WIICT 2017

Proceedings of the Workshop on Innovation on Information and Communication Technologies 2017

Institute ITACA
Information and Communication Technologies

Editors:
Dr. Carlos Fernandez-Llatas
Dra. Maria Guillen

Organizing Committee

- **Chair**:Maria Guillem, Universitat Politècnica de València, Spain
- Alberto Bonastre, Universitat Politècnica de València, Spain
- Jose Manuel Catala, Universitat Politècnica de València, Spain
- Montserrat Robles, Universitat Politècnica de València, Spain
- Carlos Fernández-Llatas, Universitat Politècnica de València, Spain
- Jose Mariano Dahoui, Universitat Politècnica de València, Spain
- Isidora Navarro, Universitat Politècnica de València, Spain
- Manuel Traver Salcedo, Universitat Politècnica de València, Spain

Scientific Committee

- **Chair**: Carlos Fernandez-Llatas, Universitat Politècnica de València, Spain
- Cenk Demiroglu, Ozyegin University, Turkey
- Johan Gustav Bellika, National Center of Telemedicine, Norway
- Yigzaw Kassaye Yitbarek, University of Tromso, Norway
- Raymundo Barrales, Universidad Autónoma Metropolitana de México
- Frank Y. Li, University of Agder, Norway
- Antonio Martinez, Tecnologías para la Salud y el Bienestar S.A., Spain
- Jose Carlos Campelo, Universitat Politècnica de València, Spain
- Juan Vicente Capella, Universitat Politècnica de València, Spain
- Antonio Mocholí, Universitat Politècnica de València, Spain
- Sara Blanc, Universitat Politècnica de València, Spain
- Carlos Baraza, Universitat Politècnica de València, Spain
- Juan Carlos Ruiz, Universitat Politècnica de València, Spain
- Joaquin Gracia, Universitat Politècnica de València, Spain
- David De Andres, Universitat Politècnica de València, Spain
- Ricardo Mercado, Universitat Politècnica de València, Spain
- Miguel Rodrigo Bort, Universitat Politècnica de València, Spain
- Alejandro Liberos, Universitat Politècnica de València, Spain
- Jorge Pedrón, Universitat Politècnica de València, Spain
- Lenin Lemus, Universitat Politècnica de València, Spain
- Vicente Traver, Universitat Politècnica de València, Spain
- Pilar Sala, Universitat Politècnica de València, Spain
- Alvaro Martinez, Universitat Politècnica de València, Spain
- Jose Luis Bayo, Universitat Politècnica de València, Spain
- Lucia Aparici, Universitat Politècnica de València, Spain
- Juan Miguel García-Gomez, Universitat Politècnica de València, Spain
- Elies Fuster, Universitat Politècnica de València, Spain
- Carlos Saez, Universitat Politècnica de València, Spain
- Miguel Angel Rodriguez-Hernandez, Universitat Politècnica de València, Spain
- Canek Portillo Jimenez, Universitat Politècnica de València, Spain
- Ángel Perles, Universitat Politècnica de València, Spain
- Julian Camilo Romero-Chavarro, Universitat Politècnica de València, Spain
- Diego Boscá Tomás, Universitat Politècnica de València, Spain
- Luis Tello-Oquendo, Universitat Politècnica de València, Spain
- Israel Leyva-Mayorga, Universitat Politècnica de València, Spain
- Ángel Sanchis-Cano, Universitat Politècnica de València, Spain
- Luis José Saiz Adalid, Universitat Politècnica de València, Spain
- Gema Ibañez, Universitat Politècnica de València, Spain
- Pedro Yuste, Universitat Politècnica de València, Spain
- Daniel Gil Tomàs, Universitat Politècnica de València, Spain
- Conrado Javier Calvo Saiz, Universitat Politècnica de València, Spain

Table of Contents

A Software-Defined Networking Architecture for 5G Internet-of-Things Communication *

Luis Tello-Oquendo[1,3], Shih-Chun Lin[2], Ian F. Akyildiz[1], Vicent Pla[3], Jorge Martinez-Bauset[3] and Vicente Casares-Giner[3]

[1] Broadband Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
[2] Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27606, USA
[3] Instituto ITACA, Universitat Politècnica de València, Valencia 46022, Spain

Abstract. Internet-of-Things (IoT) is a key enabling technology for the fifth-generation (5G) wireless systems, since it supports ubiquitous information transmissions and services for a sheer number of IoT devices. In this paper, a new 5G architecture, the so-called SoftAir, is introduced to provide dynamic and flexible IoT infrastructure, and a corresponding sum-rate analysis is carried out via an optimization approach for efficient IoT transmissions. First, SoftAir decouples control and data planes for a software-defined wireless architecture and enables effective coordinations among remote radio heads (RRHs), equipped with millimeter-wave (mmWave) frontend. Next, by introducing an innovative design of software-defined gateways (SD-GWs) as local IoT controllers in SoftAir, the great diversity of IoT applications and the heterogeneity of IoT devices is easily accommodated. In particular, these SD-GWs aggregate the traffic from heterogeneous IoT devices and perform protocol conversions between radio access networks and IoT networks. Moreover, a sum-rate optimization framework is proposed in the SoftAir architecture with respect to upstream and downstream IoT communication, where the respective sum-rates are maximized and system-level constraints from (i) quality-of-service requirements of IoT transmissions; (ii) total power limit and (iii) fronthaul capacities of mmWave RRHs are guaranteed. Simulation results validate the efficacy of our solutions, where the SoftAir solution surpasses existing IoT schemes in terms of spectral efficiency and achieves optimal date rates for 5G IoT communication.

1 Introduction

Internet of Things (IoT) is an important technology for the fifth-generation (5G) wireless systems, due to its capacity to provide connectivity for anyone/anything

* This work was supported by the US National Science Foundation (NSF) under Grant No. 1547353, in part by the Ministry of Economy and Competitiveness of Spain under Grants TIN2013-47272-C2-1-R and TEC2015-71932-REDT, and by Programa de Ayudas de Investigación y Desarrollo (PAID) of the Universitat Politècnica de València. It was completed during the research stay of L. Tello-Oquendo at the Broadband Wireless Networking Lab., Georgia Institute of Technology, Atlanta, GA 30332, USA.

C.Fernandez-Llatas and M.Guillen (Ed.):
Workshop on Innovation on
Information and Communication Technologies
(ITACA-WIICT 2017)

ISBN 978-84-697-7327-7

at any time and any location. It is anticipated that there will be 29 billion connected devices by 2022 [8], and the global mobile data traffic will achieve 49 exabytes (10^{18} bytes) by 2021 [7]. However, facing this sheer number of IoT devices, it fundamentally challenges the ubiquitous information transmissions through the backbone networks, such as cellular systems. Moreover, the heterogeneity of IoT devices (range from devices with limited resources that require only intermittent connectivity for reporting, e.g., sensors, to devices that require always-on connectivity for monitoring and/or tracking, e.g., security cameras, transport fleet) and the hardware-based, inflexible cellular infrastructure impose even greater challenges to enable efficient IoT communication.

Current IoT solutions rely on low-power wide area (LPWA) networks [14], which complement traditional cellular and short-range wireless technologies in addressing IoT applications. Several technologies, such as LoRa, NB-IoT, SIGFOX, have been developed and designed solely for applications with very limited demands on throughput, reliability, or quality-of-service (QoS) [16]. However, without a central regulation among these LPWA technologies, existing IoT solutions cannot support highly diverse QoS requirements from increasing 5G IoT applications. Furthermore, due to currently fixed and hardware-based infrastructure [5], no existing work has considered the joint architectural design of IoT networks and radio access networks (RANs), and the provision of reliable and efficient upstream/downstream IoT transmissions.

In this paper, to fully address the above challenges in 5G IoT, we introduce a new architecture, the so-called SoftAir, to support flexible IoT infrastructure and propose a sum-rate optimization framework upon SoftAir to yield optimal spectral efficiency in IoT communication. Specifically, inspired by wireless software-defined networking [2, 3], we first propose the SoftAir architecture, which decouples control and data planes for open, programmable, and virtualizable wireless forwarding infrastructure. The data plane consists of software-defined RANs (SD-RANs) and software-defined core networks; the control plane has network management tools and user applications. In SD-RANs, SoftAir centralizes the communication functionalities in the baseband server (BBS) pool and enables effective coordinations among hardware-based remote radio heads (RRHs), equipped with millimeter-wave (mmWave) frontend and multiple antennas.

In addition, we propose software-defined gateways (SD-GWs) as local IoT controllers in SoftAir. SD-GWs, serving as the bridge between IoT networks and SD-RANs, aggregate the data traffic from heterogeneous IoT devices, manage and orchestrate IoT communication, and perform protocol conversions between IoT networks and SD-RANs. This innovative design of SD-GWs enables smooth and ubiquitous information transmissions, traversing between IoT and backbone networks. Moreover, upon the SoftAir architecture, we propose a sum-rate optimization framework that maximizes upstream/downstream data rates of IoT communication. Based on the physical-layer modeling of mmWave multi-input and multi-output (MIMO) transmissions, the objective is to maximize total data rates from/to SD-GWs through optimal associations of SD-GWs and mmWave RRHs and the respective precoding schemes, while guaranteeing (i) QoS requirements from diverse IoT applications, (ii) total power limit of

mmWave RRHs, and (iii) fronthaul capacity constraints between the BBS pool and mmWave RRHs.

Our main contributions are summarized as follows:

- A 5G SoftAir architecture is introduced to provide efficient, ubiquitous IoT transmissions by supporting an unified software-defined platform for both IoT networks and cellular systems.
- An innovative design of SD-GWs as local IoT controllers is proposed to orchestrate IoT devices and perform protocol conversions between IoT networks and SD-RANs.
- Upon the SoftAir architecture, a sum-rate optimization framework is proposed that achieves optimal date rates for both upstream and downstream IoT communication.

Simulation results show that our solutions outperform existing IoT infrastructure (with hardware-based architectures and fixed IoT-RAN associations), and achieves optimal rates of 100 [Mbits/s] and 430 [Mbits/s] for upstream and downstream transmissions, respectively. Regarding densely deployed IoT, the impact of increasing (fixed) number of mmWave RRHs with fixed (increasing) number of antennas are also examined with respect to the achievable sum-rates.

To the best of our knowledge, this work is the first to provide software-defined IoT infrastructure and optimal spectral efficiency for 5G IoT communication. The rest of the paper is organized as follows. Section 2 introduces the system model. Section 3 proposes the sum-rate analysis for 5G IoT via an optimization approach. Section 4 gives the numerical results, and Section 5 concludes the paper.

2 System Model

Extended from our preliminary study in [4], Fig. 1 shows the proposed 5G SoftAir architecture that supports flexible IoT infrastructure and seamless device connectivity. Specifically, it consists of three domains: sensing, network, and application. The sensing domain enables IoT devices to interact and communicate with each other, through the data collection technologies such as wireless sensor networks, RFID, ZigBee or near-field communication. The network domain consists of three parts: (i) the centralized BBS pool, which connects to the core network via backhaul links and has software-defined BSs (SD-BSs) from real-time virtualization technology for software-implemented

Notations: We use the following notations throughout this paper: boldface lower and upper case symbols represent vectors and matrices, respectively; $\mathbf{I}_x$ denotes an x by x identity matrix; $\mathbb{C}^{x,y}$ denotes the set of $x \times y$ complex matrices. The trace, transpose, and Hermitian transpose operators are denoted by $\mathrm{tr}(\cdot)$, $(\cdot)^{\mathsf{T}}$, and $(\cdot)^{\mathsf{H}}$, respectively. We use $\mathcal{CN}(\mathbf{X}, \mathbf{Y})$ to denote the circular symmetric complex Gaussian distribution with mean matrix $\mathbf{X}$ and covariance matrix $\mathbf{Y}$; the distribution of a uniform random variable is denoted by $\mathcal{U}(\cdot)$, the distribution of a normal random variable with mean x and variance σ^2 is denoted by $\mathcal{N}(x, \sigma)$, and $\sim$ stands for "distributed as". Expectation is denoted by $\mathbb{E}[\cdot]$, variance is denoted by $\mathbb{V}[\cdot]$. $\|\mathbf{x}\|$ denotes the Euclidean norm of a complex vector $\mathbf{x}$, and $|z|$ denotes the magnitude of a complex number z.

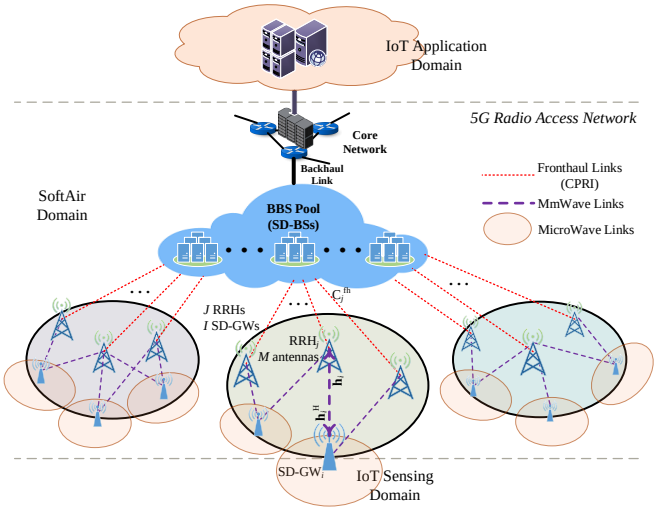


Fig. 1. A SoftAir architecture for 5G IoT communication.

baseband units (e.g., digital processing tasks), (ii) mmWave RRHs plus antennas, which are remotely controlled by SD-BSs and serve SD-GWs' transmissions, and (iii) low-latency high-bandwidth fronthaul links (fiber or microwave) using the common public radio interface (CPRI) for an accurate, high-resolution synchronization among mmWave RRHs. The objective of this network domain is to transfer the data collected from the sensing domain to the remote destination. Finally, the application domain is responsible for data processing and the provision of a wide variety of applications and services.

It is worth to note that an important component between the sensing and network domains in the SoftAir architecture is the SD-GW. In addition to alleviate high traffic loads from tremendous heterogeneous IoT devices, these SD-GWs aggregate the data from randomly deployed IoT devices and provide the Internet access to IoT networks through SD-RANs. SD-GWs also manage IoT connectivity and orchestrate IoT devices by regulating parameters in network protocols.

The network model of SoftAir consists of a SD-GW set $\mathcal{I} = \{1, \ldots, I\}$ and an associated mmWave RRH set $\mathcal{J} = \{1, \ldots, J\}$. All RRHs are connected to the BBS pool $\mathcal{B}$ via fronthaul links, where the j th fronthaul link between the $j \in \mathcal{J}$ RRH and the $b \in \mathcal{B}$ BBS has a predetermined capacity C_j^{fh}. The BBS performs most baseband processing tasks while transmission functions are realized by the RRHs using the processed baseband signals received from the BBS through the fronthaul transport network. The associations between the SD-GWs and RRHs can be determined based on the distance or channel gain from each RRH to each SD-GW. These mmWave RRHs are equipped with an array of M antennas and communicate with the single-antenna SD-GWs through mmWave links. Note that one RRH can serve a group of SD-GWs: when the j th RRH is assigned to serve the i th SD-GW, this RRH receives the SD-GW's processed baseband

signal from the BBS pool and then modifies the pre-coding vectors accordingly, as illustrated in detail in Section 3.1.

3 Sum-Rate Analysis for 5G Internet-of-Things Communication via an Optimization Approach

In the following, we first model the peculiarities of mmWave transmissions in the SoftAir architecture. Then, we formulate a sum-rate optimization framework that jointly optimizes mmWave RRHs' beamforming weights and associations between mmWave RRHs and SD-GWs for maximum upstream/downstream sum-rates, while guaranteeing QoS and system-level constraints.

3.1 Millimeter-Wave Communication

We introduce the link budget for the i th mmWave communication link between the i th SD-GW and j th RRH. Particularly we detail the path-loss l_i, channel vector $\mathbf{h}_i$, and beamforming gain G_i^{BF}, and derive the achievable upstream rate R_i^{ul} and downstream rate R_i^{dl}, respectively.

Path-Loss Considering the special characteristics of mmWave propagation (such as short-range communication, inevitable blockage effects, and sparse-scattering radio patterns) and from several research results [1,6,13], the path-loss for mmWave communication link i, l_i, can be modeled with three link-states: LoS, NLoS, or outage. We formulate the path-loss with respect to these three states for mmWave communication link i as follows:

$$l_{iL}(d_i) = (\alpha_L d_i)^{-\beta_L}; l_{iN}(d_i) = (\alpha_N d_i)^{-\beta_N}; l_{iO}(d_i) = 0, \tag{1}$$

where α_L (α_N) can be interpreted as the path-loss of the LoS (NoS) link at 1 [m] distance, and β_L (β_N) denotes the path-loss exponent of the LoS (NLoS) link. From experimental results [1,6,13], β_N value (can be up to 4) is normally higher than β_L value, i.e., 2. Then, each link-state is formulated by the channel state probabilities $\mathbb{P}_L$, $\mathbb{P}_N$, and $\mathbb{P}_O$, respectively, as follows [15]

$$\mathbb{P}_L = (1-\mathbb{P}_O)\gamma_L e^{-\delta_L d_i}; \mathbb{P}_N = (1-\mathbb{P}_O)(1-\gamma_L e^{-\delta_L d_i}); \mathbb{P}_O = \max(0, 1-\gamma_O e^{-\delta_O d_i}), \tag{2}$$

where d_i denotes the transmitter-receiver distance; the parameters γ_L (γ_O) and δ_L (δ_O) depend on both the propagation scenario and the considered carrier frequency. For computing the path-loss model, we use the values of the parameters at 73 GHz as in [12, Table I].

Channel Vector Besides the peculiarities of mmWave transmissions [1,6], the blockage information is not entirely feasible; therefore, we exploit the stochastic geometry analysis for modeling the mmWave channel vector [6]. Specifically, we model the channel vector $\mathbf{h}_i$ as follows:

$$\mathbf{h}_i = \sqrt{l_i \boldsymbol{\beta}_i}\, \boldsymbol{\xi}_i \in \mathbb{C}^{M\times 1}, \tag{3}$$

where l_i is the large-scale path-loss in power of mmWave communication link i (which might also include log-normal shadowing), $\boldsymbol{\beta}_i \in \mathbb{C}^{M\times M}$ is the covariance matrix for antenna correlations in small-scale fading, and $\boldsymbol{\xi}_i \in \mathbb{C}^{M\times 1}$ is a Gaussian vector with the distribution $\mathcal{CN}(0, \mathbf{I}_M)$ for the fast-fading. The corresponding path-loss component in (3) is modeled as follows

$$\begin{aligned} l_i = &\mathscr{I}[u < \mathbb{P}_L]\, l_{iL}(d_i) + \\ &\mathscr{I}[\mathbb{P}_L \leq u < (\mathbb{P}_L + \mathbb{P}_N)]\, l_{iN}(d_i) + \\ &\mathscr{I}[(\mathbb{P}_L + \mathbb{P}_N) \leq u \leq 1]\, l_{iO}(d_i), \end{aligned} \tag{4}$$

where $\mathscr{I}[x]$ is the indicator function, it returns 1 when x is true, and 0 otherwise; $u \sim \mathcal{U}[0, 1]$ is a uniform random variable.

Beamforming To ensure an acceptable range of communication in the multi-antenna mmWave transmissions, we introduce the precoding vectors, i.e., beamforming weights at the RRHs, where the weight vector $\mathbf{w}_i \in \mathbb{C}^{M\times 1}$ is the linear downlink beamforming vector at the j th RRH corresponding to the i th SD-GW. The beamforming gain is given as follows

$$G_i^{\mathrm{BF}} = \mathbf{w}_i^{\mathsf{H}} \boldsymbol{\beta}_i \mathbf{w}_i, \tag{5}$$

with $\boldsymbol{\beta}_i$ being the covariance matrix of the channel response vector $\mathbf{h}_i$. In the case where the fading is fully correlated between antennas, the matched filtering pre-coding method is exploited as $\boldsymbol{\beta}_i = \mathbf{h}_i^{\mathsf{H}} \mathbf{h}_i$ and $\mathbf{w}_i = \mathbf{h}_i / \|\mathbf{h}_i\|$; therefore, $G_i^{\mathrm{BF}} = \|\mathbf{h}_i\|^2$.

3.2 Achievable Sum-Rate Analysis

In the following, we first consider upstream IoT communication then the downstream case.

Upstream Transmissions (IoT Networks to SD-RANs) Following our multi-antenna mmWave transmissions characterization over a link i described previously, the received baseband signal vector $\mathbf{y} \in \mathbb{C}^M$ at the BBS $b \in \mathcal{B}$ (each element corresponds to an antenna) is

$$\mathbf{y}^{\mathrm{ul}} = \sqrt{P^{\mathrm{ul}}}\ \mathbf{H}\mathbf{x}^{\mathrm{ul}} + \eta^{\mathrm{ul}}, \tag{6}$$

where $\mathbf{H} = [\mathbf{h}_1 \cdots \mathbf{h}_I] \in \mathbb{C}^{M\times I}$, $\mathbf{h}_i \in \mathbb{C}^M$ denotes the mmWave channel corresponding to the i th SD-GW, $\mathbf{x} = [x_1 \cdots x_I]^{\mathsf{T}}$ denotes the I×1 vector containing the transmitted signals from all SD-GWs, P^{ul} is the average transmit power of each SD-GW, and $\eta^{\mathrm{ul}} \sim \mathcal{CN}(0, \sigma)$ is the zero-mean circularly symmetric Gaussian noise with noise power σ^2.

