless Networks
- Next-Generation Networking
- Optical Networks & Systems
- Signal Processing for Communications
- Wireless Communications

Please address questions regarding the Technical Symposia to TPC Chair Abbas Jamalipour (a.jamalipour@ieee.org) or TPC Vice Chairs Sonia Aissa (aissa@emt.inrs.ca) and Pascal Lorenz (lorenz@ieee.org). Please address questions regarding a specific symposium or a SAC symposium track to the respective symposium co-chairs or the SAC symposium track chairs listed on the conference website.

Accepted and presented technical and workshop papers will be published in the IEEE GLOBECOM 2018 Conference Proceedings and submitted to IEEE Xplore®. Author requirements of accepted paper are available on the conference website.

### TUTORIALS

Proposals are invited for half- or full-day tutorials in communications & networking. Please address questions regarding tutorials to Tutorial Co-Chairs Robert Schober (robert.schober@fau.de) & Vincent Wong (vincentw@ece.ubc.ca).

### WORKSHOPS

Proposals are invited for half- or full-day workshops on communications and networking. Please address questions regarding workshops to Workshop Co-Chairs Ying-Dar Lin (ydlin@cs.nctu.edu.tw) & Winston Seah (winston.seah@ecs.vuw.ac.nz).

## IMPORTANT DATES

**Technical Symposia Papers**
15 April 2018

**Workshop Proposals**
25 February 2018

**Tutorial Proposals**
18 March 2018

## ORGANIZING COMMITTEE

**General Chair**
Arif Al-Hammadi
Khalifa University

**General Co-Chair**
Mohammed Al-Mualla
Ministry of Education, UAE

**Executive Chair**
Bayan Sharif
Khalifa University

**Executive Vice Chair**
Khaled B. Letaief
HKUST

**TPC Chair**
Abbas Jamalipour
University of Sydney

**TPC Vice Chairs**
Sonia Aissa
Université du Québec
Pascal Lorenz
University of Haute Alsace

**Tutorial Co-Chairs**
Robert Schober
Friedrich-Alexander Universität
Vincent Wong
University of British Columbia

**Workshop Co-Chairs**
Ying-Dar Lin
National Chiao Tung University
Winston Seah
Victoria University of Wellington

**Conference Operations Chair**
Mahmoud Al-Qutayri
Khalifa University

**IF&E Chair**
Thanos Stouraitis
Khalifa University

**Patronage Chair**
Fahem Al Nuaimi
Ankabut

**Finance Chair**
Khulood Al Ali
Khalifa University

IEEE

globecom2018.ieee-globecom.org

IEEE ComSoc™
IEEE Communications Society

# SCIENTIFIC ASSOCIATION FOR INFOCOMMUNICATIONS

## Who we are

Founded in 1949, the Scientific Association for Infocommunications (formerly known as Scientific Society for Telecommunications) is a voluntary and autonomous professional society of engineers and economists, researchers and businessmen, managers and educational, regulatory and other professionals working in the fields of telecommunications, broadcasting, electronics, information and media technologies in Hungary.

Besides its 1000 individual members, the Scientific Association for Infocommunications (in Hungarian: Híradástechnikai Tudományos Egyesület, HTE) has more than 60 corporate members as well. Among them there are large companies and small-and-medium enterprises with industrial, trade, service-providing, research and development activities, as well as educational institutions and research centers.

HTE is a Sister Society of the Institute of Electrical and Electronics Engineers, Inc. (IEEE) and the IEEE Communications Society.

## What we do

HTE has a broad range of activities that aim to promote the convergence of information and communication technologies and the deployment of synergic applications and services, to broaden the knowledge and skills of our members, to facilitate the exchange of ideas and experiences, as well as to integrate and harmonize the professional opinions and standpoints derived from various group interests and market dynamics.

To achieve these goals, we…

- contribute to the analysis of technical, economic, and social questions related to our field of competence, and forward the synthesized opinion of our experts to scientific, legislative, industrial and educational organizations and institutions;
- follow the national and international trends and results related to our field of competence, foster the professional and business relations between foreign and Hungarian companies and institutes;
- organize an extensive range of lectures, seminars, debates, conferences, exhibitions, company presentations, and club events in order to transfer and deploy scientific, technical and economic knowledge and skills;
- promote professional secondary and higher education and take active part in the development of professional education, teaching and training;
- establish and maintain relations with other domestic and foreign fellow associations, IEEE sister societies;
- award prizes for outstanding scientific, educational, managerial, commercial and/or societal activities and achievements in the fields of infocommunication.

## Contact information


President: **GÁBOR MAGYAR, PhD** • *elnok@hte.hu*

Secretary-General: **ERZSÉBET BÁNKUTI** • *bankutie@ahrt.hu*

Operations Director: **PÉTER NAGY** • *nagy.peter@hte.hu*

International Affairs: **ROLLAND VIDA, PhD** • *vida@tmit.bme.hu*

Address: H-1051 Budapest, Bajcsy-Zsilinszky str. 12, HUNGARY, Room: 502

Phone: +36 1 353 1027

E-mail: *info@hte.hu*, Web: *www.hte.hu*