Let $\mathbf{A}$ be the M×I linear detection matrix (which depends of the channel matrix $\mathbf{H}$) used by the BBS $b \in \mathcal{B}$ to separate the received signal into user streams. The BBS processes its received signal vector and obtains the estimated

channel matrix (assuming no estimation errors) by multiplying the detection matrix with the Hermitian-transpose of the linear receiver as

$$\tilde{\mathbf{y}}^{\mathrm{ul}} = \mathbf{A}^{\mathsf{H}}\mathbf{y}^{\mathrm{ul}}. \tag{7}$$

Then, substituting (6) into (7) gives $\tilde{\mathbf{y}}^{\mathrm{ul}} = \mathbf{A}^{\mathsf{H}}\mathbf{H}\mathbf{x} + \mathbf{A}^{\mathsf{H}}\eta^{\mathrm{ul}}$. The i th element of $\tilde{\mathbf{y}}^{\mathrm{ul}}$ can be written as $\tilde{y}_i^{\mathrm{ul}} = \sqrt{P_i^{\mathrm{ul}}}\mathbf{a}_i^{\mathsf{H}}\mathbf{H}\mathbf{x} + \mathbf{a}_i^{\mathsf{H}}\eta^{\mathrm{ul}}$, where $\mathbf{a}_i$ is the i th column of $\mathbf{A}$. By the matrix multiplication, we further get $\tilde{y}_j^{\mathrm{ul}} = \sqrt{P_i^{\mathrm{ul}}}\mathbf{a}_i^{\mathsf{H}}\mathbf{h}_i x_i + \sum_{k=1,k\neq i}^{I}\sqrt{P_k^{\mathrm{ul}}}\mathbf{a}_i^{\mathsf{H}}\mathbf{h}_i x_k + \mathbf{a}_i^{\mathsf{H}}\eta^{\mathrm{ul}}$, where x_i denotes the i th element of $\mathbf{x}$ and $\mathbf{h}_i$ is the i th column of $\mathbf{H}$. Then, the signal-to-interference-plus-noise ratio (SINR) achieved by the i th SD-GW, γ_i^{ul}, is given by

$$\gamma_i^{\mathrm{ul}} = P_i^{\mathrm{ul}}|\mathbf{a}_i^{\mathsf{H}}\mathbf{h}_i|^2 / \textstyle\sum_{k=1,k\neq i}^{I} P_k^{\mathrm{ul}}|\mathbf{a}_i^{\mathsf{H}}\mathbf{h}_k|^2 + \|\mathbf{a}_i\|^2\sigma^2. \tag{8}$$

Assuming an ergodic channel [10], the achievable uplink rate of the i th SD-GW is given by

$$R_i^{\mathrm{ul}} = B\log_2(1+\gamma_i^{\mathrm{ul}}), \tag{9}$$

where B denotes the wireless transmission bandwidth. We define the uplink sum-rate [bits/s/Hz] per cell as

$$C^{\mathrm{ul}} = \textstyle\sum_{i=1}^{I} R_i^{\mathrm{ul}}. \tag{10}$$

Downstream Transmissions (SD-RANs to IoT Networks) The received baseband signal $y^{\mathrm{dl}} \in \mathbb{C}$ at the i th SD-GW is

$$y^{\mathrm{dl}} = \sqrt{P_j^{\mathrm{dl}}}\mathbf{h}_i^{\mathsf{H}}\,\mathbf{s} + \eta^{\mathrm{dl}}, \tag{11}$$

where $\mathbf{s} \in \mathbb{C}^M$ is the signal vector intended for the i th SD-GW with P_j^{dl} average power, $\eta^{\mathrm{dl}} \sim \mathcal{CN}(0,\sigma^2)$ is the receiver noise. We assume channel reciprocity, i.e., the downlink channel $\mathbf{h}_i^{\mathsf{H}}$ is the Hermitian transpose of the uplink channel $\mathbf{h}_i$. The transmit vector $\mathbf{s}$ is given as

$$\mathbf{s} = \sqrt{\upsilon}\textstyle\sum_{i=1}^{I}\mathbf{w}_i x_i^{\mathrm{dl}} = \sqrt{\upsilon}\,\mathbf{W}\,\mathbf{x}^{\mathrm{dl}}, \tag{12}$$

where $\mathbf{W} = [\mathbf{w}_1 \cdots \mathbf{w}_I] \in \mathbb{C}^{M\times I}$ is a pre-coding matrix (i.e. the network beamforming design) and $\mathbf{x}^{\mathrm{dl}} = [x_1 \cdots x_I]^{\mathsf{T}} \in \mathbb{C}^I$ contains the data symbols for the i th SD-GW. The parameter υ normalizes the average transmit power per RRH to $\mathbb{E}[\frac{P_j^{\mathrm{dl}}}{I}\mathbf{s}^{\mathsf{H}}\,\mathbf{s}] = P_j^{\mathrm{dl}}$, i.e., $\upsilon = \left(\mathbb{E}\left[\frac{1}{I}\mathrm{tr}(\mathbf{W}\,\mathbf{W}^{\mathsf{H}})\right]\right)^{-1}$.

The associated SINR γ_i^{dl}, achieved by the i th SD-GW is

$$\gamma_i^{\mathrm{dl}} = \upsilon|\mathbf{h}_i^{\mathsf{H}}\mathbf{w}_i|^2 / \textstyle\sum_{k=1,k\neq i}^{I}\upsilon\,|\mathbf{h}_i^{\mathsf{H}}\mathbf{w}_k|^2 + \sigma^2. \tag{13}$$

Since the SD-GWs do not have any channel estimate, we provide an ergodic achievable rate based on the techniques developed in [11, Theorem 1] as follows

$$R_i^{\mathrm{dl}} = B_i(1-\kappa)\log_2(1+\gamma_i^{\mathrm{dl}}), \tag{14}$$

where B_i is the bandwidth allocated to the i th SD-GW, κ accounts for the spectral efficiency loss due to signaling at RRHs. The downlink sum-rate [bits/s/Hz] per cell is

$$C^{\mathrm{dl}} = \textstyle\sum_{i=1}^{I} R_i^{\mathrm{dl}}. \tag{15}$$

3.3 Optimization Framework

We jointly optimize associations between RRHs and SD-GWs so that the SD-GW sum-rate is maximized, and the QoS requirements of SD-GWs and system-level constraints (in terms of beam-forming weights and fronthaul capacity) are satisfied simultaneously.

Besides supporting almost pervasive device connectivity, IoT applications demand services with different rate requirements; therefore, we formulate those requirements in terms of SINR coverage at SD-GWs and achieved sum-rate per cell at SD-RANs. Given ϑ as the minimum tolerable SINR over a link i, the SINR constraints of SD-GWs can be formulated as

$$\gamma_i \geq \vartheta, \forall i \in \mathcal{I}, \tag{16}$$

where γ_i is computed by (8) or (13) in case of uplink or downlink transmissions.

On the other hand, given the pre-coding vector at the j th RRH for the i th SD-GW, the transmitter power used by this RRH to serve the i th SD-GW is $\mathbf{w}_i^{\mathsf{H}}\mathbf{w}_i$ [9]. Let $P_j^{\text{r-max}}$ denote the maximum power of the j th RRH, we impose the constraints on RRHs' downlink beamforming weights as

$$\textstyle\sum_{i=1}^{I} \mathbf{w}_i^{\mathsf{H}}\mathbf{w}_i \leq P_j^{\text{r-max}}, \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{17}$$

where (17) limits the total transmit power of RRHs. Additionally, the per-fronthaul capacity constraints (neglecting the fronthaul capacity consumption for transferring compressed beamforming vector) are formulated by

$$C \leq C_j^{\text{fh}}, \forall j \in \mathcal{J} \tag{18}$$

where C is computed by (10) in uplink transmissions or by (15) in downlink transmissions.

We aim to maximize the total achievable uplink/downlink rate from/to SD-GWs, and the corresponding optimization framework is proposed as follows:

$$\begin{aligned} \text{Find} \quad & P_i^{\text{ul}}, P_j^{\text{dl}}, \mathbf{w}_i, \forall i \in \mathcal{I}, j \in \mathcal{J} \\ \text{maximize} \quad & C = \textstyle\sum_{i=1}^{I} R_i, \\ \text{subject to} \quad & \gamma_i \geq \vartheta, \textstyle\sum_{i=1}^{I} \mathbf{w}_i^{\mathsf{H}}\mathbf{w}_i \leq P_j^{\text{r-max}}, C \leq C_j^{\text{fh}}, \\ & \forall i \in \mathcal{I}, j \in \mathcal{J}, \end{aligned} \tag{19}$$

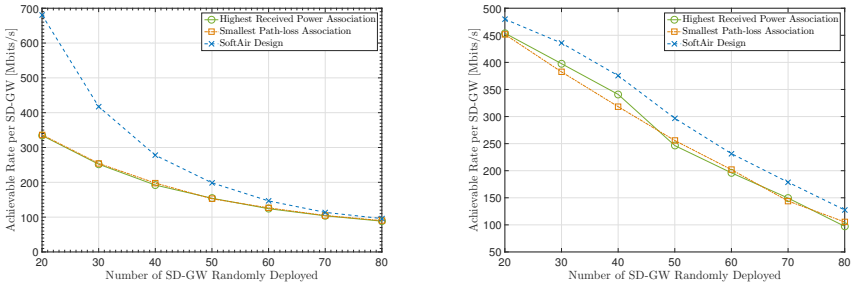
where R_i is computed by either (9) or (14) depending on uplink or downlink transmission, respectively. The optimization variables $P_i^{\text{ul}}, P_j^{\text{dl}}, \mathbf{w}_i$ take values from a discrete set that leads the optimization framework to an integer programming problem. The size of this problem allows to solve it by commercial solvers through exhaustive searching methods that yields a solution in few seconds. The acquired solutions are then processed by the BBS pool and SD-GW local controller for optimal upstream/downstream transmissions.

4 Numerical Results

In this section, we present the simulation results of our proposed designs in Section 3.3 with that of the following benchmark associations [15] for existing IoT communication:

– *Highest received power association.* In this scheme, we consider the conventional cellular network association based on uplink or downlink reference signals which undergo both path-loss and shadowing. Thus, each SD-GW will be served by the RRH providing the highest received power to it.
– *Smallest path-loss association.* In this scheme, a SD-GW will be served by the RRH with the smallest path-loss to it. This association comes to the fact that user equipments might be unable to consider random fluctuations by shadowing because the pronounced blockage impact on received signals produce less slowly-varying shadowing in mmWave transmissions.

In all experiments, each evaluation point represents the average value of 10^4 samples. We examine both the spectral efficiency and the achievable rate per SD-GW in SoftAir. Towards this, we have a set $\mathcal{J}$ of $J = 12$ RRHs, each one equipped with $M = 4$ antennas; the coverage area of every RRH has a radius of 200 [m]. They serve several SD-GW densities ranging from 30 to 80 SD-GWs randomly distributed in the coverage area of RRHs. The wireless bandwidth is set as $B = 500$ [MHz], and the carrier frequency is set as 73 [GHz]. The channel vectors are generated according to the mmWave communication modeling in Section 3.1, where the three-state path-loss model with log-normal shadowing is considered. The transmit power constraint for each RRH is $P_j^{\text{r-max}} = 45$ [dBm]. The maximum transmission power of each SD-GW is set as 23 [dBm] and the thermal noise power is assumed to be -101 [dBm/Hz]. Moreover, we assume that all RRHs possess the same fronthaul capacity, i.e., $C_j^{\text{fh}} = 6$ [bps/Hz], $\forall j \in \mathcal{J}$. As 64 QAM is set as the highest constellation supported in the systems, the maximum spectrum efficiency per data stream is 6 [bps/Hz].



(a) Upstream IoT Transmissions. (b) Downstream IoT Transmissions.

Fig. 2. Achievable rate for the SoftAir design and two existing IoT solutions with conventional mmWave schemes.

Fig. 2 depicts the uplink and downlink rate as a function of the deployed SD-GWs. The rate achieved by our solution is up to 50% higher than conventional solutions. It implies that as the number of served SD-GWs increases, both uplink and downlink rate per SD-GW decrease due to larger interference. We further consider densely deployed IoT and explore the impact of increasing number of RRHs and antennas on achievable sum-rates. In particular, the network

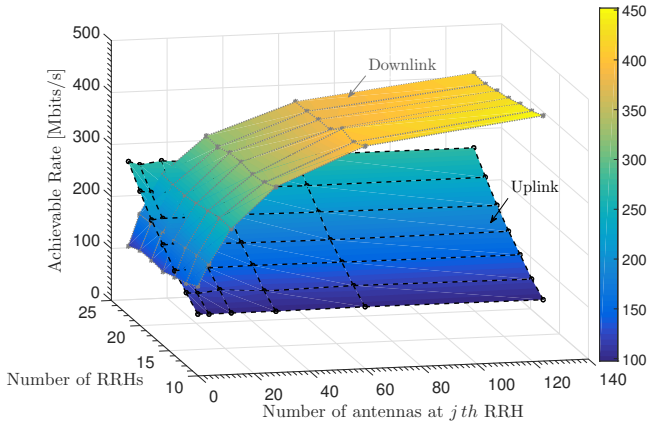


Fig. 3. Impact of increasing number of RRHs or/and antennas at RRHs on achievable sum-rates, where the served 80 SD-GWs are randomly deployed.

hereby has 12 associated RRHs with 4 antennas; each RRH serves 80 randomly deployed SD-GWs in the coverage area of the RRH. Fig. 3 indicates that either increasing the number of RRHs or increasing the number of antennas at RRHs affects the achievable rate depending on the direction of the transmission. The increasing number of RRHs significantly improves the achievable uplink rate at SD-GWs whereas the achievable downlink rate experiences small changes. On the other hand, the increasing number of antennas at RRHs greatly improves the achievable downlink rate at SD-GWs whereas the achievable uplink rate remained steady.

To sum up, our SoftAir solution provides ultrahigh data rates (i.e., at least 430 [Mbits/s] rate in downlink and 100 [Mbits/s] rate in uplink through mmWave transmissions) for each SD-GW in densely deployed scenarios, and a decision for increasing the number of RRHs or the number of antennas at RRHs can be made according to IoT applications.

5 Conclusion

In this paper, we introduced a 5G SoftAir architecture and provided optimal sum-rates for both upstream and downstream IoT communication. First, by jointly exploiting mmWave frontend, MIMO, and virtualization, the SoftAir system is proposed that gives software-defined infrastructure and enables effective coordinations among mmWave RRHs. Furthermore, SD-GWs are designed in SoftAir as local controllers that manage and orchestrate IoT transmissions between IoT networks and SD-RANs. Moreover, a sum-rate optimization framework is proposed in SoftAir, where total date rates of upstream/downstream IoT communication is maximized through optimal associations between mmWave RRHs and SD-GWs. Simulation results validate the superiority of our solutions than conventional IoT schemes, where the SoftAir solution achieves optimal spectral efficiency for 5G IoT communication.

References

1. Akdeniz, M.R., et al.: Millimeter wave channel modeling and cellular capacity evaluation. IEEE Journal on Selected Areas in Communications 32(6), 1164–1179 (June 2014)
2. Akyildiz, I.F., Lin, S.C., Wang, P.: Wireless software-defined networks (W-SDNs) and network function virtualization (NFV) for 5G cellular systems: An overview and qualitative evaluation. Computer Networks 93, Part 1, 66–79 (2015)
3. Akyildiz, I.F., Nie, S., Lin, S.C., Chandrasekaran, M.: 5G roadmap: 10 key enabling technologies. Computer Networks 106, 17–48 (2016)
4. Akyildiz, I.F., Wang, P., Lin, S.C.: SoftAir: A software defined networking architecture for 5G wireless systems. Computer Networks 85, 1–18 (2015)
5. Akyildiz, I.F., Wang, P., Lin, S.C.: Softwater: Software-defined networking for next-generation underwater communication systems. Ad Hoc Networks 46, 1–11 (2016)
6. Bai, T., Heath, R.W.: Asymptotic SINR for millimeter wave massive MIMO cellular networks. In: 2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). pp. 620–624 (June 2015)
7. Cisco: Cisco visual networking index (VNI): Global mobile data traffic forecast update, 2016-2021 (Feb 2017), http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html
8. Ericsson: Ericsson mobility report (Nov 2016), https://www.ericsson.com/mobility-report
9. Ha, V.N., Le, L.B., Dao, N.D.: Coordinated multipoint transmission design for Cloud-RANs with limited fronthaul capacity constraints. IEEE Transactions on Vehicular Technology 65(9), 7432–7447 (Sept 2016)
10. Hoydis, J., Brink, S.T., Debbah, M.: Massive MIMO in the UL/DL of cellular networks: How many antennas do we need? IEEE Journal on Selected Areas in Communications 31(2), 160–171 (Feb 2013)
11. Jose, J., Ashikhmin, A., Marzetta, T.L., Vishwanath, S.: Pilot contamination and precoding in multi-cell TDD systems. IEEE Transactions on Wireless Communications 10(8), 2640–2651 (Aug 2011)
12. Lin, S.C., Akyildiz, I.F.: Dynamic base station formation for solving NLOS problem in 5G millimeter-wave communication. In: IEEE INFOCOM 2017 (May 2017)
13. Rappaport, T.S., MacCartney, G.R., Samimi, M.K., Sun, S.: Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design. IEEE Transactions on Communications 63(9), 3029–3056 (Sept 2015)
14. Raza, U., Kulkarni, P., Sooriyabandara, M.: Low power wide area networks: An overview. IEEE Communications Surveys Tutorials PP(99) (2017)
15. Renzo, M.D.: Stochastic geometry modeling and analysis of multi-tier millimeter wave cellular networks. IEEE Transactions on Wireless Communications 14(9), 5038–5057 (Sept 2015)
16. Sinha, R.S., Wei, Y., Hwang, S.H.: A survey on LPWA technology: LoRa and NB-IoT. ICT Express (2017)

Simulation of a Communication System in Castile-La Mancha for a National Radiation Dose Data Bank

Franklin G. Jiménez-Peralta, Ángel Gómez-Sacristán, and Miguel A. Rodríguez-Hernández.

ITACA, Universitat Politècnica de València, Valencia, Spain
fraji@teleco.upv.es, agomez@upv.es, marodrig@upvnet.upv.es

Abstract. The objective of this paper is to simulate a communications network associated to a National Radiation Dose Data Bank in Castile-La Mancha (Spain). The National Radiation Dose Data Bank is being developed for the control of the amount of radiation that a patient receives in diagnostic tests or curative treatments. Using a simulation tool based on Omnet ++, which uses modules of the INET Framework, a network of convergent accesses with sources of heterogeneous IP traffic was designed and configured to define several scenarios to measure the operation of the network. This tool allows the analysis of important parameters in the network performance such as bandwidth used, packet loss, end to end delay, and jitter.

1 Introduction

Ionizing radiation is a kind of energy released by the atoms in the form of electromagnetic waves capable of causing ionization in the medium that it crosses. It is able to extract electrons of their states bound to the atom, arriving in this way to modify the state of the matter. [1]

Ionizing radiation is a type of energy that can be found in various situations, either naturally in sources such as soil, water or vegetation, or artificially by sources as nuclear power generation, medical equipment for diagnostic and therapeutic purposes.

The exposure to ionizing radiation doses could lead to several dangerous effects if this is not controlled, also the doses of low-level radiation are accumulative and could have harmful effects on people's health. However, due to its use in various fields (industry, medicine, research) one should analyze the pros and cons of its use.

One of the fields where the use of ionizing radiation has been widespread and particularly beneficial is medicine, where its applications range from the sterilization of surgical material to radiological diagnostic tests or curative treatments of diseases. Despite the benefits of the use of ionizing radiation within medicine, the control of doses received by people is necessary. The European Union in its directive 2013/59 / EURATOM decrees that since 2018 all member countries should establish requirements for equipment used in radiology should have the ability to transfer the amount of radiation information produced by the equipment during the procedure to the scanning record. [2]

C.Fernandez-Llatas and M.Guillen (Ed.):
Workshop on Innovation on
Information and Communication Technologies
(ITACA-WIICT 2017)

ISBN 978-84-697-7327-7

In a “Smart-Hospital” [3] environment, Human Type Communications (HTC) and Machine Type Communications (MTC), are sharing a common telecommunications infrastructure. Radiological equipment is a typically MTC source that generates data (text and images) in a DICOM format [8].

Smart Hospitals usually are connected to a convergent access of the Next Generation Network (NGN) in order to share data and use of applications stored in remote servers. The objective of this paper is to define a communication system which allows the transmission and exchange of radiological information of patients between each of the health centers and the National Dosimetry Center. This system called "National Dose Bank to Patients”, will allow a record of the radiation doses that patients receive throughout their lives.

This paper focuses on the the optimization of the communications scenario of a Smart Hospital using a simulation tool based on Omnet ++ [4] which is specifically designed to analyze convergent access networks with heterogeneous IP traffic sources. Simulation results will be able to analyze important parameters in the performance of the network and thus evaluate the quality of communication.

The structure of this article is as follows: Section 2 shows the description of the traffic sources with the source size to use. Section 3 describes the scenario simulation and results, and finally, Section 4 contains the conclusions and future work.

2 Traffic Sources

Based on the needs of the "National Patient Dose Data Bank" project, we worked with the Machine Type Communications (MTC) category, which is related to the transfer of information, medical histories, medical images and patient monitoring.

Within this category, a degree of priority was assigned depending on how critical the service is within the network. In this way and taking as reference MacroLAN service [5], which divides the traffic into four categories: (multimedia, gold, silver and management category), it was assigned the category silver for data traffic since it will be a traffic of lower priority, where the delay is secondary, because it is necessary to store a patient's medical history, but not immediately.

The parameters used to define the behavior of data sources are: the size and number of files and the way of sending that information.

The modeling of MTC sources is done by sending files through transmission windows or distributed sessions throughout the simulation period. In each window one or more information files are sent simulating images and/or text resources, the profiles of a data source are shown in Fig.2.

The configurable parameters will be: the number of sessions to be generated, the number of files in each session, the separation between files, as well as the size of each file to send, each of these parameters can take fixed values or variables through a statistical distribution, the configuration is done through XML (Extended Markup Language) files.

Fig.1. Profile of a data source.

Within the parameters to be configured in a data source the size and type of text resource that is generated is important because these profiles together with the number of equipment and the retransmission time will define the amount of traffic in the network.

2.1 Source size

For simulating the size of the information transmitted from the radiological equipment, two types of files have been chosen, which will contain all the information associated with a patient, necessary to control his radiological activity.

The first case with a file of a size 1000 bytes corresponds to a frame containing several fields such as: patient identification information, radiation dose received, type of observation made, among others [6].

The second file to be used will have a size of 160 KBytes, corresponds to the frame size of the system called "Dosewatch" [7], which is a radiation dose management solution designed to collect and analyze automatically irradiated patients.

This frame apart from containing the radiological information of the patients, includes a compressed image which can be used for medical review but not for diagnosis. It also works under the DICOM standard (Digital Imaging and Communication in Medicine) [8], which allows the transmission of medical images in digital format.

3 Scenario simulation and results

The first phase of the project was carried out for the region of Castile-La Mancha, for the configuration of sources were provided real data provided by the National Dosimetry Center on the amount and distribution of radiological equipment. In this phase, the scenario is characterized by a health center that aggregates all the radiological devices corresponding to each province of the Castile-La Mancha region. Each radiological device generates data traffic by sending the files with the medical information associated with a patient.

The region of Castile-La Mancha has 5 provinces: Toledo, Ciudad Real, Cuenca, Guadalajara and Albacete, the scheme of the network in Omnet ++ is shown in Figure 2.

The scheme of the Smart Hospital network of this project is composed of several modules: radiological center, MPLS cloud and router. Radiological center acts as traffic generator, it can characterize heterogeneous traffic sources, in this project there are nine data traffic corresponding to nine radiological equipment used in health centers. MPLS cloud performs traffic engineering, simulating the process from the traffic generation in the radiological center until traffic reaches the private cloud, using multiprotocol Label Switching (MPLS). Router is mainly used to perform QoS quality of service functions at the entrance of the private cloud which is the site where all network traffic is joined.

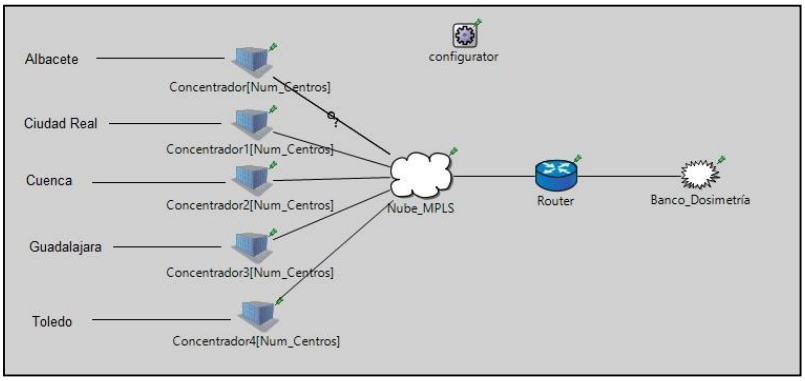


Fig. 2. Topology of the communications network for the region of Castile-La Mancha implemented in Omnet ++.

For the characterization of the generating sources, we have the information shown in Table 3, corresponding to the nine types of radiological equipment and their respective distribution in each province of Castile-La Mancha.

Table 1. Distribution of equipment in Castile-La Mancha. [9]

Radiological Equipment	**Albacete**	**Ciudad Real**	**Cuenca**	**Guadalajara**	**Toledo**	**Total**
Grafía	17	16	7	6	21	67
Grafía and Escopía	5	6	2	2	7	22
Computed Tomography	6	7	3	3	8	27
Mammography	6	6	2	2	7	23
Orthopatographs	2	2	2	2	3	11
Dental	17	38	11	9	23	98
Portable	6	7	3	4	13	33
Arco	17	14	4	10	27	72
Densitometers	2	0	0	0	1	3
Total	78	96	34	38	110	356

Once the traffic generating sources have been defined, the next step is to describe the quality of service (QoS) parameters taken into account which are: queue management and packet marking, each of these parameters is described below.

As for queue management, the silver traffic category is assigned a large queue size (700 packets), with the objective of minimizing packet loss, also assigning green and yellow frame dropping probabilities and a maximum queue size to discard all green frames.

The marking of IP packets is done according to the following parameters, which will be contracted with the carrier: CIR (Commited Information Rate) or average long-term traffic rate guaranteed by the operator; CBS (Commited Burst Size) or burst size compromised, relative to CIR; PIR (Peak Information Rate) o maximum traffic rate guaranteed by the operator, can never be greater than the capacity provided by the operator; PBS (Peak Burst Size) or burst size allowed, relative to the PIR.

For the treatment of the yellow frames, in this case as all traffic will be classified with silver quality, if the volume is greater than the PIR all the traffic will be discarded.

The following is an example of the configuration parameters of the router at the exit of the radiological center.

Table 2. QoS values in router radiological center.

Traffic Type	**CIR**	**CBS**	**PIR**	**PBS**
Silver	20Mbps	2.5MB	100Mbps	12.5MB

Once the QoS parameters are defined, the general configuration of the scenario is completed. A simulation period of 60 minutes is defined in permanent regime, the MPLS delay is calculated using time values given by network carrier between the National Dosimetry Center and the provinces. Finally, the analysis of both individual and global bandwidth, analysis of packet loss and delay in the network is performed.

After defining the applications and the general configuration, the radiological centers are configured, the scenario will have five centers, one for each province of the Castile-La Mancha region, whose output link will be 100Mbps.

Figures 3 to 6 show different results that reflect the behavior of the network, important parameters such as: bandwidth consumption, packet loss analysis, end-to-end delay and jitter.

First, we show the results in terms of the bandwidth required in the link for the province of Toledo which has a greater number of sources, then shows the global or input bandwidth in the private cloud, and finally, extreme to extreme delay and jitter.

All these first results correspond to the 160KBytes file. Figure 3 shows the bandwidth consumed in the province of Toledo, it is around 5Mbps.

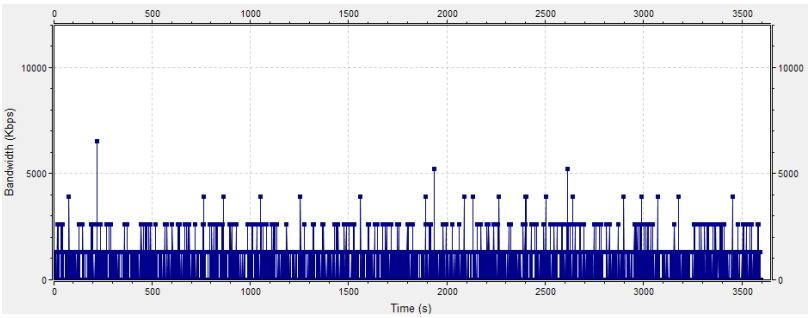


Fig.3. Bandwidth concentrator of Toledo.

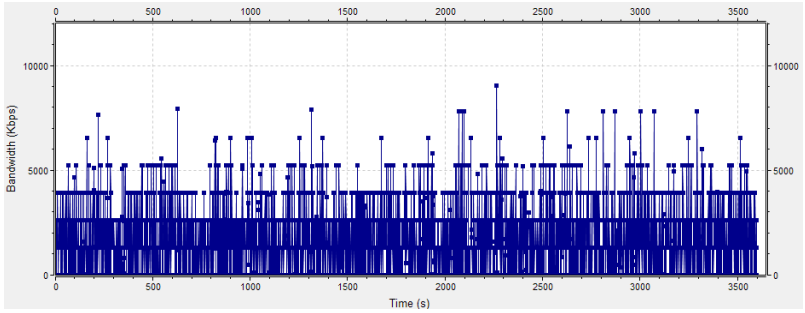


Fig.4. Traffic to the entrance of the servers (Valencia).

As shown in Fig. 4, the bandwidth required to avoid congestion at the entrance to the servers is approximately 9Mbps, so sending such information will not lead to congestion in the network.

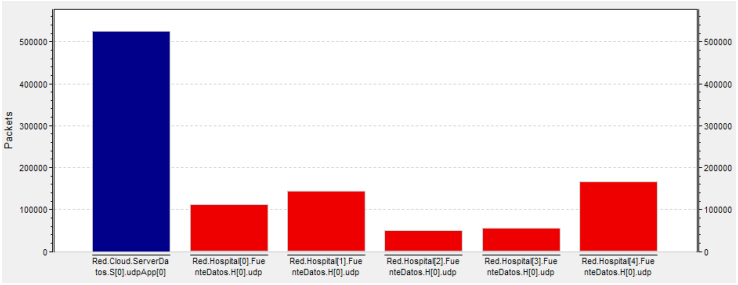


Fig.5. Packages sent and received by the servers (Valencia - CND).

Figure 5 shows the number of packets sent and received, the bars in red represent the packets sent, the bar in blue represents the packets received, in the same way it is observed that there is no loss of information.

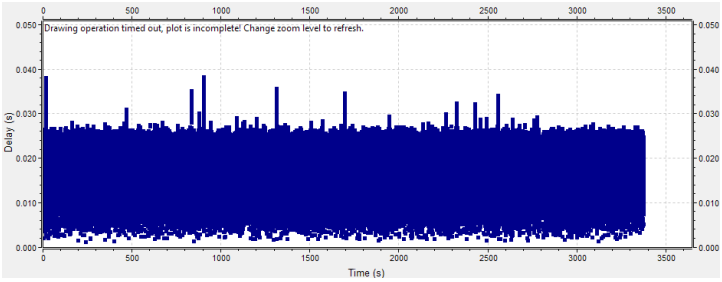


Fig.6. Delay in data traffic.

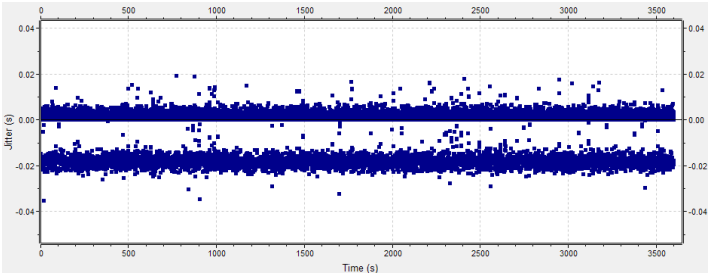


Fig.7. Jitter in data traffic.

Finally, the results of packet delay in the network transmission process and jitter or delay variations are presented (Fig. 6 and 7). The results present low values due to a scenario without congestion.

4 Conclusions and future work

The design and simulation of the communication system associated to a National Radiation Dose Data Bank was done. This system will allow the control of the amount of radiation that a patient receives in diagnostic tests or curative treatments.

The traffic graphs show that the impact of the communication networks of each center will not be high, but due to the number of radiological studies carried out daily in all regions of Spain, an analysis by means of the simulation of the communications network is necessary.

The simulation time for the scenario was 60 minutes, with random file submissions every 5 minutes, considering an average case of patient care where radiological observations are performed in the morning or at a certain time of each day.

The architectural model developed allows the simulation of other types of traffic as transmission of DICOM images for diagnosis, or national epidemiological studies. On the other hand, a database with all the information of the radiological inspection in Spain opens many opportunities to improve the attendance to the patients within the concept of Smart Hospital.

As a future work, the analysis of the network could be carried out at the national level, for which only real data are required by the National Dosimetry Center on parameters such as: number and distribution by province of the radiological centers and the number and distribution of equipment, since the network at the national level is already designed, in addition analysis is expected with the addition of other types of traffic, such as voice and data traffic.

References

1. World Health Organization - Ionizing Radiation, http://www.who.int/mediacentre/factsheets/fs371/es/
2. European Union. Council Directive 2013/59/EURATOM, December 5th 2013.
3. Angel Gomez-Sacristan, Miguel A. Rodriguez-Hernandez, and Victor Sempere, Evaluation of Quality of Service in Smart-Hospital Communications, Journal of *Medical Imaging and Health Informatics, vol. 5, pp. 1864-1869, 2015*.
4. OMNeT++ User Manual Version 4.4.1, http://www.omnetpp.org/
5. Telefónica, Macrolan Service. Available in: http://www.movistar.es/grandes-empresas/soluciones (2014).
6. Rodriguez-Hernandez, M. A., Rodrigo-Boix, I., Vilar-Palop, J., & Llorca-Domaica, N. (2017, March). Preliminary study of the communications system for a National Bank of Patients Doses. In *Global Medical Engineering Physics Exchanges/Pan American Health Care Exchanges (GMEPE/PAHCE)*, 2017 (pp. 1-4). IEEE.
7. DoseWatch, http://www3.gehealthcare.com/en/products/dose_management/dosewatch
8. DICOM Standard, http://dicom.nema.org/
9. Activity Report of the National Dosimetry Center year 2015, 2015.
10. Metro Ethernet, https://sites.google.com/site/3cuelectronica/home/ethernet/metroethernet
11. David Gomez Cuadrado, *Simulation of All-IP access network with configurable traffic sources*. Polytechnic University of Valencia, Spain, 2014.
12. A. Gomez-Sacristan, M. A. Rodriguez-Hernandez, V. Sempere, "Telecom Design Services in Smart-Hospital Communications", *2016 Global Medical Engineering Physics Exchanges/Pan American Health Care Exchanges (GMEPE/PAHCE)*, pp. 133: 1-6, 2016.
13. International Telecommunication Union, ITU-T Recommendation Y.2001: Next generation Networks General Overview (2004).
14. International Telecommunication Union, ITU-T Recommendation Y.2112: A QoS control architecture for Ethernet-based IP access networks (2007).
15. International Telecommunication Union, ITU-T Recommendation Y.2113: Ethernet QoS control for next generation networks (2009).
16. Cisco Systems Inc., 'Cisco Catalyst 3650 Configuration Guide' (http://www.cisco.com/c/en/us/support/switches/catalyst-3650-seriesswitches/products-installation-and-configuration-guides-list.html) (2014).

Collision handling in the LTE-A random access procedure: common assumptions and their impact on performance

Israel Leyva-Mayorga, Luis Tello-Oquendo, Vicent Pla and Jorge Martinez-Bauset

ITACA, Universitat Politècnica de València, Valencia, Spain
{isleyma, luiteloq, vpla, jmartinez}@upv.es

Abstract. Wireless communication systems are evolving at an incredibly fast pace. Cellular networks are a clear example of this fast evolution, as 5G is expected to be fully standardized and operative within the next few years, which is expected to unleash the true potential of massive machine-to-machine (M2M) communications and the Internet of Things (IoT). The current LTE-A (4G) cellular system will, most likely, serve as a base for the development of the future 5G cellular system. As a consequence, the performance evaluation of the random access (RA) procedure of LTE-A under massive machine-to-machine (M2M) communication scenarios is a hot research topic. Specifically, identifying the limitations of the RA procedure and of the Random Access Channel (RACH) is of utmost importance in order to attain ultra-reliable M2M communications in 5G. Nevertheless, there is still controversy among the research community on the correct modeling of the RA procedure. For instance, there is no consensus on the manner in which the cellular base station (eNB) reacts to collisions; i.e., simultaneous transmissions of the same preamble (access codeword). In this study, we describe the two different outcomes and the two major assumptions regarding the collision handling in the RA procedure (under each assumption only one of the two outcomes is possible). Then, we compare the performance under these two assumptions. Results show that the selected assumption highly impacts the performance of the RA procedure. However, each one of the two possible outcomes may occur with a certain probability in a real implementation. Therefore, the results obtained under each assumptions may serve as lower and upper bounds to the real performance of the network.

1 Introduction

In the near future, the deployment of 5G networks will motivate a technological revolution in which everything is connected. As such, great efforts are being made to develop protocols, architectures and communication systems that are capable of achieving ubiquitous and ultra-reliable communications between a massive number of wireless devices. In this regard, current systems, such as 4G

C.Fernandez-Llatas and M.Guillen (Ed.):
Workshop on Innovation on
Information and Communication Technologies
(ITACA-WIICT 2017)

ISBN 978-84-697-7327-7

cellular networks (LTE-A) will serve as a base for the future development of the Internet of things (IoT) [1, 2]. Specifically, the current LTE-A system has a widely deployed infrastructure, and is capable of providing with ubiquitous coverage along with global connectivity [3]. Because of this, LTE-A networks present the best solution to date for the interconnection of mobile devices (known as user equipments, UEs, in LTE-A).

In LTE-A, the UEs access the cellular base station (eNB) by means of the random access (RA) procedure; it is performed through the random access channel (RACH) and comprises a four-message handshake: preamble transmission (only allowed in predefined time-frequency resources called random access opportunities, RAOs), random access response (RAR), connection request and contention resolution messages. Each of the steps that comprise the RA procedure are explained in detail in Section 2.

The RA procedure of LTE-A was designed to handle human-to-human (H2H) traffic, which is characterized by relatively infrequent access requests and large amounts of transmitted data. On the other hand, machine-to-machine (M2M) applications, in which the devices communicate autonomously [1, 2, 4], are characterized by a large number of synchronized access requests and small amounts of transmitted data. The RA procedure is not efficient at handling such a high number of synchronized access requests. Consequently, M2M applications may lead to severe congestion in the RACH. Furthermore, due to the rapid increase in the number of interconnected devices, the frequency and severity of congestion will surely increase in the coming years. As a consequence, the correct performance evaluation of the LTE-A RA procedure is of prime importance.

Despite the extensive literature related to the modeling [5, 6] and performance evaluation of the RA procedure under M2M scenarios [7–9], there is still no consensus on the behavior of the system in the presence of collisions. For instance, it is clear that collisions occur whenever multiple UEs transmit the same preamble simultaneously. However, the eNB may either detect the collision immediately or continue with the RA procedure. The possible causes for these two different outcomes and their implications in the reminder of the RA procedure are described in detail in Section 2.

Based on the two possible outcomes during the RA procedure, two main assumptions are made in the literature. In the first one, every time a preamble collision occurs, the eNB does not decode the transmitted preamble (*Msg1*); hereafter, we refer to this as the *CMsg1* assumption. In the second one, the eNB correctly decodes the preamble and the RA procedure continues until the transmission of the connection request message (*Msg3*); hereafter, we refer to this as the *CMsg3* assumption.

The *CMsg1* assumption is made in most of the literature [5–10] as it is recommended by the 3GPP for the performance analysis of the LTE-A RA procedure. Nevertheless, the *CMsg3* assumption is supported in studies such as [11] and also in works such as [12]. In recent studies such as [13] the performance analysis of the RA procedure under both assumptions has been conducted by modifying a module of the LTE-Sim simulator. In the latter study, the authors consider

that the *CMsg3* assumption is the scenario with realistic assumptions on the detection of preambles. Results show that the performance of the LTE-A RA procedure under each assumption is widely different. Nevertheless, little to no detail on the employed simulator and on the implications of the assumptions are described.

In this study, we evaluate the performance analysis of the LTE-A RA procedure under both assumptions; we also describe in detail the causes for each possible outcome and their implications. The rest of the paper is organized as follows. The RA procedure and the two possible outcomes are described in Section 2, along with their possible causes and implications on the RA procedure. The results derived from our performance of the RA procedure under both assumptions are presented in Section 3. Conclusions are presented in Section 4

2 LTE-A random access (RA) procedure

The RA procedure, as depicted in Fig. 1, is initiated by the UEs in order to shift from idle to connected-mode; i.e., for the initial connection between a UE and the eNB. The RA procedure is performed as follows [14–17].

Preamble transmission(*Msg1*): At the beginning of the RA procedure each UE randomly selects one out of the R available preambles (orthogonal sequences) and sends it to the eNB in a predefined time-frequency resources, known as random access opportunities RAO (*Msg1*). Due to the orthogonality of the different preambles, multiple UEs can access the eNB in the same RAO using different preambles. If a preamble is transmitted (with sufficient power) by exactly one UE, it is decoded by the eNB. On the other hand, two outcomes are possible if multiple UEs transmit the same preamble in a RAO:

The eNB does not decode the transmitted preamble (*CMsg1*). This may occur if: (i) the eNB determines that the preamble was transmitted by multiple UEs, i.e., based on the received signal power and the time shift between the multiple received copies of the preamble; (ii) the interference caused by the multiple preamble transmissions is too high so that the eNB is not capable of decoding the preamble; or (iii) all the preamble transmissions are lost due to a wireless channel error. Regardless of the cause, the UEs will not receive the RAR message within the RAR window. Then, the implicated UEs will detect the collision.

The eNB correctly decodes the transmitted preamble (*CMsg3*). This may occur if: (i) the received power from one of the preamble transmissions is significantly higher than that of the other simultaneous transmissions of the same preamble; i.e., the capture effect, or (ii) all but one of these preamble transmissions are lost due to a wireless channel error. Therefore, the multiple UEs that transmitted the preamble will receive the RAR message and continue with the RA procedure by sending *Msg3*. The eNB will receive multiple *Msg3*s with different data in the same reserved resources and will not transmit *Msg4* in response.

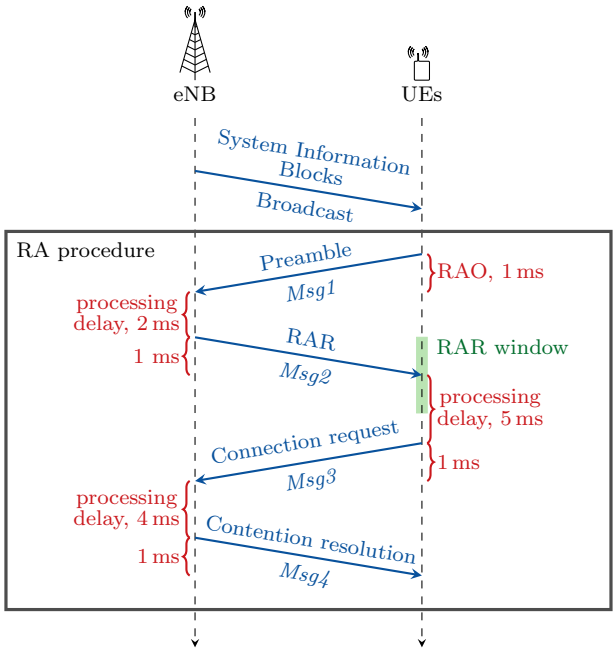


Fig. 1. Timing diagram of the RA procedure in LTE-A.

Random access response (RAR, *Msg2*): The eNB computes an identifier for each of the successfully decoded preambles and schedules the transmission of a RA response (RAR) message (*Msg2*). It includes, among other data, information about the identification of the detected preamble, time alignment, uplink grants (reserved uplink resources) for the transmission of *Msg3*, the backoff indicator, and the assignment of a temporary identifier. Exactly two subframes after the preamble transmission has ended (this is the time needed by the eNB to process the received preambles), the UE begins to wait for a time window, RAR window, to receive an uplink grant from the eNB.

There can be up to one RAR message in each subframe, but it may contain multiple uplink grants (up to three); each of which is associated to a successfully decoded preamble. The downlink resources per RAR message are limited and the length of the RAR window, W_{rar}, is fixed. Consequently, there is a maximum number of uplink grants that can be sent within the RAR window. Only the UEs that receive an uplink grant can proceed with the transmission of *Msg3*.

Connection request (*Msg3*) and contention resolution (*Msg4*): After receiving the corresponding uplink grant, the UE adjusts its uplink transmission time according to the received time alignment and transmits a scheduled connection-request message, *Msg3*, to the eNB using the reserved uplink resources. The RA procedure is concluded when the eNB sends the contention resolution message (*Msg4*) to the UEs in response to the connection request

message. Hybrid automatic repeat request (HARQ) is used to protect the *Msg3* and *Msg4* transmissions. If a UE does not receive *Msg4* within the Contention Resolution Timer, then it declares a failure in the contention resolution and schedules a new access attempt after backoff.

Backoff: If the RA procedure fails and if the maximum number of preamble transmissions (notified by the eNB through the *System Information Block 2*, SIB2 [15]), has not been reached, failed UEs ramp up their power and re-transmit a new randomly chosen preamble in a new RAO. For this, the UE waits for a random time, chosen uniformly between zero and the *Backoff indicator* and then performs a new preamble transmission at the next RAO. The Backoff indicator is defined by the eNB and its value ranges from 0 to 960 ms. The UEs are only aware of a failed preamble transmission if no uplink grant has been received at the end of the RAR window. Backoff is also performed by the UEs that do not receive the Contention resolution message (*Msg4*) after the maximum number of allowed Connection request messages (*Msg3*) has been reached.

When the maximum number of allowed preamble transmissions has been reached and the RA fails, the network is declared unavailable by the UE, a RA problem is indicated to upper layers, and the RA procedure is terminated.

As mentioned above, a RA fails whenever multiple UEs transmit the same preamble simultaneously. However, the implications of the two different outcomes, *CMsg1* and *CMsg3*, widely differ. The implications of these two outcomes, as briefly described in [13], are now discussed.

If the *CMsg1* outcome occurs, the colliding UEs do not receive an uplink grant within the RAR window; then these UEs will perform backoff. As mentioned above, W_{rar} is the length of the RAR window in subframes. Let t_{cmsg1} be the number of subframes needed for a UE to be aware of a failed preamble transmission; i.e., the number of subframes elapsed between the preamble transmission and the end of the RAR window. Since the preamble transmission is performed in one subframe and two subframes are needed for the eNB to process the transmitted preambles, we have

$$t_{cmsg1} = 3 + W_{\text{rar}}. \tag{1}$$

The timing diagram of the *CMsg1* outcome is shown in Fig. 2.

On the other hand, if the *CMsg3* outcome occurs, the colliding UEs will receive an uplink grant within the RAR window and proceed to transmit the Connection request message (*Msg3*). These UEs will transmit their Connection request messages (*Msg3*) in the same reserved resources; since the *Msg3* transmissions include a unique identifier for each UE, these transmissions are different from one another. As a result, the eNB will not be able to decode the transmitted *Msg3*s and will send a NACK message instead of *Msg4*. This process will be performed a total of $H_{\max}$ times, where $H_{\max}$ is the maximum number of *Msg3* and/or *Msg4* transmissions allowed; the round-trip-time (RTT) of the *Msg3* transmissions is $T_{Msg3} = 8$ ms. Afterwards, the colliding UEs will perform backoff. Let t_{cmsg3} be the number of subframes needed for a UE to be aware of $H_{\max}$ failed *Msg3* transmissions. Also let w_{rar} be the number of subframes

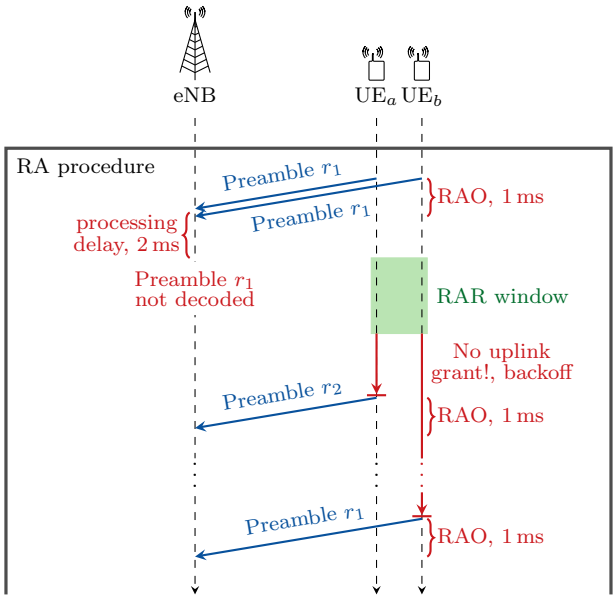


Fig. 2. Timing diagram of a failed RA procedure given that the eNB cannot decode the preamble transmitted by multiple UEs (*CMsg1*).

elapsed since the beginning of the RAR window until the reception of the uplink grant; w_{rar} may be different each time the *CMsg3* outcome occurs and depends on the number of preambles decoded at each RAO. Since the preamble and uplink grant transmissions are performed in one subframe, five subframes are needed for the UEs to process the uplink grant and two subframes are needed for the eNB to process the transmitted preambles, we have

$$t_{cmsg3} = 4 + w_{\text{rar}} + 5 + H_{\max} T_{Msg3}. \tag{2}$$

The timing diagram of the *CMsg3* outcome is shown in Fig. 3.

As it can be seen, the *CMsg3* outcome implies the wastage of uplink grants in UEs whose *Msg3* transmissions will eventually collide; i.e., whose RA will invariably fail. Furthermore, the delay of UEs involved in the *CMsg3* outcome is much longer that for those involved in the *CMsg1* outcome.

As mentioned above, the two most common assumptions that are made during the performance analysis of the LTE-A RA procedure are: (i) the *CMsg1* outcome occurs every time multiple UEs transmit the same preamble simultaneously and (ii) the *CMsg3* outcome occurs every time multiple UEs transmit the same preamble simultaneously. Hereafter, we denote these two assumptions simply as the *CMsg1* and the *CMsg3* outcomes.

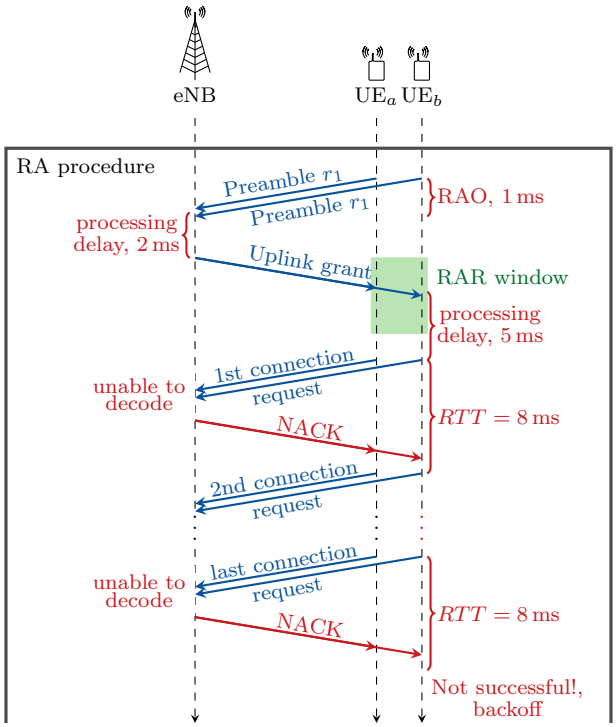


Fig. 3. Timing diagram of a failed RA procedure given that the eNB decodes a preamble transmitted by multiple UEs (*CMsg3*).

3 Performance analysis of the LTE-A RA procedure

In this section, we evaluate the performance of the RA procedure under both the *CMsg1* and the *CMsg3* assumptions. We have selected a typical RACH configuration, i.e., *prach-ConfigIndex* = 6, where RAOs occur every $T_{\text{RAO}} = 5$ ms [4]. Two different Traffic models are considered. In the first one, Traffic model 1, N UE arrivals are randomly distributed throughout the distribution period, of length $t_d = 12000$ RAOs (60 seconds), following a uniform distribution. In the second one, Traffic model 2, $N = 30000$ UEs are randomly distributed throughout the distribution period, of length $t_d = 2000$ RAOs (10 seconds), following a Beta(3,4) distribution. These and other configuration parameters selected for the performance analysis of the RA procedure are enlisted in Table 1 [4, 17]. The characteristics of the two different traffic models are shown in Table 2.

Results are obtained by means of a simulator coded in C that strictly adheres to the RA procedure as described in the specifications [14–17] and to the suggested behavior for its performance analysis [4]. The accuracy of our simulator has been confirmed during previous studies [18]. In order to obtain the results under the *CMsg3* assumption, we have modified one of the modules of our sim-

Table 1. Basic configuration parameters and processing delays for a typical RACH configuration.

Parameter	**Setting**
PRACH configuration index, *prach-ConfigIndex*	6
RAO periodicity, T_{RAO}	5 ms
RAR window length, W_{rar}	5 subframes
Number of available preambles	54
Available uplink grants per subframe	3
Available uplink grants per RAR window, N_{UG}	15
Maximum number of preamble transmissions	10
Maximum number of *Msg3* and *Msg4* transmissions, $H_{\max}$	5
Backoff indicator	20 ms
Preamble processing delay	2 subframes
Uplink grant processing delay	5 subframes
Msg3 processing delay	4 subframes
Msg3 RTT	8 subframes
Msg4 RTT	5 subframes
Re-transmission probability for *Msg3* and *Msg4*	0.1
Detection probability for the kth transmitted preamble	$1 - 1/e^k$

Table 2. Traffic models for the performance evaluation of the LTE-A RA procedure.

Parameter	**Traffic model 1**	**Traffic model 2**
Number of UE arrivals, N	$12 \times 10^4, \ldots, 36 \times 10^4$	3×10^4
Distribution period, t_d	12×10^3 RAOs	2×10^3 RAOs
Distribution of UE arrivals	Uniform	Beta(3,4)

ulator according to the MAC protocol specifications [14]. Simulations are run j times until the cumulative results obtained up to the j simulation differ from those obtained up to the $(j-1)$th simulation by less than 1%.

As a starting point, we evaluate the performance of the LTE-A RA procedure under Traffic model 1; the number of UE arrivals, N, has been selected in order to achieve an average number of UE arrivals per RAO between 10 and 30 for a long distribution period. The access success probability, i.e., the probability of successfully completing the RA procedure under both assumptions, *CMsg1* and *CMsg3*, for the given number of UE arrivals is shown in Fig. 4. As it can be clearly seen, the success probability is significantly lower for the *CMsg3* assumption when compared to the *CMsg1* assumption. This is mainly due to the wastage of uplink grants in UEs that will later fail the RA procedure. Specifically, the success probability under the *CMsg3* assumption drops rapidly if an average of ≈ 10 or more UE arrivals occur per RAO. On the other hand, the success probability under the *CMsg1* assumption begins to drop when the average number of UE arrivals approaches the available number of uplink grants per RAR window, $N_{\text{UG}} = 15$. Consequently, the average number of UEs that

can be effectively served under *CMsg1* is much higher than that under *CMsg3* assumption.

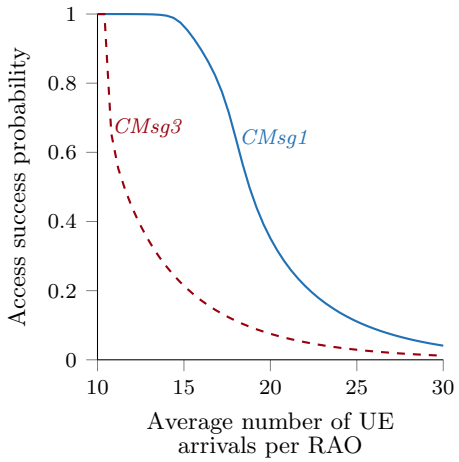


Fig. 4. Access success probability for the given average number of UE arrivals per RAO under both assumptions: *CMsg1* and *CMsg3*.

Next we proceed to evaluate difference on performance of the RA procedure under the *CMsg1* and *CMsg3* under Traffic model 2. For this, the expected number of successful UE accesses per RAO under both assumptions is shown in Fig. 5a; the expected number of UE arrivals and the number of available uplink grants, N_{UG} are also shown as context. As it can be seen, during the period of highest congestion, i.e., between the 500th and the 1500th RAO, the expected number of successful accesses under assumption *CMsg3* is very close to zero. Furthermore, the highest expected number of successful accesses under this assumption is close to ten. On the other hand, the expected number of successful accesses under assumption *CMsg1* is always higher than or equal to that under assumption *CMsg3* and its highest value is very close to $N_{\mathrm{UG}} = 15$.

We have also studied the access delay under both assumptions, defined as the time elapsed between the first preamble transmission and the successful completion of the RA procedure. Fig. 5b shows the cumulative distribution function (CDF) of the access delay under both assumptions. It can be clearly seen that the CDF of the access delay under the *CMsg1* assumption increases more rapidly than the one obtained under the *CMsg3* assumption. Hence, the access delay under the former is shorter than the one obtained under the latter.

4 Conclusion

In this paper, we have described the two possible outcomes of the LTE-A RA procedure for the case in which multiple UEs transmit the same preamble simultaneously. We have also discussed the implications of each outcome; which were

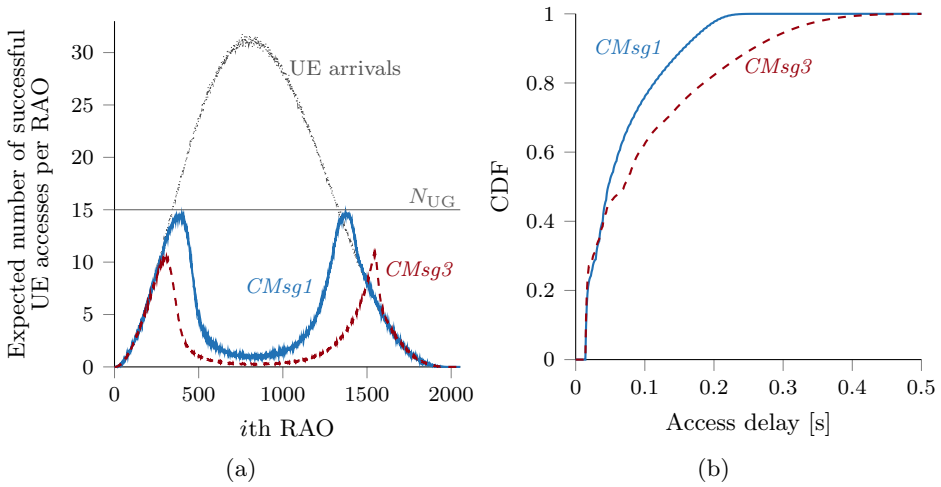


Fig. 5. (a) Expected number of successful UE accesses per RAO and (b) CDF of access delay under Traffic model 2 and both assumptions: *CMsg1* and *CMsg3*.

then confirmed by the results obtained by simulation under both, *CMsg1* and *CMsg3* assumptions. Specifically, the probability of completing the RA procedure is lower and the access delay is longer under the *CMsg3* assumption when compared to the *CMsg1* assumption. In other words, the *CMsg1* outcome is the best-case scenario for the performance of the RA procedure. Conversely, the *CMsg3* outcome is the worst-case scenario.

In a real implementation, both outcomes may occur with a certain probability. Consequently, the performance of the RA procedure in a real implementation will be located in between the results obtained under the *CMsg1* and *CMsg3* assumptions. Therefore, evaluating the performance under the *CMsg1* and *CMsg3* assumptions provides the upper and lower bounds to the performance of the LTE-A RA procedure.

Acknowledgment

This research has been supported in part by the Ministry of Economy and Competitiveness of Spain under Grants TIN2013-47272-C2-1-R and TEC2015-71932-REDT. The research of I. Leyva-Mayorga was partially funded by grant 383936 CONACYT-Gobierno del Estado de México 2014. The research of L. Tello-Oquendo was supported in part by Programa de Ayudas de Investigación y Desarrollo (PAID) of the Universitat Politècnica de València.

References

1. Ghavimi, F., Chen, H.H.: M2M Communications in 3GPP LTE/LTE-A Networks: Architectures, Service Requirements, Challenges, and Applications. IEEE Commun. Surv. Tutorials **17**(2) (2015) 525–549
2. Verma, P.K., Verma, R., Prakash, A., Agrawal, A., Naik, K., Tripathi, R., Khalifa, T., Alsabaan, M., Abdelkader, T., Abogharaf, A.: Machine-to-Machine (M2M) Communications: A Survey. J. Netw. Comput. Appl. **66** (2016) 83–105
3. 3GPP: Architecture enhancements to facilitate communications with packet data networks and applications (Mar 2016)
4. 3GPP: Study on RAN Improvements for Machine-type Communications. TR 37.868, 3GPP (Jul 2011)
5. Wei, C.H., Bianchi, G., Cheng, R.G.: Modeling and analysis of random access channels with bursty arrivals in OFDMA wireless networks. IEEE Trans. Wirel. Commun. **14**(4) (2015) 1940–1953
6. Cheng, R.G., Chen, J., Chen, D.W., Wei, C.H.: Modeling and analysis of an extended access barring algorithm for machine-type communications in LTE-A Networks. IEEE Trans. Wirel. Commun. **14**(6) (2015) 2956–2968
7. Lin, T.M., Lee, C.H., Cheng, J.P., Chen, W.T.: PRADA: Prioritized random access with dynamic access barring for MTC in 3GPP LTE-A networks. IEEE Trans. Veh. Technol. **63**(5) (2014) 2467–2472
8. Tavana, M., Shah-Mansouri, V., Wong, V.W.S.: Congestion control for bursty M2M traffic in LTE networks. In: Proc. IEEE International Conference on Communications (ICC). (2015) 5815–5820
9. Suyang Duan, Shah-Mansouri, V., Wong, V.W.S.: Dynamic access class barring for M2M communications in LTE networks. In: Proc. IEEE Global Telecommunications Conference (GLOBECOM). (2013) 4747–4752
10. Wei, C.H., Cheng, R.G., Tsao, S.L.: Performance analysis of group paging for machine-type communications in LTE networks. IEEE Trans. Veh. Technol. **62**(7) (2013) 3371–3382
11. Osti, P., Lassila, P., Aalto, S., Larmo, A., Tirronen, T.: Analysis of PDCCH Performance for M2M Traffic in LTE. IEEE Trans. Veh. Technol. **63**(9) (2014) 4357–4371
12. Dahlman, E., Parkvall, S., Skold, J.: 4G: LTE/LTE-Advanced for Mobile Broadband. (2013)
13. Andrade, T.P.C.D., Astudillo, C.A., Sekijima, L.R., da Fonseca, N.L.S.: The Random Access Procedure in Long Term Evolution Networks for the Internet of Things. IEEE Wirel. Commun. **55**(3) (2017) 124–131
14. 3GPP: Medium Access Control (MAC) protocol specification. TS 36.321 V13.0.0, 3GPP (Feb 2016)
15. 3GPP: Radio Resource Control (RRC); Protocol specification. TS 36.331 V13.0.0, 3GPP (Jan 2016)
16. 3GPP: Physical layer procedures. TS 36.213 V13.0.0, 3GPP (May 2016)
17. 3GPP: Feasibility study for Further Advancements for E-UTRA. TR 36.912 V13.0.0, 3GPP (Jan 2016)
18. Leyva-Mayorga, I., Tello-Oquendo, L., Pla, V., Martinez-Bauset, J., Casares-Giner, V.: Performance analysis of access class barring for handling massive M2M traffic in LTE-A networks. In: Proc. IEEE International Conference on Communications (ICC). (2016) 1–6

GNSS Reflectometry Application in an Integrated IoT System

Sara Blanc[1], José Luis Bayo[1], Francisco Rodriguez[1], Angel Martín[2], Ana Belén Anquela[2], Sara Ibáñez[3], Héctor Moreno[3]

[1] Instituto ITACA, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022, Valencia, Spain
sablacla@disca.upv.es
[2] Dpto. de Ingeniería Cartográfica, Geodesia y Fotogrametría, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022, Valencia, Spain
[3] Dpto. de Producción Vegetal, Centro Valenciano del Estudios del Riego (CVER), Universidad Politécnica de Valencia, Camino de Vera s/n, 46022, Valencia, Spain

Abstract. Sensors, Intelligent sensors and now High Performance Sensors are the new challenge in the IoT paradigm. High performance sensors are complex sensors able to acquire a high quantity of heterogeneous data by themselves in short time cycles with some performance capabilities. Examples are the next generation of Global Navigation Satellite System (GNSS) sensors or agro-context-aware wireless sensor networks. This project works on adapting multi constellation GNSS Reflectometry to specific agro-field services that drawn together with weather measures, probes for crops and other external data to defining the Internet of Agro-Services.

1 Introduction

Smart farming and agriculture innovation tend towards the use of many heterogeneous distributed sensors to collect data of different types (Internet of Things, Internet of Agro). Actions to be taken over the crops by farmers, such as optimum water use in irrigation and fertirrigation, will depend on elaborated decision making processes based on real data. In this research line, our project is focused on developing a new system to soil moisture direct observation by GNSS reflectometry and combined with meteo sensors to measure important factors such as the evapotranspiration in real time. Each sensor defines a micro-heterogeneous measurement system able to acquire data from multi -ubiquitous sources (Things). Thus, it is expected a high quantity of data (big data) collected in the crop which drive us into two problems. Firstly, the quantity of observable raw data can be higher than the volume managed by the Internet transmission link (actually constrained in agro-field scenarios). For example, SigFox connection is a low cost solution but with restrictions over the number of daily messages. Secondly, such a quantity of data easily derives into *infoxication* when the magnitude of the final system has not been clearly defined attending to its scope (or to the farmer expectations) and all measurable raw data "seems" to be important. Moreover, alt-

C.Fernandez-Llatas and M.Guillen (Ed.):
Workshop on Innovation on
Information and Communication Technologies
(ITACA-WIICT 2017)

ISBN 978-84-697-7327-7

hough a clear definition of the agro-system scope would be done, new requirements use to appear soon after their deployment. Due to the difficulties of managing sensors installed in crops and associated costs it is desirable to actuate remotely over them to reconfigure their processes. Therefore, information should flow in both directions. From the crops to cloud services and friendly applications designed for farmers in one hand, and from the cloud services to the crop installed sensors in the other hand, being able to recognize different context demands. *Sensoring* is then a not simple action but a performance of information management and data protection within a scalable and portable middleware useful to support many different interrelated tasks sharing a common interface. Such a middleware must be able to both acquire and treat (compute) raw data coming from non-homogeneous sensors installed in-the-field as well as to manage data from interoperated Internet connected services.

Additionally to the local farmer field, intelligent agro-services also depend on out-of-field data such as climate data (past and present), crops characterization, tillage working, meteorological services, or agro-data acquired in near fields (other GNSS sensors or agro-WSN) overall integrated in an Internet of Agro (IoA) system. The result is a complete business ecosystem that covers intelligent and high performance sensors designers, upper-sensor-level middleware developers, Internet-link operators, cloud services developers and the most important, the farmer.

This observation system can be considered as a pilot project for data integration services. Data transferred from and to the crops is treated as an agro-context-aware service in a Service Oriented Architecture (SOA) supported by choreograph middleware software.

2 Edge-cutting technologies

Soil moisture observations has been traditionally taken by ground probes [1]. Adding network connection capacities to these probes enhances data reading management. Probes remote control and accessing transforms them into sensors.

Due to soil features, the number of probes sensors ranges from a few ones to a high numbers of probes. As many testing points more accurate will be soil moisture measurement. However, probes (including the connectors, cables and data logger device) increase the cost of the high-resolution soil moisture map determination. A first action to reduce costs is the development of Wireless Smart Sensor Networks to integrate every in-the-field sensor in networking. Ground sensors, weather stations and any other in-the-field data acquisition system is composed in a service (or more than one) accessed through a hotspot gateway reducing Internet traffic (compared with individualized accesses sensor by sensor). In Agro-IoT applications such a clustering is very desirable due to the frequent Internet access constrains in arable areas. Data can be filtered, temporally stored (in case of connectionless faults), aggregated and buffered in different context queues. Although reducing connectors, cables and data logger costs, ground probes are expensive and hardly used for covering the full extension. Therefore, making decision processes over irrigation, fertirrigation or water footprint are biased by the specific soil features and tillage over the crop testing point, not gen-

erally transposable. Thus, advancement in the use of novel techniques, less expensive but more comprising, is highly necessary.

GNSS reflectometry, supported by single and multiple constellations, i.e. GPS, GLONASS and GALILEO is potentially a new low-cost technology but a challenge considering it already experimental in soil moisture observation. Global Navigation Satellite Systems (GNSS) lie in L band, which is the most sensitive band for soil moisture microwave remote sensing. Several studies have shown that, with specially designed GNSS antenna and receiver systems [2], it is possible to estimate near-surface soil moisture. These measurements were carried out with custom designed systems based on two GNSS antennas: one tracking the direct signals from the satellites and the other oriented toward the ground to track the reflected signal. However, it is possible to estimate soil moisture with standard single ground-based dual-frequency geodetic GNSS receivers. These geodetic type GNSS instruments, which are optimized to track the direct signals from the satellites, could be successfully applied to measure the reflected signals too avoiding the need to build a double GNSS antenna system, reducing the costs of the technique.

Performed studies using observations from GPS constellation showed very good correlation with land soil moisture sensors at a 10 cm depth. A challenge comes from the upcoming GALILEO full operation constellation. GALILEO signal power emission is higher compared with GPS or GEOSAT, so soil moisture information is expected at 20-25 cm depth. But GPS and GEOSAT satellite constellations cannot be discarded because GNSS reflectometry using multi constellation will increase horizontal resolution (as more satellites more tracks in ground obtained) and vertical resolution (information at different depths).

Resolution improvement is a significant advantage. However, the quantity of raw data expected per lecture is very high, in the order of several Mb, not affordable in real-time by typical Internet connections in arable areas. Therefore, GNSS signal treatment and pre-data processing should be considered as part of a Smart sensoring process. Additionally to quantity, data quality is also a keystone. When sensor lectures are distributed along the day in a fixed sample frequency raw data includes useful as well as not useful data obtained by satellites out of the ideal azimuth range. But smart sensoring could implement discontinuously lectures attending to intelligent algorithms that adapt cycles to day time for any specific GNSS sensor localization.

Due to the above mentioned reasons, some computational potential is required to manage both in-situ systems (WSN, GNSS sensor, etc.) as well as information but saving hardware resources. Hotspots are really constrained in memory and energy, issues that can be relieved by an embedded choreograph engine. The arable area is virtualized to the farmer by the abstraction of every in-the-field sensor into a unique integrated service by transforming the hotspot is a smart intelligent sensor that brings together control over the sensors, information management and M2M interoperability (data are transferable out of the field).

Machine learning algorithms, geo-mapping and so on are implemented and executed in cloud. Data transferring from the field into the cloud is the feature that transform a sensor into an IoT sensor. Crossed information (data) obtained from out-of-the-field services increases the possibilities in the decision making process design at the busi-

ness level. For example, climate evolution, meteorological historical data or new crop features will change agro system from reactive into pro-active systems.

2 Technical Project Description

This project in focused on an specific agro-IoT service: soil moisture direct observation. The scenario in Fig. 1 includes 7 states:

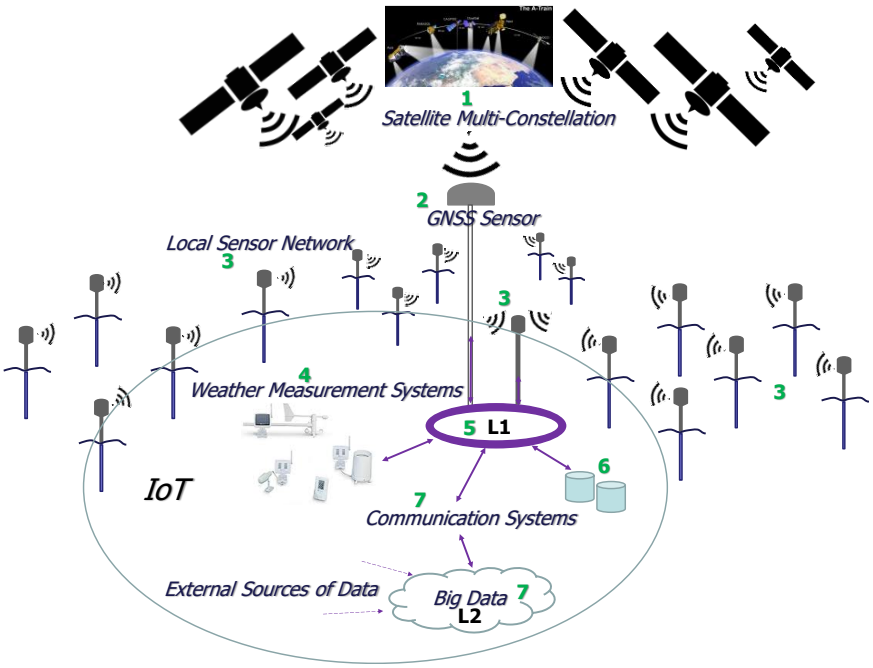


Fig. 1 Agro-IoT scenario.

The scenario in Fig. 1 includes 7 states:

1.- The multi-constellation Global Navigation Satellite System (GNSS) reflectometry will provide soil moisture information increasing horizontal resolution (more satellites used, so more tracks in ground could be obtained) and vertical resolution (information at different depths). Moreover, a combined solution that includes multi-satellite constellation, including GALILEO, could provide more precision to measures (horizontal and vertical resolution)
2.- Global Navigation Satellite Systems (GNSS) lie in L band, which is the most sensitive band for soil moisture microwave remote sensing. Geodetic and low-cost GNSS antenna transforms SNR (Signal-to-Noise ratio) uploaded in different L frequencies

within 5 to 30 seconds interval, and other GNSS data (coordinates of the antenna, time, elevation and azimuth of the satellites), into digital variables.
3.- Ground probe sensors support GNSS data calibration. Soil moisture observations are traditionally obtained by ground probes. A high quantity of ground probes increases the cost of an equivalent high-resolution soil moisture map determination, but they are needed to obtain input references for GNSS training. Although they will not be part of the final system, ground probes are the mean to obtain a significant sample of input data. Moreover, this data will be obtained from different arable fields and crops to analyze their influence on soil moisture measurement.
4.- Agro- IoT feeds on multi-source heterogenic information for making decision processes. Evapotranspiration (or PET) is the amount of water that would be evaporated and transpired if there were sufficient water available. Thus, in-the-filed direct observation requires of additional instruments to the ground probes.
5.- (L1) Hotspot based on a choreograph service paradigm. Services under choreography perform independently but with real time communication capabilities to interchange data in order to carry out simple service or complex workflow orchestrated processes. Moreover, portability and scalability are inherent to choreography while interoperability rests on the integration of choreography with REST API desing.
6.- High performance sensors imply high quantity of data. Agro-field scenarios use to be constrained on Internet facilities. Thus, choreography services need to include local storage, filtering, information management and REST M2M communication with high level services (in cloud).
7.- (L2) Data (big data) can hardly processed with computational local resources due to their energy and memory limitations. IoT cloud storage and orchestration computing will suitably supports the expected outcome agro- applications and services.

3 Conclusions

In September 2016 EIP-AGRO Focus Group in Water and Agriculture presented its annual report about strategies at farm level [3]. This report strengthened the understanding about the necessity of investing in water saving as essential resource which has suffer negative impacts in the last decade. *"Water availability may be increased with strategies that reduce water losses or increases the capacity to store the water to be used by crops or livestocks"*.
The poster presents an on-going research project in the framework of Agro-IoT services. It combines edge-cutting technologies development in on-farm agro-sensors and advancement in technology to support new and more enhanced decision support systems (DSS).

References

1. Dalton, F.N., Van Gemuchten, M. Th (1986). The time domain reflectometry method for measuring soil water content and salinity. Geoderma 38: 237-250.

2. Larson, K.M., Small, E.E., Gutmann, E. D., Bilich, A.L., Axelrad, A., Braun, J.J. (2008): Using GPS multipath to measure soil moisture fluctuations: initial results. GPS solutions, 12(3), 173-177.

3. EIP-AGRO focus group report (September 2016). Available at https://ec.europa.eu/eip/agriculture/sites/agri-eip/files/eip-agri_fg_water_and_agriculture_final-report_en.pdf (last access on 14th September 2017)

A NS3 model to evaluate Sigfox features in different environments

Pablo Pardal Garcés[1], José Carlos Campelo Rivadulla[1], and Alberto Miguel Bonastre Pina[1]

[1] Instituto ITACA, Universitat Politècnica de València, Camino de Vera s/n, 46022, València, España,
pabparg1@teleco.upv.es, {jcampelo, bonastre}@itaca.upv.es

Abstract. Nowadays we live in a communicated world in which the concept of Internet of Things (IoT) is becoming more important. This concept refers to the digital interconnection of any object with Internet through different technologies. Some of these technologies, such as LTE, WiFi or Bluetooth, may be considered "traditional", as far as they are being used for computed communications. However, other technologies, have been developed recently for the IoT world. These are the so called Low-Power Wide Area Networks (LPWAN), such as LoRa, Sigfox or RPMA among others. These new technologies, which aim to become future standards, offer many interesting capabilities, which have not yet been proven.

Sigfox is one of the most promising technologies in this line, and thus many new devices use it. In order to be able to verify the characteristics of these devices, a NS3-model of the Sigfox communication technology has been implemented. With this module is possible to study many of its characteristics, such as coverage, power consumption and error rate, in different environments.

1 Introduction

In recent years, the concept of Internet of Things (IoT) has arisen as an idea through which the Internet model would focus on total connectivity, where all objects that we can imagine could be interconnected by means of digital networks. The IoT is one of the great technological bets of today

This model has become one of the most important research and business topics in many technology companies. The radical increase in the number of devices connected to Internet in the last ten years has led to the fact that the number of connected devices is higher than the number of inhabitants on the planet.

In a closer future most of the information that travels through the communication networks will be generated by objects connected, related to multiple application fields, such as home automation, smart cities, smart cars, industry, energetic efficiency or telemedicine, among many others.

C.Fernandez-Llatas and M.Guillen (Ed.):
Workshop on Innovation on
Information and Communication Technologies
(ITACA-WIICT 2017)

ISBN 978-84-697-7327-7

The long list of advantages in different sectors justifies the effort in the improvement of communication networks. It is necessary to advance, among others, in aspects such as network and device security, bandwidth distribution, support infrastructures or device consumption.

That is why new communication technologies oriented to the Internet of Things have emerged to cover the shortcomings of traditional technologies, such as short-range wireless networks (e.g., Bluetooth or Z-Wave), wireless local area networks (e.g., WiFi) or cellular networks (e.g., LTE). All of them have drawbacks when applied to IoT traffic. Usually, objects require the transmission of a small amount of bytes, with low power consumption but long distances, to avoid infrastructure requirements. When dealing with this traffic, the non-cellular technologies have an elevated power consumption and a short-range that make them inaccurate for these communications. On the other hand, cellular technologies are compliant with the requirement of long-ranges but the consumption specifications – at least ten years of battery lifetime – are not satisfied.

In this way, the so-called LPWAN networks have been developed around this compromise between long distance and low power. They are expected to take an important role in the IoT sector due to their unique characteristics. They offer ranges from a few meters to a tens of kilometers whereas they allow a battery life for more than ten years. The LPWAN long-range and low-power consumption networks offer the best capabilities to implement many IoT architectures.

There are several open discussions in the community about which LPWAN communication technology offers the best capabilities. In [1] we can see an extensive summary of the LPWAN paradigm and a complete description of many of the different new technologies, such as Sigfox, LoRa, Ingenu or Weightless-Sig, among others.

In this paper we have focused on the study of Sigfox because it is, from our point of view, one of the most promising LPWAN communication technology. Sigfox is based in an open standard from the ETSI, called Low Throughput Network (LTN), and if offers the best coverage at this moment, leading the international market ahead of its competitors.

Simulation techniques have been selected to study this new technology. A simulation model allows many performance studies, offering a wide variety of possible scenarios, many of them not available for real observation.

Due this reason and the protagonist role of Sigfox, we implemented a NS-3 module to study the performance of a Sigfox network. With this module is possible to simulate the behavior of Sigfox networks in different environments. This will be very helpful for comparative evaluation with other technologies, such as Zigbee, LTE or WiFi, that also may be considered suitable for some IoT applications.

The rest of the paper is structured as follows. In Sections 2 and 3 the principles and characteristics of the Sigfox technology and the radio interface of the ETSTI LTN standard, respectively, are briefly described. In Sec. 4 the NS-3 model of Sigfox communication technology are explained. Finally, the conclusions of this work may be found in Section 5.

2 Sigfox

As stated before, one of the most relevant communication technologies today is Sigfox Network, based on the LTN standard of ETSI and on the UNB transmission technology. Sigfox is a French company founded in 2009 that offers LPWAN services. Its connectivity solution relies on an infrastructure of antennas and base stations completely independent of existing networks, such as 2G, 3G or 4G. The main objective of Sigfox is to allow every object to connect to the Internet anywhere with autonomous power supplies. It is currently the main service provider in the IoT sector in Europe.

The operation of this network is very similar to cellular networks because both of them are based on the placement of several receiving and transmitting stations. The difference between SigFox and cellular stations is that devices and sensors that are connected to the SigFox network are not linked to a single specific base station. Any station can receive the information from any device and transmit it to the cloud. Thereby, the coverage offered by a Sigfox network reaches distances of up to 50 km in rural environments and 3-10 km in urban environments due to the UNB transmission technology principles.

To stablish this communication, the SigFox network uses the unlicensed bidirectional ISM radio bands – reserved for Industrial, Scientific and Medical applications, that cannot be used for telecommunications. Many ISM bands are defined, but the most popular is that located around 902 MHz in the United States, and 868 MHz in Europe . For LTN applications, a bandwidth of 200 kHz is reserved, thus enabling 2000 channels of 100 Hz. Each message can be transmitted in any of these channels, with a bit rate from 100 bps up to 600 bps. Those rates vary according with the local normative. The European uplink spectrum is shown in Fig. 1.

Sigfox limits the maximum number of messages that can be sent per day is 140. This limitation is introduced to comply with the spectrum use regulations. The European regulation provides that a transmission duty cycle of 1% must be met for the 868MHz band. Therefore, a device is not authorized to transmit more than 1% of the time every hour and the transmission of a message can take up to six seconds. This allows up to six messages per hour.

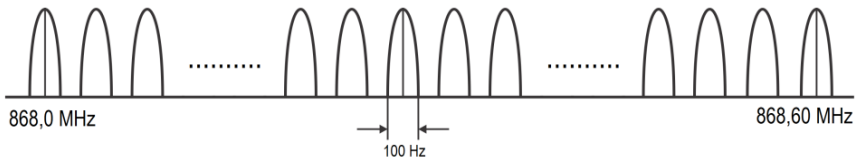


Fig. 1. Uplink spectrum channels. Sigfox defines a bandwidth of 200 kHz.

3 Low Throughput Network (LTN) radio interface

The ETSI LTN communication standard is defined in three specification documents, which specify the use cases of the standard [2], its functional architecture [3], and the operation of the protocol and its interfaces [4] . In the present project, only the radio interface between access points and end-points is considered relevant in order to model the Sigfox communication technology. This interface defines two possible radio implementations that use the same radio spectrum: UNB that uses ultra-narrow band communication and OSSS, that uses orthogonal sequence spread spectrum technologies. In this work only UNB transmission technology has been considered in our models, so OSSS-based implementation of Sigfox will not be used in this paper.

3.1 UNB

The UNB transmission technology is based in the utilization of ultra-narrow spectrum channels to reach great distances with a minimum energy requirement. Transmission of data at small bandwidths helps to avoid interference and allows large numbers of end-points in a given cell.

To access the communication medium, UNB uses the so-called *Random Frequency and Time Division Multiple Access (RFTDMA)* technique. It consists of a random access to the wireless medium both in time and frequency domain, without any contention based-protocol. This technique is described in [5].

The UNB solution is dedicated to transmit the uplink messages, and downlink transmissions are available as a response of these messages or as broadcast messages. The synchronization between both links makes that the end-points open a fixed reception window after each uplink transmission. During this window, the access points transmit the downlink messages.

The characteristics of each link are described below.

3.1.1 Uplink

Uplink radio specifications:

- **frequency band:** 200 kHz within the 868,00 to 868,60 MHz band
- **channelization mask:** 100 Hz (600 Hz in the USA)
- **number of channels:** 2.000 channels
- **baud rate:** 100 baud (600 baud in the USA)
- **modulation scheme:** BPSK
- **maximum transmission power:** compliant with local regulation (typically 25 mW)
- **sensitivity:** better than -135 dBm
- **duty cycle:** 1%

3.1.2 Downlink

Downlink radio specifications:

- **frequency band:** 869,40 to 869,65
- **channelization mask:** dynamic selection
- **baud rate:** 600 baud
- **modulation scheme:** GFSK
- **maximum transmission power:** compliant with local regulation (typically 500 mW)
- **duty cycle:** 10%

Fields of the UNB downlink MAC frame (see Fig. 3):

- **preamble:** 32 bits
- **frame sync:** 13 btis
- **flags:** 2 bits
- **FCS:** 8 bits
- **authentication:** 16 bits
- **error codes:** variable
- **payload:** variable

Fields of the UNB uplink MAC frame: (see Fig. 2)

- **preamble:** 4 Bytes
- **frame synchronization:** 2 Bytes
- **end-point ID:** 4 Bytes
- **payload:** 0-12 Bytes
- **authentication:** variable length
- **frame check sequence:** 2 Bytes (CRC)

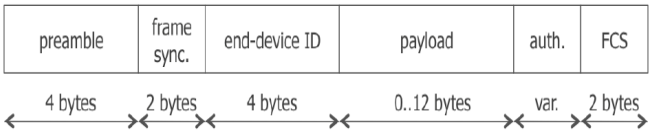


Fig. 2. Uplink UNB MAC frame.

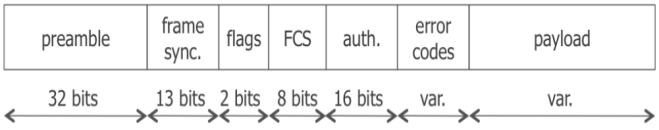


Fig. 3. Downlink UNB MAC frame.

4 NS-3 Sigfox model

As already mentioned, a large part of the IoT market bets by the use of Sigfox in their applications[6]. Thus, in addition to the free character of the ETSI LTN standard, we decide to develop a NS-3 Sigfox model to simulate networks operating with this communication technology. This model allows obtaining information regarding the power consumption, the coverage, the link quality, the error rate, the randomness of the RFTDMA protocol or the content of the frames.

In Fig. 4 we can see the structure of the developed Sigfox model. The net-device module is the basis of the model and is formed by two main modules, phy-layer and mac-layer, which are described as below. In addition, we are going to describe the propagation loss model used in the simulations, the Okumura-Hata loss model.

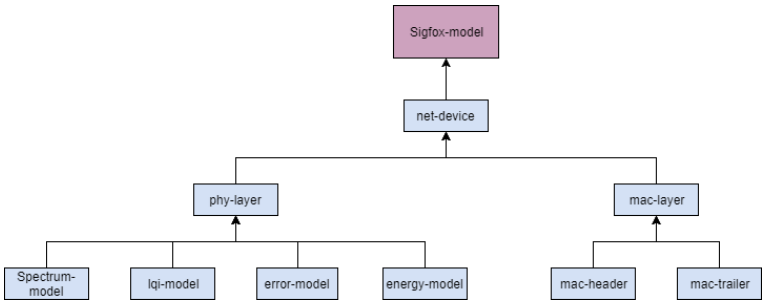


Fig. 4. Structure of the NS-3 Sigfox model.

4.1 MAC layer

This module designs the MAC layer. Typically, the main function of MAC layers is the control to the access [7] to the physical transmission medium. In this case, due to the simplicity of the RFTDMA medium access protocol, this function is implemented in the PHY layer. By this way, the present MAC layer only forms the data packets to be transmitted to the lower layer. To do this, it uses two support modules to create the trailer and the header of the packets. The payload is added by the main MAC layer module.

4.2 PHY layer

This model works as a state machine with four states: sleep, transmission, reception and transient. The main operation stage of the device is the sleep mode. From this state, the device may be switched on to the transmission mode by order of the upper layer. It can also be switched on to the receiving mode after the detection in the wire-

less medium of the preamble established by the protocol. Once information has been transmitted or received, the device switches to a transient state to change consecutively to the initial sleep mode. Each one of the states has a different power consumption which were established following the specifications of the Telit LE51 chip, which can operate in a Sigfox mode.

As discussed in the previous point, one of the characteristics of this physical layer is that it also implements the medium access protocol. This protocol works merely by selecting a random wireless communication channel from those available without any contention based-protocol. Thus, the PHY layer at the transmission state only select a random channel to transmit.

In addition to the main module, that manages the physical layer, other four modules define the wireless link and its physical characteristics, study the quality of this link, evaluate the error rate in the communication and obtain information about the power consumption that has the modeled device. To complete the evaluation of the physical characteristics of the wireless link we use the Okumura-Hata propagation loss model, which is described below.

4.3 Okumura-Hata propagation loss model

The Okumura-Hata propagation loss model is one of the most used to predict the losses caused by the medium in the range 100 MHz to 1900 MHz. It is based on an extensive battery of measurements taken during a project carried out in Tokyo. This model is considered one of the simplest but still very representative in terms of its precision in the calculation of the losses in the wireless medium. It has become the method used for planning of mobile systems in Japan.

This model defines the propagation losses with the following formula (1):

$$L_b = 69.55 + 26.16\ \log f\ - 13.82\ \log h_b - a(h_m) + (44.9 - 6.55 \log h_b) \log d_m \quad (1)$$

Where h_m is a correction factor that depends on the height of the antenna. This formula may be calculated for different environments as follows:

1. Rural areas

$$L_b = L_b(urban) - 4.78 \log(f)^2\ + 18.33 \log f\ - 40.94 \quad (2)$$

2. Suburban areas

$$L_b = L_b(urban) - 2\left[\log\left(\frac{f}{28}\right)\right]^2 - 5.4 \quad (3)$$

3. Urban areas
    - Medium or small sized cities ($1 \leq h_m \leq 10$ m):

$$a(h_m) = (1.1 \log f - 0.7)h_m - (1.56 \log f - 0.8) \quad (4)$$

    - Big cities:

$$a(h_m) = \begin{cases} 8.29(\log 1.54h_m)^2 - 1.1, & f \leq 200\ MHz \\ 3.2(\log 11.75h_m)^2 - 4.97, & f \geq 400\ MHz \end{cases} \quad (5)$$

5 Conclusions and future work

The scope of this work has been to study and create a simulation model to evaluate Sigfox, one of the most promising LPWAN technologies. After a brief introduction about the IoT and LPWAN paradigm, we discussed the characteristics of the Sigfox communication technology and the ETSI LTN standard. Then, we implemented a NS-3 Sigfox model to simulate a whole Sigfox network, allowing us to obtain information about important specifications of the technology such as power consumption, coverage, link quality or error rate, among others. Thus, this model will be very helpful for comparative evaluation with other technologies, such as Zigbee, LTE and WiFi, being considered the suitability of them for many IoT applications

As a future work, aforementioned communication parameters will be studied, evaluating the accomplishment of offered features. In addition, it is intended to verify the model by comparison with real devices in all the three environments described (rural, suburban and urban). Finally, new proposal of communication techniques will be studied in future works.

References

1. U. Raza, P. Kulkarni and M. Sooriyabandara, "Low Power Wide Area Networks: An Overview," in IEEE Communications Surveys & Tutorials, vol. 19, no. 2, pp. 855-873, Second-quarter 2017.
URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7815384&isnumber=7936707 Accessed June 2017
2. Low Throughput Networks (LTN); Use Cases for Low Throughput Networks.
3. Low Throughput Networks (LTN); Functional Architecture.
4. Low Throughput Networks (LTN); Protocols and Interfaces.
5. Claire Goursaud, Jean-Marie Gorce. Dedicated networks for IoT : PHY / MAC state of the art and challenges. EAI endorsed transactions on Internet of Things, 2015, <10.4108/eai.26-10-2015.150597>. <hal-01231221>

6. Isaac Brown, A Detailed Breakdown of LPWAN Technologies and Providers. [Online] Available: http://web.luxresearchinc.com/hubfs/Insight_Breakdown_of_LPWAN_Technologies.pdf?t=1461874447328 Accessed June 2017
7. M. T. Do, C. Goursaud and J. M. Gorce, "On the benefits of random FDMA schemes in ultra narrow band networks," 2014 12th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), Hammamet, 2014, pp. 672-677.
URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6850364&isnumber=6850267 Accessed June 2017
8. K. E. Benson, Q. Han, K. Kim, P. Nguyen and N. Venkatasubramanian, "Resilient Overlays for IoT-Based Community Infrastructure Communications," 2016 IEEE First International Conference on Internet-of-Things Design and Implementation (IoTDI), Berlin, 2016, pp. 152-163.
URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7471359&isnumber=7471335 Accessed June 2017
9. ETSI, “System Reference document (SRdoc); Short Range Devices (SRD); Technical characteristics for Ultra Narrow Band (UNB) SRDs operating in the UHF spectrum below 1 GHz”

Choreography technologies for Interaction of Internet of Things systems

Aroa Lizondo[1], Jose-Luis Bayo-Montón[1], Antonio Martinez-Millana[1], Iván Ballester, Weisi Han, Vicente Traver[1], Carlos Fernandez[1]

[1] Instituto ITACA, Universitat Politècnica de València, Camino de Vera s/n, 46022, Valencia, Spain

Abstract. Internet of Things (IoT) defines a framework in which sensors, devices and actuators can be managed in a ubiquitous and distributed way. So far, this concept has been extensively used in several contexts (SmartCities and Farms among others) and is currently expanding to other domains.The capabilities offered by the IoT in the health field can improve people health care quality. Connecting wearable eHealth sensors and environmental sensors into the IoT paradigm could be an easy and fast way to deploy complex telemedicine interventions.

1 Introduction

A lot of architectures and platforms, which integrate multiple healthcare services in different equipment, have been proposed in recent years, aiming at providing an environment to track vital biosignals. Pervasive computing enables evolving the concept of autonomous systems that could possibly integrates various devices and technologies considering the environmental context to perform a set of rule-based actions and message exchanging [1].

This paper presents the use of a system of choreography of services, the Choreographer, which has proved to be a very useful tool [2], [3] to develop these platforms, since it allows the communication and the transfer of the monitored data among the different systems and users.

2 Materials

The Choreographer is a semantic engine, capable of connecting registered services and sensors actuators in a distributed way based on Choreography principles [4].
It uses a Service Oriented Architecture (SOA) where network resources are made available through services which can be accessed through standard methods and without the need of knowing how they were implemented internally.

C.Fernandez-Llatas and M.Guillen (Ed.):
Workshop on Innovation on
Information and Communication Technologies
(ITACA-WIICT 2017)

ISBN 978-84-697-7327-7

The Choreographer can be characterized by the following components:

- Services: represent functions within the process.
- Messages: information that is exchanged between services. They are sent using a specific eXtensible Markup Language (XML) message protocol called eXtensible MeSsaGe (XMSG).

The Choreographer allows the intercommunication among services providing tools for registering, multicast and broadcast communication, message filtering, etc.
This system allows an easy deployment, efficiency and the independence of the program allowing the intercommunication among platforms developed in different and technologies (Java, DotNet, Android).

3 Applications

To demonstrate the versatility of uses of the Choreographer, four projects carried out in recent years in which it has been used as a central core are described below.

3.1 MOSAIC

The objective of MOSAIC is **to improve existing modeling techniques applied in Diabetes** for a better risk, diagnose. Models will be improved by enriching them with datasets coming from studies and from hospital information systems [5]
This view describes the capabilities, structure, responsibilities and specifications of the MOSAIC components and how they interact among them through services. The functional view categorizes the services into three types: application, interoperability and system services.

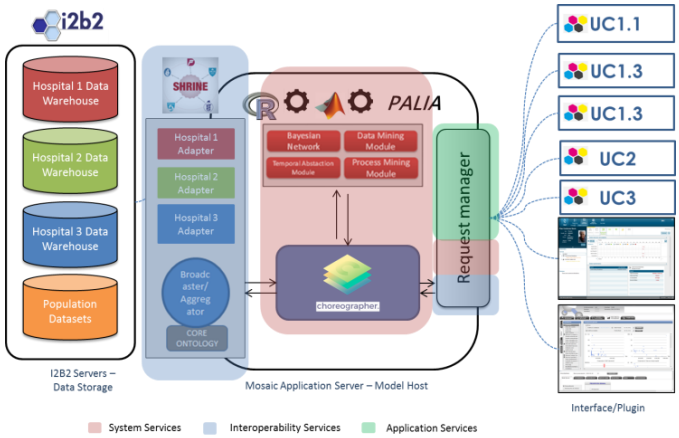


Fig. 1. MOSAIC System Architecture

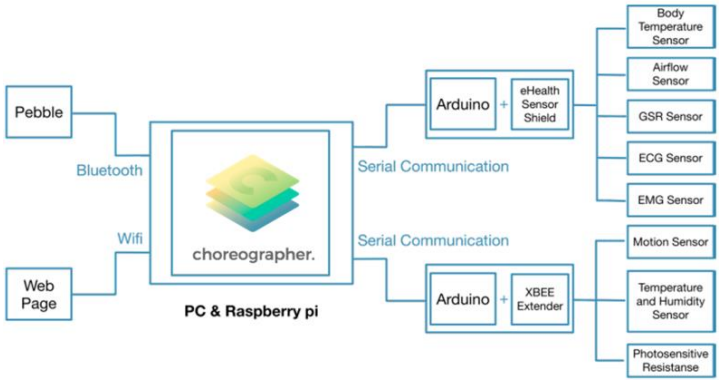
The term application encompasses all services that directly support one of the five main scenarios (aka Use Cases). Interoperability services covers all functionalities that can be reused by any component within the system (e.g. can be included in ETL processes that prepare input data for the algorithms). Finally, system services cover all logic (including functional logic and infrastructure) that is common to multiple scenarios

The core of the component is a message dispatcher engine (the Choreographer) and a data base that contains the services that are registered (declared) within it. The services may be connected to the core locally, when the services are allocated in the same computer of the choreographer core (e.g.: Model Services), or remotely by using a TCP protocol service wrapper (e.g.: ETL services)

3.2 Pervasive Healthcare Monitoring System (PHMS)

This is an embedded system with a custom lite protocol for the connection of cheap **wearable devices and environmental monitoring** devices based on prototyping eHealth solutions (Arduino, Raspberry Pi, bio-signals kit and SmartHome kit).

The whole system can be divided into two parts — Environmental information detec-



tion part and Biometric information detection part. The biometric sensors are chosen from Cooking Hacks, eHealth Biometric Sensor Platform and the environment sensors are chosen from Arduino Smart Home kit. All components are interconnected using the Choreographer.

Fig. 2. Architecture of the PHMS

On one side of the system, all of the five biometric sensors are integrated on one Arduino board to detect and collect biometric information at the same time. Three environment sensors are integrated on another Arduino board, and these sensors can run concurrently as well.

On the other side of the system, different elements are used to display the collected data (webpage). To enable the communication between the sensors and the displaying interfaces, a software is created on top of a proprietary Choreography Engine made with Microsoft .NET technology. This engine is connected to the display interfaces by using Wi-Fi and communicate with the Arduino board through serial communication. The choreographer software is deployed in two computing elements: a desktop computer and a Raspberry Pi 3.

3.3 Integration of robotic systems into an IoT platform

The intention of this project is to develop a proof-of-concept, functional social robot under the SABIEN-ITACA's framework, the Chorepgrapher.

This robot, is working under an IoT platform. This feature is used for connecting the robot to the environment, so the robot will be able to gather data from other computers and devices in the local network. The robot is able to move, walk around and perform basic tasks. You can talk to it, since it has a basic artificial intelligence. It is able to interact with the user both by speech and by text.

This project can be conceived as a first step towards a fully functional social robot able to take care of elder people, keep control of patients and send alarms and messages through the net to the doctors.

The SABEIN's Choreographer is running on a Raspberry with Microsoft's Win 10 IoT Core under the .NET framework.

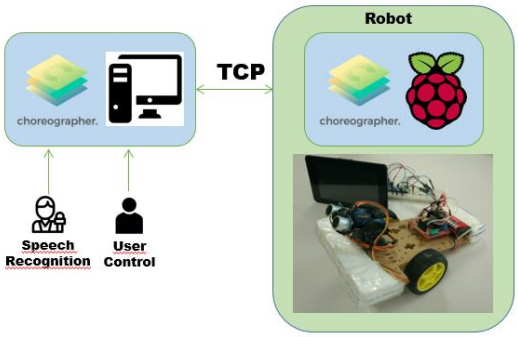


Fig. 3. Architecture of the integration of a social robot

At the diagram, there are only two Choreographers connected to each other. However, it is possible to connect as many Choreographers as desired among them. This could be useful, for example, for making the robot receive data from the computers in a hospital, download a patient's history, or updating its environmental database from a weather station.

3.4 FASyS: Absolutely Safe and Healthy Factory

The project aims to guarantee in an integrated way the safety and the continuous well-being of the worker in the factories of the future, allowing the workers to become the key factors of competitiveness and differentiation of the new production model.
One of the biggest problems in achieving this is the lack of a complete set of information regarding the worker's health. Therefore, Fasys eliminates timely follow-up and bets on continuous health monitoring.

The technology allows to collect a sufficiently complete set of information to achieve a holistic care of the worker, achieving continuous and personalized monitoring. This Fasys vision generates a large amount of data (for example, worker monitoring, environment monitoring, medical decision support systems, systems of action protocols, treatment systems and adherence to therapy).

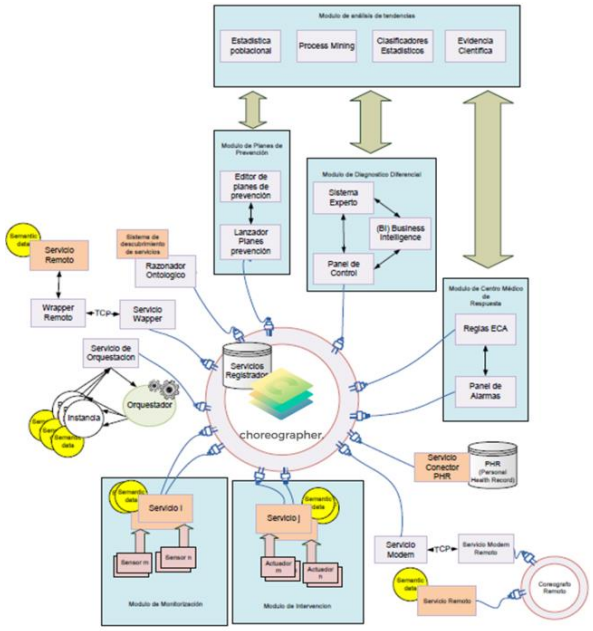


Fig. 4. FASyS System Architecture

The architecture will get the interoperability of the different existing blocks in the previous scheme. The previous figure shows, graphically, the scheme of the proposed architecture for the Fasys decision support system. This architecture is formed by the choreographer as the central core, to which are connected the different modules and subsystems that are necessary to carry out a continuous monitoring of the worker's health in the Fasys environment.

4 Conclusions

With the presentation of these four projects in which the Choreographer is used in health environments we highlight the usefulness and the possibilities of this tool.

Acknowledgments

Author Aroa Lizondo García: This work has been funded by the Ministry of Economy and Competitiveness: Program for the Promotion of Young Employment and Implementation of the Youth Guarantee 2014 (PEJ-2014-A-06813).
The subsidized activity is part of the National System of Youth Guarantee and is co-financed in the framework Of the Youth Employment Operational Program, with financial resources from the Youth Employment Initiative (IEJ) And the European Social Fund (ESF) for the 2014-2020 programming period.

References

1. Webber, T.; César M.; Leonardo A.A.; Rubem D. R.; Letícia B.P.; Pervasive Computing Integration on Healthcare Environments, (2012).
2. Fernandez-Llatas, C.; Mocholí, J.B.; Sanchez, C; Sala, P; Naranjo, J.C, Process choreography for Interaction simulation in Ambient Assisted Living environments (2010)
3. Fernández-Llatas, C.; Mocholí, J.B.; Moyano, A.; Meneu, T. Semantic Process Choreographyfor Distributed Sensor Management. SSW (2010) 32–37.
4. Blanc, S.; Bayo-Monton, J.L.; Campelo, J.C.; Fernandez-Llatas, C., Process Choreography in Wireless Sensor and Actuator Networks: a proposal to achieve Network Virtualization
5. A. Martinez-Millana et al., "From data to the decision: A software architecture to integrate predictive modelling in clinical settings," 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan (2015), 8161-8164.

WBAN-IBC COMMUNICATIONS: Approaches, Perspectives and Applications.

Juan Félix Villegas Méndez, José Carlos Campelo Rivadulla and Alberto Miguel Bonastre Pina.

[1] Instituto ITACA, Universitat Politècnica de València, Camino de Vera s/n, 46022, València, España,
jufelvi91@gmail.com, {jcampelo, bonastre}@itaca.upv.es

Abstract. In the field of networks, a Wireless Body Area Network (WBAN) has been defined as a network designed for the communication of different low power wireless devices located in and over the human body. Those may be used for different purposes, such as the implementation of small-sized sensors, for the monitoring of vital parameters of the human body or control and authentication applications. All of them need to exchange information between these devices, besides to being able to send their data to an external network.
One of the network specification of the IEEE 802.15.6 standard, called HBC, uses the human body itself as a channel for signal transmission by means of the electrostatic field coupling technique. This technique has also been called Body Channel Communication (BCC) and Intra Body Communication (IBC).
This paper summarizes the identifying, analyzing and synthesizing studies that have been elaborated about WBAN, considering the IEEE 802.15.6 standard, as well as the basics of IBC networks and the proposed applications for this kind of networks.

Key words: Wireless Body Area Network (WBAN), Body Area Network (BAN), Ultrawide Band (UWB), Narrow Band (NB), Intrabody Communications (IBC), Human Body Communications (HBC), transmitter, receiver, transceiver, IEEE 802.15.6., transmission through the skin.

1 Introduction

Body area networks have become a promising line of research, and this is most evident with the increased use of wireless networks and the constant miniaturization of electronic devices, which has led to a new approach of BAN networks: the Wireless Body Area Network (WBAN). They enable devices to be used to adhere to the clothing, the human body or even be implanted under the skin. These devices are capable of establishing wireless communications links, in order to be used in different applications, such as those related to health, entertainment, and even authentication and security.

A WBAN consists, typically, of a set of interconnected wireless sensor nodes, which send their information to a personal device (PD), which is responsible for collecting

C.Fernandez-Llatas and M.Guillen (Ed.):
Workshop on Innovation on
Information and Communication Technologies
(ITACA-WIICT 2017)

ISBN 978-84-697-7327-7

information from the sensor nodes and allows the connection of the body network with other networks, the so-called *interbody networks*. The personal device PD is denoted as BCU (Body Control Unit), Body Gateway or Personal Server (PS).

IEEE 802 Committee has established a Working Group called IEEE 802.15.6 for the standardization of WBAN [1]. Its purpose is to establish a communication standard for low-power devices, interconnected by means of this type of networks. This standard recognizes three physical layer specifications for WBAN: Narrow Band PHY (NB), Ultra Wideband Physical Layer (UWB), and the physical layer of communications across the human body: Human Body Communications (HBC).

This paper has been structured in five sections, as follows:

Section 2 describes an analysis of WBAN communications, synthesizing the IEEE 802.15.6 standard. Section 3 presents a synthesis of the existing PHY layers recognized by the IEEE 802.15.6 standard for WBANs. Section 4 also shows the upper link layer, or MAC layer for WBAN networks.

Section 5 presents the basics about IBC, considering existing studies on the characteristics of the skin as a communication channel, IBC channel models that have been proposed, besides exposing the techniques of capacitive coupling and galvanic coupling of the electrodes to the human body. Some proposed applications for IBC communication systems are also presented. Finally, section 6 presents the conclusions.

2 Wireless Body Area Networks

WBANs can be defined as a collection of low-power, miniaturized nodes, capable of establishing a wireless link, that monitor the functions of a human body and its environment [2], however, applications of this type of networks can go beyond the monitoring of body functions, and focus on other fields such as security and entertainment.

A wireless body area network is basically composed of small devices adhered or implanted to the human body, capable of establishing a wireless communication link between them. A general specification of the devices can be made, classifying them into sensors and actuators [3]. The former are designed to measure certain parameters of the human body, such as heart beats, body temperature. The actuators perform a determined action according to the measurements provided by the sensors, or through some action that the user determines, for example, there are actuators that allow the supply of insulin automatically.

It is important to highlight the difference of a WBAN with a WPAN (Wireless Personal Network), since the latter is a network in the environment around the person, reaching distances from tens, and sometimes hundreds, of meters. A WBAN is a network designed for the interconnection of devices deployed in or over a human body, with distances that can reach up to 2 meters. A WPAN network uses technologies such as Bluetooth (IEEE 802.15.1) or ZigBee (IEEE 802.15.4). A WBAN network uses the NB (Narrow Band), UWB (Ultra Wide Band) and BCC (Body Channel Communication) links, as specified in the IEEE 802.15.6 standard.

The IEEE 802.15.6 standard mentions that current personal area networks (PANs) do not fulfill medical requirements for proximity to human tissues and communication requirements relevant to some application environments [1]. In addition, do not support the combination of reliability, QoS (quality of service), low power, data rate and non-interference required to address the range of Body Area Network (BAN) applications.

It is possible to use the techniques designed within the traditional wireless sensor networks (WSN) [4]; however, WBAN networks have unique and important characteristics, regarding energy limitation, density and displacement of sensors, propagation characteristics, data rates, latency, mobility, safety and security.

2.1 WBAN Architecture.

The WBAN architecture can be classified into three levels, illustrated in **Fig 1**.

- **Intra-WBAN Communication:** Communications with ranges of up to two meters between devices attached, or implanted to the human body or devices around it.
- **Inter-WBAN Communication:** Communications between the BCU device and other access points (APs), which may be part of the infrastructure, or be dynamically located to respond to emergencies.
- **Beyond-WBAN Communication:** It is mainly focused on metropolitan areas, where, through a Gateway, allows the connection of the previous levels to networks like Internet, or cellular networks.

According to the IEEE 802.15.6 Standard [14], WBAN nodes can be classified into three subcategories, as follows:

- **Surface Nodes:** adhered to the surface of the human body, on the skin, or at a maximum distance of 2 centimeters from the human body.
- **External nodes:** they do not have direct contact with the surface of the human body, and they are located at a distance of few centimeters of the human body, until a maximum distance of 5 meters from it.
- **Implanted nodes:** These are implanted directly under the skin or within other tissues of the human body.

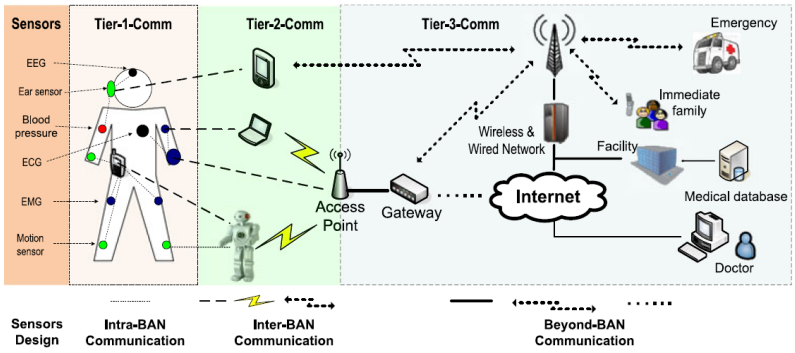


Fig 1. WBAN architecture [5].

2.2 The IEEE 802.15.6. Standard.

The 802.15.6 standard was approved and published in 2012. It was designed for wireless communications around or within a human body [6]. It was defined as the short-range wireless communications standard in or near a body, but not limited to humans.

The frequencies, in which these communications operate, according to the same standard, are in the ISM bands (Industrial, Scientific and Medical). In addition to meeting stringent criteria to avoid interference, the standard states that support for quality of service (QoS), extremely low power and data rates of up to 10 Mbps should be included. The standard also considers the effects of transmission and reception of information due to the presence of a person, considering also the changes in these characteristics due to the movement of the users.

Network Topology.

According to the IEEE 802.15.6 standard, nodes must be organized in star topologies, with one or two hops [7]. There will be a unique node acting as a hub (that may be called Coordinator, BCU, PD, PS or Body Gateway) for each WBAN.

The exchange of data frames in a star topology at one hop is performed directly between the nodes and the hub. In a two-hop star network topology, the exchange of frames to the hub can be performed through a retransmission node. **Fig 2** shows the WBAN topologies.

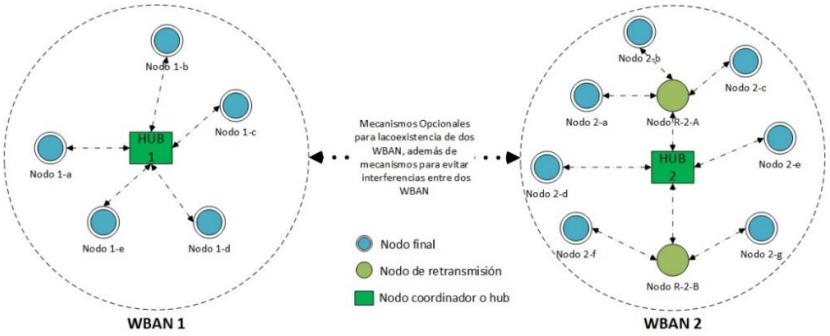


Fig 2. WBAN topologies according the IEEE 802.15.6. Standard.

Reference model for WBAN.

The IEEE 802.15.6 standard defines a medium access control layer (MAC) for three different physical layers (PHY). These physical layers use different frequency ranges. They are Narrow Band (NB), Ultra-Wide Band (UWB) and a physical layer specifically designed for Intrabody communications (IBC, HBC, BCC), as shown in **Fig 3**. The first two operate through Radio waves, using the frequency bands determined by the standard, but the latter is a designed for communications through human tissues, essentially the skin.

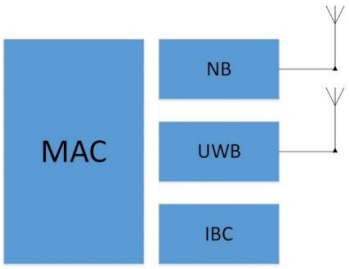


Fig 3. MAC and PHY layers for WBAN, according the IEEE 802.15.6 standard.

3 WBAN PHY Layer.

The IEEE 802.15.6 standard supports three PHY operational layers, two mandatory and one optional. The PHY UWB (Ultra Wide Band) and PHY HBC (Human Body Communication) layers are mandatory; the PHY NB layer (Narrow Band) is optional. According with the application (medical or non-medical, over the body, with implanted sensors or communications with devices close to the body), a PHY layer should be chosen.

Narrow Band PHY.

The use of the PHY NB layer (Narrow Band) is oriented to the communication of devices over a human body or implanted nodes [8]. The PHY NB layer can be used for short-range bi-directional wireless transmission between devices in medical applications. Such devices must be designed to be highly reliable in the ISM bands [9].

The PHY NB operation bands are: 402 to 405 MHz, 420 to 450 MHz, 863 to 870 MHz, 9092-928 MHz, 950 to 958 MHz, 2360 to 2400 MHz and 2400 to 2483.5 MHz.

Depending on the frequency, the data rates that NB supports have a range from 57.5 kbps to 971.4 kbps, very close for a data rate of 1 Mbps.

A physical frame is called PPDU (Physical Protocol Data Unit). In PHY NB has three main components: a PLCP preamble, a PLCP header and the PSDU (Physical Layer Service Data Unit). The first two are control fields are used to synchronize the transmission and to identify transmission parameters as the data rate, the payload length, among others. The PSDU is the last component of a PPDU frame, contains a MAC header, the payload or MAC body, and a Frame Check Sequence (FCS).

Ultrawide Band PHY.

UWB physical layer defines eleven channels, numbered from 0 to 10. Each channel has a bandwidth of 499.2 MHz. In addition, two frequency bands are defined for PHY UWB, the so-called *High Band* and *Low Band*. Channels numbered from 0 to 2 are assigned to the *Low Band*, with center frequencies at 3494.4 MHz, 3993.6 MHz and 4492.8 MHz; respectively. The *High Band* is formed by the channels 3 to 10, whose central frequencies are located at 6489.6 MHz; 6988.8 MHz; 7488.0 MHz; 8486.4 MHz; 8985.6 MHz, 9484.8 MHz and 9984.9 MHz, for each of the channels. The available data rates for PHY UWB vary from 395 kbps to 12,636 Mbps [10].

Three fields conform the PPDU frame for PHY UWB: a synchronization header (SHR), physical layer header (PHR) and the respective Physical Layer Service Data Unit (PSDU), which contains the payload.

Human Body Communications PHY.

The physical layer of HBC (PHY HBC) uses a communication technology called Electric Field Communication (EFC). HBC does not use RF propagation [11], so it does not require antennas (NB and UWB does), but it requires the use of electrodes attached to the skin, to use the human body as a communications channel.

Regarding the frequency band of operation for PHY HBC, the standard only mentions that a compatible device must be able to support data transmissions and receptions in the 21 MHz band [1]. Some authors mention that HBC operates with a bandwidth of 5.25 MHz [6], while others mention that the bandwidth is 4 MHz, and operates in two frequency bands, centered at 16 MHz and 27 MHz [12] [13]. Recent studies of the standard ratify the use of the 21 MHz band for HBC, with scalable data rates from 164 to 1312 kbps [14].

Similar to the NB PHY layer, the HBC PHY layer encapsulates the PSDU frame with the payload information, adding control and error correction fields, as well as bits for framing. Fig 4 details the structure of a HBC PHY frame.

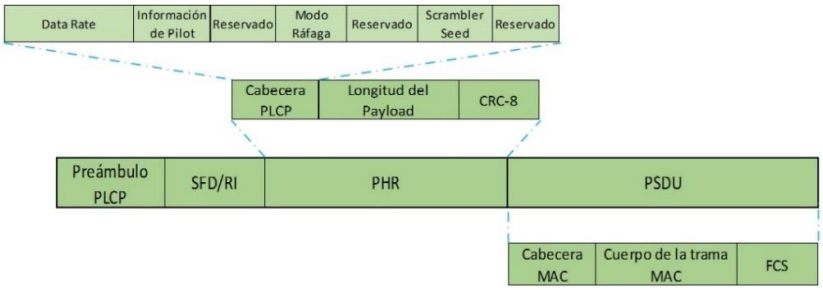


Fig 4. Structure of HBC PHY frames, according to the IEEE 802.15.6 standard.

4 MAC Layer.

The IEEE 802.15.6 standard proposes a single MAC layer, above the three physical layers. The MAC layer must be flexible, in addition to supporting contention-based control mechanisms and free contention mechanisms [15].

A MAC frame consists of three fields: a MAC header that is conformed by 56 bits or 7 octets, a frame body, containing the payload or variable length, with a maximum size of 255 bytes, i.e. per each frame can be transferred up to 255 bytes of useful information. The MAC frame is concatenated with an FCS (Frame Check Sum) sequence of 18 bits in length. Fig 5 describes the structure of a MAC frame.

The MAC header is composed of other subfields: ***Frame Control***, consisting of 32-bit control, contains information about the type of frame being sent. ***Receiver identifier***, an 8-bit field for an abbreviated address of the frame receiver, this may be the address of a node or NID, or the address of the hub, or HID. ***Transmitter Identifier***, an 8-bit field for an abbreviated address of the frame transmitter, may be the address of a node or NID, or the address of the hub, or HID. Finally a ***BAN Identifier***, which is an 8-bit field for an abbreviated address that identifies the BAN or WBAN in general.

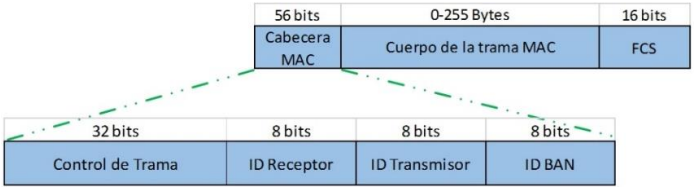


Fig 5. Format of a MAC frame, according to the IEEE 802.15.6 standard.

According to the IEEE 802.15.6 standard, in the MAC layer, three categories of medium access mechanisms are defined [12]. The first one is Random Access Mechanism, where the hub can use a protocol for allocating resources such as slotted ALOHA or CSMA/CA (Carrier Sensor Multiple Access / Collision Avoidance), which will depend on the PHY layer used. When using a UWB PHY layer, the hub considers the ALOHA

protocol. If it is NB PHY, the hub considers the CSMA / CA protocol. No protocol has been established for the PHY HBC layer.

The second one is called un-programmed and improvised access mechanism, where the hub sends query orders or publish orders without prior notice. For this, it uses mechanisms of allocation of publication resources (downlink) or consultation (uplink).

Finally, the third one is the Programmed Access and per query Scheduled Access Mechanism. Where the scheduled access mechanism is used to achieve scheduled assignments of uplink, downlink and bilink type; and the programmed access mechanism is used for query or publication assignments.

5 Intrabody Communications

IBC emerged because communication between devices require low power consumption, information security, frequency reuse and resistance to interference, besides allowing portability and not obstructing the mobility of the user, so, an IBC device requires a small form factor.

Although the standard defines HBC communications as those that use the electric field communication technology (EFC), also called capacitive coupling, various articles and publications proposed by other authors, defines another transmission mechanism, called galvanic coupling.

IBC (Intrabody Communication) is referred as a transmission technique for electrical signals based on the use of the human body as a physical transmission medium [16]; Fig 6 illustrates the connection scheme between the nodes and the hub, and in turn between the hub and an inter-body network.

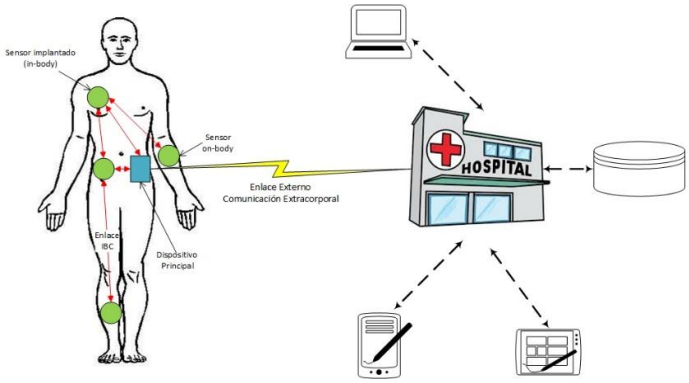


Fig 6. IBC network scheme.

5.1 IBC Transmission Systems.

There are two transmission systems for IBC: capacitive coupling and galvanic coupling. In the first method, the electric signal is controlled by an electric potential, while in the second method it is controlled by a differential current flow.

Capacitive Coupling. Researchers in the department of MIT discovered the capacitive coupling of the human body to its surroundings, where certain parts of the near field can be exploited to make the human body act as a medium for data transmission [17].

Thomas G. Zimmerman reported in [18] and [19] capacitive coupling as a system for PAN personal area networks. He exposed the near electric field that is produced by a transmitter attached to the body through an electrode, leaving the second electrode loose as a ground connection. At the other side, a receiver is attached to the body in the same way.

Fig **7** illustrates the electric fields produced by an IBC transmitter attached to the body. The couplings between the electrodes, the air and the external earth can be modeled as capacitances, which is why it is called the capacitive coupling [20]. As one of the paths of the signal is the capacitive return, the technique becomes dependent on external conditions [21].

The capacitive coupling system is basically formed by the following elements, illustrated in

Fig **7**. On the side of the transmitter an encoder, a coupler and the respective electrodes coupling to the human body and the ground electrode. For its part, the receiver will consist of two electrodes, as well as the transmitter, a signal detector that passes through the human body and through the ground return path, and a decoder.

Galvanic Coupling. This coupling method is achieved by coupling an alternating current in the human body [11]. Two electrodes are attached to the human body, both at the transmitter and at the receiver, and a ground reference is not required as in the capacitive coupling method. The signal is applied differentially on the two electrodes, so that two current flows, the first between the electrodes of the transmitter end. The other current flow is of smaller magnitude, and is established between both ends of the system. The information carriers of the galvanic coupling will be ionic fluids as a result of the injection of the differential currents through the electrodes where the human body can be visualized as a transmission line or waveguide.

Based on the studies of the flowing currents, product of the galvanic coupling, the same blocks as the capacitive coupling can be used: an encoder and a coupler in the transmitter; and a detector and the respective decoder in the receiver. As shown in Fig 8, in this case both electrodes are attached to the human body, which acts as a transmission line, and ground electrodes are not required.

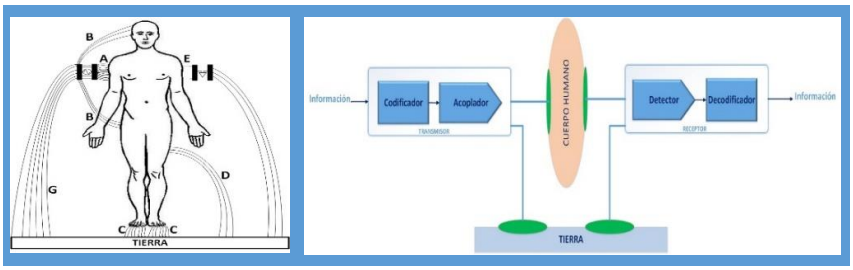


Fig 7. Capacitive coupling method.

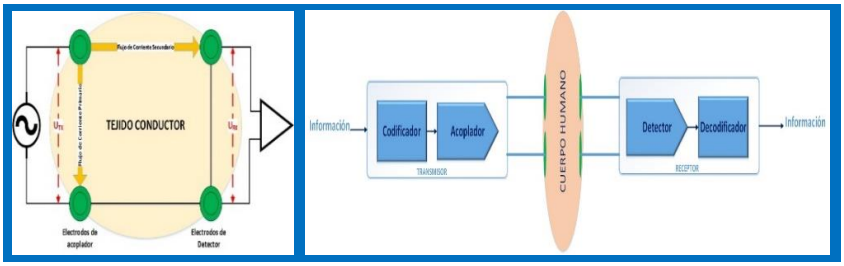


Fig 8. Galvanic coupling method

The studies developed, based on the capacitive coupling systems, have allowed the development of transceiver devices that can reach a data transmission rate above 10 Mbps, and with low power consumption. In the case of galvanic coupling devices can reach a data transmission rate bordering 2 Mbps, and with low power consumption.

The galvanic coupling system is not dependent on the environment around the human body, and some authors report that it is possible to use it for the communication of devices attached to the body and implanted devices [22].

5.2 IBC Applications.

According to the coupling mechanisms presented above, the conclusions of the authors have pointed out that the capacitive coupling systems respond effectively to applications that require longer transmission distances, however, their main disadvantage is that they are susceptible to Interference, such as radiation from other nearby devices. On the other hand, galvanic coupling systems have the advantage of being more robust against interference and body movements; however, they are limited when larger transmission distances are required.

Initially, the applications were medical oriented. Subsequently, there were proposals for security and authentication applications. Finally, the increase in transmission rates allowed proposals for new applications in the area of domestic consumption and entertainment.

Shinigawa et al. proposed different applications for IBC, and they performed an IBC network, based on capacitive coupling, that supports half-duplex communications at a transmission rate of 10 Mbps, called RedTacton, which was reported in [23]. It was designed to be able to establish contact communications, so they proposed multiple applications, as shown in Fig 9. However, there was no continuity to these studies, so that commercial products using RedTacton were not materialized, therefore these applications remained only as proposals for future implementations of the IBC communications.

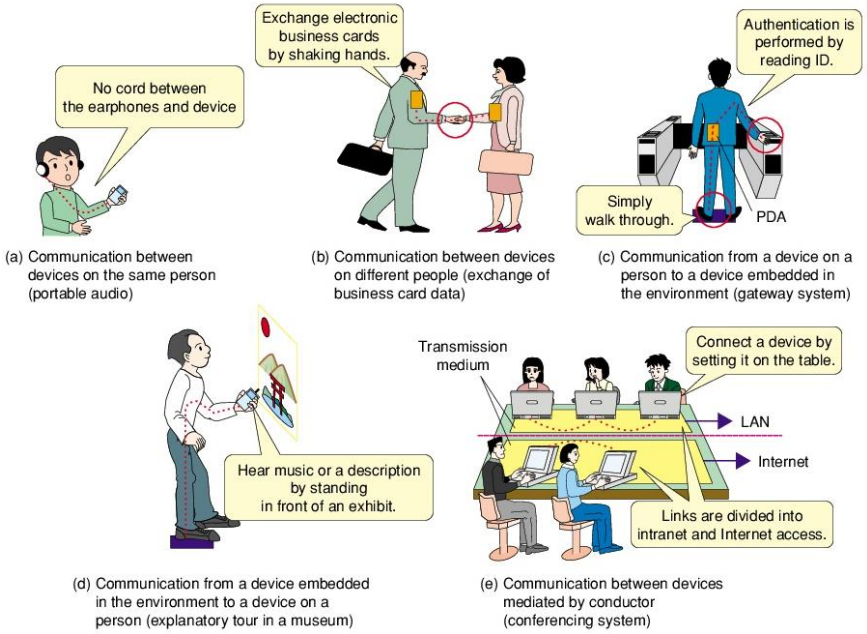


Fig 9. Applications proposed for IBC by RedTacton [23].

6 Conclusions

This article responds to the goal of being able to realize a survey of WBANs, focused mainly in the basis of IBC networks. The work approach, from the perspective of computer and network engineering, has allowed not only an analysis of various publications containing information on WBAN-IBC networks, but also a breakdown of the applications proposed by different authors, for IBC transceivers designed and implemented by them under various parameters and technologies.

It is important to consider the emergence of WBANs, which emerged with a medical orientation, and through different publications, have acquired other approaches to its application, seeking to solve the difficulties of Personal Area Networks (PANS).

Finally, it is important also, to consider the IEEE 802.15.6 standard as a starting point for the study of the WBANs, as well as the approach to the IBC networks that performs this work. From standardized devices there will be possible to achieve applications with not proprietary technologies, and it will be possible to implement higher layers over the first two layers (MAC and PHY) mentioned in the standard.

References

[1] IEEE Standards Association and others, *802.15. 6-2012 IEEE Standards for Local and Metropolitan Area Networks--Part 15.6: Wireless Body Area Networks*, IEEE, 2012.

[2] Ullah, Sana and Higgins, Henry and Braem, Bart and Latre, Benoit and Blondia, Chris and Moerman, Ingrid and Saleem, Shahnaz and Rahman, Ziaur and Kwak, Kyung Sup, «A comprehensive survey of wireless body area networks,» *Journal of medical systems*, vol. 36, nº 3, pp. 1065-1094, 2012.

[3] Latré, Benoît and Braem, Bart and Moerman, Ingrid and Blondia, Chris and Demeester, Piet, «A survey on wireless body area networks,» *Wireless Networks*, vol. 17, nº 1, pp. 1-18, 2011.

[4] Arya, Aashima and Bilandi, Naveen, «A review: Wireless body area networks for health care,» *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 2, nº 4, pp. 3800-3806, 2014.

[5] Chen, M., Gonzalez, S., Vasilakos, A., Cao, H., & Leung, V. C., «Body area networks: A survey,» *Mobile networks and applications*, vol. 16, nº 2, pp. 171-193, 2011.

[6] Li, Miaoxin and Zhuang, Mingjie, «An overview of physical layers on wireless body area network,» de *Anti-Counterfeiting, Security and Identification (ASID), 2012 International Conference on*, IEEE, 2012, pp. 1-5.

[7] Miček, Juraj and Karpiš, Ondrej and Ševčík, Peter, «Body area network: analysis and application areas,» *Int. J. Eng. Res. Dev.(IJERD)*, vol. 6, nº 8, pp. 22-26, 2013.

[8] A. Reichman, «Standardization of body area networks,» de *Microwaves, Communications, Antennas and Electronics Systems, 2009. COMCAS 2009. IEEE International Conference on*, IEEE, 2009, pp. 1-4.

[9] Seyedi, Mir Hojjat and Lai, Daniel Tze Huei, «A Novel Intrabody Communication Transceiver for Biomedical Applications,» 2014.

[10] Saleem, Shahnaz and Ullah, Sana and Yoo, Hyeong Seon, «On the security issues in wireless body area networks.,» *JDCTA*, vol. 3, nº 3, pp. 178-184, 2009.

[11] Roa Romero, Laura María, Reina Tosina, Luis Javier, Callejón Leblic, María Amparo, Naranjo Hernández, David, Estudillo Valderrama, Miguel Ángel, «Intrabody Communication,» de *Handbook of Biomedial Telemetry*, Willey, 2014, pp. 252-300.

[12] Movassaghi, Samaneh and Abolhasan, Mehran and Lipman, Justin and Smith, David and Jamalipour, Abbas, «Wireless body area networks: A survey,» *IEEE Communications Surveys \& Tutorials*, vol. 16, nº 3, pp. 1658-1686, 2014.

[13] Kwak, Kyung Sup and Ullah, Sana and Ullah, Niamat, «An overview of IEEE 802.15. 6 standard,» de *Applied Sciences in Biomedical and Communication Technologies (ISABEL), 2010 3rd International Symposium on*, IEEE, 2010, pp. 1-6.

[14] Seyedi, MirHojjat and Kibret, Behailu and Lai, Daniel TH and Faulkner, Michael, «A survey on intrabody communications for body area network applications,» *IEEE Transactions on Biomedical Engineering*, vol. 60, nº 8, pp. 2067-2079, 2013.

[15] Bae, J., Song, K., Lee, H., Cho, H., & Yoo, H. J., « A low-energy crystal-less double-FSK sensor node transceiver for wireless body-area network,» *IEEE Journal of Solid-State Circuits*, vol. 47, nº 11, pp. 2678-2692, 2012.

[16] Lee, Junghyup and Kulkarni, Vishal Vinayak and Ho, Chee Keong and Cheong, Jia Hao and Li, Peng and Zhou, Jun and Da Toh, Wei and Zhang, Xin and Gao, Yuan and Cheng, Kuang Wei and others, «30.7 A 60Mb/s wideband BCC transceiver with 150pJ/b RX and 31pJ/b TX for emerging wearable applications,» de *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2014 IEEE International*, IEEE, 2014, pp. 498-499.

[17] Ghamari, M., Arora, H., Sherratt, R. S., & Harwin, W., «Comparison of low-power wireless communication technologies for wearable health-monitoring applications.,» de *Computer, Communications, and Control Technology (I4CT), 2015 International Conference on*, IEEE, 2015, pp. 1-6.

[18] T. G. Zimmerman, «Personal Area Networks (PAN): Near-Field Intra-Body Communication,» Massachusetts Institute of Technology, Massachusetts, 1995.

[19] T. G. Zimmerman, «Personal area networks: near-field intrabody communication.,» *IBM systems Journal, 35(3.4)*, pp. 609-617, 1996.

[20] M. A. Callejón Leblic, «Contribuciones a la caracterización del cuerpo humano como medio de transmisión en las técnicas de comunicaciones intracorporales.,» Sevilla, 2016.

[21] H. Schwan, «Electrical properties of tissues and cell suspensions: mechanisms and models,» de *Engineering in Medicine and Biology Society, 1994. Engineering Advances: New Opportunities for Biomedical Engineers. Proceedings of the 16th Annual International Conference of the IEEE*, vol. 1, IEEE, 1994, pp. A70-A71.

[22] Gabriel, Camelia and Gabriel, Sami and Corthout, E, «The dielectric properties of biological tissues: I. Literature survey,» *Physics in medicine and biology*, vol. 41, nº 11, p. 2231, 1996.

[23] Shinagawa, Mitsuru and Ochiai, Katsuyuki and Sakamoto, Hideki and Asahi, Toshiaki, «Human area networking technology: Redtacton,» *NTT Technical Review*, vol. 3, nº 5, pp. 41-46, 2005.

[24] M. S. Wegmüller, «Intra-body communication for biomedical sensor networks,» ETH ZURICH, 2007.

[25] I. G. Lim, H. I. Park, T. W. Kang, C. H. Hyoung , S. W. Kang, J. H. Hwang, K. S. Kim, J. B. Kim, S. E. Kim, J.-k. Kim, H. J. Myoung, S. B. Hyun, K. H. Park, B. G. Choi y T. Y. Kang, «Frequency selective digital transmission apparatus». Patente US Patent 8,422,535, 16 April 2013.

[26] Nie, Zedong and Liu, Yuhang and Duan, Changjiang and Ruan, Zhongzhou and Li, Jingzhen and Wang, Lei, «Wearable biometric authentication based on human body communication,» de *Wearable and Implantable Body Sensor Networks (BSN), 2015 IEEE 12th International Conference on*, IEEE, 2015, pp. 1-5.

[27] Li, Huan-Bang and Takizawa, Kenichi and Kohno, Ryuji, «Trends and standardization of body area network (BAN) for medical healthcare,» de *Wireless Technology, 2008. EuWiT 2008. European Conference on*, IEEE, 2008, pp. 1-4.

[28] Hyoung, Chang-Hee and Kang, Sung-Weon and Park, Seong-Ook and Kim, Youn-Tae, «Transceiver for human body communication using frequency selective digital transmission,» *ETRI journal*, vol. 34, nº 2, pp. 216-225, 2012.

[29] Song, Yong and Hao, Qun and Zhang, Kai and Wang, Ming and Chu, Yifang and Kang, Bangzhi, «The simulation method of the galvanic coupling intrabody communication with different signal transmission paths,» *IEEE Transactions on Instrumentation and Measurement*, vol. 60, nº 4, pp. 1257-1266, 2011.

[30] Bae, Joonsung and Cho, Hyunwoo and Song, Kiseok and Lee, Hyungwoo and Yoo, Hoi-Jun, «The signal transmission mechanism on the surface of human body for body channel communication,» *IEEE Transactions on microwave theory and techniques*, vol. 60, nº 3, pp. 582-593, 2012.

[31] Cho, Namjun and Yoo, Jerald and Song, Seong-Jun and Lee, Jeabin and Jeon, Seonghyun and Yoo, Hoi-Jun, «The human body characteristics as a signal transmission medium for intrabody communication,» *IEEE transactions on microwave theory and techniques*, vol. 55, nº 5, pp. 1080-1086, 2007.

[32] Bae, Joonsung and Yoo, Hoi-Jun, «The effects of electrode configuration on body channel communication based on analysis of vertical and horizontal electric dipoles,» *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, nº 4, pp. 1409-1420, 2015.

[33] Gabriel, Sami and Lau, RW and Gabriel, Camelia, «The dielectric properties of biological tissues: III. Parametric models for the dielectric spectrum of tissues,» *Physics in medicine and biology*, vol. 41, nº 11, p. 2271, 1996.

[34] Gabriel, S and Lau, RW and Gabriel, Camelia, «The dielectric properties of biological tissues: II. Measurements in the frequency range 10 Hz to 20 GHz,» *Physics in medicine and biology*, vol. 41, nº 11, p. 2251, 1996.

[35] Ramli, Sofia Najwa and Ahmad, Rabiah, «Surveying the wireless body area network in the realm of wireless communication,» de *Information Assurance and Security (IAS), 2011 7th International Conference on*, IEEE, 2011, pp. 58-61.

[36] Hachisuka, Keisuke and Terauchi, Yusuke and Kishi, Yoshinori and Sasaki, Ken and Hirota, Terunao and Hosaka, Hiroshi and Fujii, Katsuyuki and Takahashi, Masaharu and Ito, Koichi, «Simplified circuit modeling and fabrication of intrabody communication devices,» *Sensors and actuators A: physical*, vol. 130, pp. 322-330, 2006.

[37] Chen, Xi Mei and Mak, Peng Un and Pun, Sio Hang and Gao, Yue Ming and Vai, Mang I and Du, Min, «Signal transmission through human muscle for implantable medical devices using galvanic intra-body communication technique,» de *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, IEEE, 2012, pp. 1651-1654.

[38] Wegmueller, Marc Simon and Oberle, Michael and Felber, Norbert and Kuster, Niels and Fichtner, Wolfgang, «Signal transmission by galvanic coupling through the human body,» *IEEE Transactions on Instrumentation and Measurement*, vol. 59, nº 4, pp. 963-969, 2010.

[39] Kohno, Ryuji and Hamaguchi, Kiyoshi and Li, Huan-Bang and Takizawa, Kenichi, «R&D and standardization of body area network (BAN) for medical healthcare,» de *Ultra-Wideband, 2008. ICUWB 2008. IEEE International Conference on*, vol. 3, IEEE, 2008, pp. 5-8.

[40] Pun, Sio Hang and Gao, Yue Ming and Mak, PengUn and Vai, Mang I and Du, Min, «Quasi-static modeling of human limb for intra-body communications with experiments,» *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, nº 6, pp. 870-876, 2011.

[41] Yin, Xuefeng and Chen, Jiajing and Tian, Meng and Zhang, Nan and Zhong, Zhimeng and Lu, Stan X, «Personal authentication using the fingerprints of intra-body radio propagation channels,» de *Medical Information and Communication Technology (ISMICT), 2013 7th International Symposium on*, IEEE, 2013, pp. 159-163.

[42] Oguma, Hisashi and Nobata, Naoki and Nawa, Kazunari and Mizota, Tsutomu and Shinagawa, Mitsuru, «Passive keyless entry system for long term operation,» de *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2011 IEEE International Symposium on a*, IEEE, 2011, pp. 1-3.

[43] Ng, Kwan-Hoong, «Non-ionizing radiations--sources, biological effects, emissions and exposures,» de *Proceedings of the international conference on non-ionizing radiation at UNITEN*, 2003.

[44] Digilent Inc., *Nexys 4™ FPGA Board Reference Manual,* Digilent Inc., 2016.

[45] Callejón, M Amparo and Reina-Tosina, Javier and Naranjo-Hernández, David and Roa, Laura M, «Measurement issues in galvanic intrabody communication: Influence of experimental setup,» *IEEE Transactions on Biomedical Engineering*, vol. 62, nº 11, pp. 2724-2732, 2015.

[46] Song, Seong-Jun and Lee, Seung Jin and Cho, Namjun and Yoo, Hoi-Jun, «Low power wearable audio player using human body communications,» de *Wearable Computers, 2006 10th IEEE International Symposium on*, IEEE, 2006, pp. 125-126.

[47] Pan American Health Organization, «La cantidad de personas mayores de 60 años se duplicará para 2050; se requieren importantes cambios sociales,» Pan American Health Organization, 30 Septiembre 2015. [En línea]. Available: http://www.paho.org/hq/index.php?option=com_content&view=article&id=91&Itemid=220&lang=es. [Último acceso: 23 Febrero 2017].

[48] Kim, Sang Don and Lee, Seung Eun, «Intra-body Communication Modem on FPGA with AHB-Lite Bus Interface,» *淡江理工學刊*, vol. 18, nº 4, pp. 381-385, 2015.

[49] Barnes, FS and Greenebaum, B, «Handbook of Biological Effects of Electromagnetic Fields. Bioengineering and Biophysical Aspects of Electromagnetic Fields, ; CRC,» *CRC/Taylor \& Francis*, pp. 115-152, 2007.

[50] Postow, Elliot and Polk, Charles, Handbook of biological effects of electromagnetic fields, Boca Raton: CRC Press, 1996.

[51] Shimamoto, Shigeru and Alsehab, Adbullah M and Kobayashi, Nao and Dovchinbazar, Dagvadorj and Ruiz, Jordi Augud, «Future applications of body area communications,» de *Information, Communications \& Signal Processing, 2007 6th International Conference on*, IEEE, 2007, pp. 1-5.

[52] Xu, Ruoyu and Ng, Wai Chiu and Zhu, Hongjie and Shan, Hengying and Yuan, Jie, «Equation environment coupling and interference on the electric-field intrabody communication channel,» *IEEE Transactions on biomedical engineering*, vol. 59, nº 7, pp. 2051-2059, 2012.

[53] Partridge, Kurt and Dahlquist, Bradley and Veiseh, Alireza and Cain, Annie and Foreman, Ann and Goldberg, Joseph and Borriello, Gaetano, «Empirical Measurements of Intrabody Communication Performance Under Varied Physical Configurations,» de *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology*, Orlando, Florida, ACM, 2001, pp. 183-190.

[54] Xu, Ruoyu and Zhu, Hongjie and Yuan, Jie, «Electric-field intrabody communication channel modeling with finite-element method,» *IEEE transactions on biomedical engineering*, vol. 58, nº 3, pp. 705-712, 2011.

[55] Seyedi, MirHojjat and Lai, Daniel Tze Huei, «Effect of limb joints and limb movement on intrabody communications for body area network applications,» *J Med Biol Eng*, vol. 34, nº 3, pp. 276-83, 2014.

[56] Ghamari, Mohammad and Arora, Harneet and Sherratt, R Simon and Harwin, William, «Comparison of low-power wireless communication technologies for wearable health-monitoring applications,» de *Computer, Communications, and Control Technology (I4CT), 2015 International Conference on*, IEEE, 2015, pp. 1-6.

[57] Xu, Ruoyu and Zhu, Hongjie and Yuan, Jie, «Circuit-coupled FEM analysis of the electric-field type intra-body communication channel,» de *Biomedical Circuits and Systems Conference, 2009. BioCAS 2009. IEEE*, IEEE, 2009, pp. 221-224.

[58] Hanson, Mark A and Powell Jr, Harry C and Barth, Adam T and Ringgenberg, Kyle and Calhoun, Benton H and Aylor, James H and Lach, John, «Body area

sensor networks: Challenges and opportunities,» *Computer IEEE*, vol. 42, nº 1, 2009.

[59] Hernandez, Marco and Miura, Ryu, Body Area Networks using IEEE 802.15. 6: Implementing the ultra wide band physical layer, Academic Press, 2014.

[60] Plonsey, Robert and Barr, Roger C, Bioelectricity: a quantitative approach, Springer Science & Business Media, 2007.

[61] Razak, AHA and Zayegh, Aladin and Begg, RK and Seyedi, M and Lai, DT, «BFSK modulation to compare intra-body communication methods for foot plantar pressure measurement,» de *GCC Conference and Exhibition (GCC), 2013 7th IEEE*, IEEE, 2013, pp. 172-176.

[62] Kobayashi, Takumi and Shimatani, Yuichi and Kyoso, Masaki, «Application of near-field intra-body communication and spread spectrum technique to vital-sign monitor,» de *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, IEEE, 2013, pp. 4517-4520.

[63] Seyedi, MirHojjat and Cai, Zibo and Lai, Daniel TH and Rivet, Francois, «An energy-efficient pulse position modulation transmitter for galvanic intrabody communications,» de *Wireless Mobile Communication and Healthcare (Mobihealth), 2014 EAI 4th International Conference on*, IEEE, 2014, pp. 192-195.

[64] Cavallari, Riccardo and Martelli, Flavia and Rosini, Ramona and Buratti, Chiara and Verdone, Roberto, «A survey on wireless body area networks: Technologies and design challenges,» *IEEE Communications Surveys \& Tutorials*, vol. 16, nº 3, pp. 1635-1657, 2014.

[65] Akyildiz, Ian F and Su, Weilian and Sankarasubramaniam, Yogesh and Cayirci, Erdal, «A survey on sensor networks,» *IEEE Communications magazine*, vol. 40, nº 8, pp. 102-114, 2002.

[66] Barakah, Deena M and Ammad-uddin, Muhammad, «A survey of challenges and applications of wireless body area network (wban) and role of a virtual doctor server in existing architecture,» de *Intelligent Systems, Modelling and Simulation (ISMS), 2012 Third International Conference on*, IEEE, 2012, pp. 214-219.

[67] Ullah, Sana and Mohaisen, Manar and Alnuem, Mohammed A, «A review of IEEE 802.15. 6 MAC, PHY, and security specifications,» *International Journal of Distributed Sensor Networks*, vol. 9, nº 4, p. 950704, 2013.

[68] Shinagawa, Mitsuru and Fukumoto, Masaaki and Ochiai, Katsuyuki and Kyuragi, Hakaru, «A near-field-sensing transceiver for intrabody communication based on the electrooptic effect,» *IEEE Transactions on instrumentation and measurement*, vol. 53, nº 6, pp. 1533-1538, 2004.

[69] Bae, Joonsung and Song, Kiseok and Lee, Hyngwoo and Cho, Hyunwoo and Yoo, Hoi-Jun, «A low energy crystal-less double-FSK transceiver for wireless body-area-network,» de *Solid State Circuits Conference (A-SSCC), 2011 IEEE Asian*, IEEE, 2011, pp. 181-184.

[70] Wang, Hao and Tang, Xian and Choy, Chiu Sing and Leung, Ka Nang and Pun, Kong Pang, «A 5.4-mW 180-cm transmission distance 2.5-Mb/s advanced

techniques-based novel intrabody communication receiver analog front end,» *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 23, nº 12, pp. 2829-2841, 2015.

[71] Cho, Hyunwoo and Lee, Hyungwoo and Bae, Joonsung and Yoo, Hoi-Jun, «A 5.2 mW IEEE 802.15. 6 HBC standard compatible transceiver with power efficient delay-locked-loop based BPSK demodulator,» *IEEE journal of solid-state circuits*, vol. 50, nº 11, pp. 2540-2559, 2015.

[72] Wang, Hao and Choy, CS, «A 170cm transmission distance, high speed IntraBody Communication receiver design and its application to FPGA audio player,» de *2013 IEEE International Conference of IEEE Region 10 (TENCON 2013)*, 2013.

[73] Chung, Ching-Che and Chang, Chi-Tung and Lin, Chih-Yu, «A 1 Mb/s--40 Mb/s human body channel communication transceiver,» de *VLSI Design, Automation and Test (VLSI-DAT), 2015 International Symposium on*, IEEE, 2015, pp. 1-4.

[74] Lin, Yu-Tso and Lin, Yo-Sheng and Chen, Chun-Hao and Chen, Hsiao-Chin and Yang, Yu-Che and Lu, Shey-Shi, «A 0.5-V biomedical system-on-a-chip for intrabody communication system,» *IEEE Transactions on Industrial electronics*, vol. 58, nº 2, pp. 690-699, 2011.

[75] Song, Seong-Jun and Cho, Namjun and Yoo, Hoi-Jun, «A 0.2-mW 2-Mb/s digital transceiver based on wideband signaling for human body communications,» *IEEE Journal of Solid-State Circuits*, vol. 42, nº 9, pp. 2021-2033, 2007.

[76] Bae, J., Song, K., Lee, H., Cho, H., & Yoo, H. J., «A 0.24-nJ/b wireless body-area-network transceiver with scalable double-FSK modulation.,» *IEEE Journal of Solid-State Circuits*, vol. 47, nº 1, pp. 310-322, 2012.

[77] Shih, Horng-Yuan and Chang, Yu-Chuan, «68.4 µW 400 MHz intrabody communication receiver front-end for biomedical applications,» *Electronics letters*, vol. 48, nº 3, pp. 143-144, 2012.

Modelling and Performance Analysis of MAC Protocols for Heterogeneous WSN

Canek Portillo-Jimenez, Jorge Martinez-Bauset and Vicent Pla

ITACA Research Insitute
Universitat Politècnica de València, Spain

Abstract. The modelling and performance evaluation of Wireless Sensor Networks (WSN) are of capital importance for their correct design and successful deployment. WSN have experienced an important resurgence, especially through applications that integrate the Internet of Things. In that sense, a WSN can be constituted of different classes of nodes, having different characteristics or requirements (heterogeneity). In this work, a performance study of WSN is performed considering heterogeneous scenarios and diverse medium access priorities. To accomplish that, an analytical model is developed with a pair of two-dimensional Discrete-Time Markov Chains. Scenarios with two classes of nodes forming the network were studied. Performance parameters such as packet average delay, throughput and consumed energy, are obtained and validated by simulation, showing accurate results.

1 Introduction

A WSN can be defined as a collection of nodes that performs sensing, processing and communication activities in a cooperative fashion with limited energy resources. Some examples of WSN applications are medical, industrial, agriculture and environmental monitoring. Furthermore the technological evolution of sensors promises to facilitate the integration of WSN with Internet of Things (IoT), enhancing applications for WSN, such as smart grid, smart water, smart transport systems and smart homes [1].

Sensors nodes are considered energy-constrained devices as they are battery-supplied. One of the major contributors of energy consumption is the energy spent in the channel access process, which is performed by the nodes in order to transmit data [2]. The access of the nodes to the medium is coordinated by the Medium Access Control (MAC) protocol. Recent WSN MAC protocol developments employ Duty-Cycling (DC) operation to save energy and maximize the lifetime of battery-powered sensor nodes. In WSN MAC with DC, sensors are put to sleep periodically to save energy, waking up during the packet exchange periods. The S-MAC was the first MAC protocol for WSN to implement the DC technique, and is also one of the most popular [3].

In addition to the energy issue, it is important to note that a WSN may be constituted of different classes of nodes with different traffic patterns and

C.Fernandez-Llatas and M.Guillen (Ed.):
Workshop on Innovation on
Information and Communication Technologies
(ITACA-WIICT 2017)

ISBN 978-84-697-7327-7

even different priority requirements (heterogeneous scenario). For instance, heterogeneous WSN deployments, where emergency situations may arise, such as fires, earthquakes or some medical applications, need to send data to the destination node as soon as the event occurs. Thus, they need to have priority in the transmission of the information, in relation to other nodes constituting the WSN.

On the other hand, the modelling and performance study of WSN are of capital importance for their design and successful deployment. There are currently some analytical models to evaluate the performance of MAC in WSN, and specifically the S-MAC protocol. Recent examples of modelling and performance analysis with applications to the S-MAC can be found in [4–7]. In particular, in those papers the authors have modelled the protocols using Discrete-Time Markov Chains (DTMC). In [4, 5], they model the S-MAC with discrete time Markov chains (DTMC), but assuming that the states of the nodes are mutually independent. In fact, there is a certain dependency among nodes which is considered in [6]. In addition, all these models were designed considering homogeneous scenarios, where all nodes behave in the same way in terms of loads, communication and energy capabilities. They do not consider the possibility of different loads or different access priorities to the communication channel.

As we have seen, there are analytical models for WSN MAC protocol that incorporate Markov chains, but as far as we know, there is no one that incorporates in the analysis the existence of heterogeneous scenarios with different classes of nodes that have different characteristics or requirements and with the possibility of prioritization. The main contribution of this work is the analytical modelling and performance evaluation of the S-MAC WSN, incorporating different classes of nodes for heterogeneous scenarios. Furthermore, the operation of nodes with different loads and the possibility of managing different access priorities are considered. The model is based on two two-dimensional Discrete-Time Markov Chains (2D-DTMC). This model extends the capabilities of the models of [6] to enable the modelling of different classes of nodes with different priority assignments in heterogeneous network scenarios.

The remainder of the article is distributed as follows. In section 2 the transmission model is presented, where the general scenario is explained. The mathematical modelling of the system is presented in section 3. The analysis to obtain the performance parameters is developed in section 4. The results and their discussion are set out in section 5. Finally the conclusions are presented in section 6.

2 Transmission model

S-MAC protocol

In S-MAC the time is divided into cycles of equal duration T, and each cycle consists of an *active* period and a *sleep* period. The *active* period is subdivided into two parts: the *sync* period of fixed-duration T_{sync}, where $SYNC$ packets

are exchanged, and the *data* period, where the $DATA$ packets are exchanged (see Figure 1).

In a *sync* period, nodes choose a *sleep-awake* schedule and exchange it with its neighbours through $SYNC$ packets. These packets include the address of the node that sends the packet and the start time of its next *active* period. With this information, nodes are able to coordinate to wake up together at the beginning of each *sync* period. To send $SYNC$ packets, they use Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) mechanism for a contention-based access to the channel. CSMA/CA is based on the generation of a random backoff time and a carrier sensing procedure. If the channel is unoccupied when the backoff timer expires, then the node transmits the $SYNC$ packet.

Nodes also use CSMA/CA to transmit $DATA$ packets during the *data* period. They generate new backoff times at each *data* period initiation and perform carrier sense. If the channel is idle when the backoff timer expires, then de node can transmit the $DATA$ packet using the $RTS/CTS/DATA/ACK$ handshake. When a winning node receives a CTS in response to its previous RTS, it transmits one $DATA$ packet. In S-MAC, a node goes to *sleep* until the beginning of the next *sync* period when: i) it loses the contention (hears a busy medium before its backoff expires); ii) it encounters an RTS collision; iii) after a successful transmission (only one packet per cycle is sent).

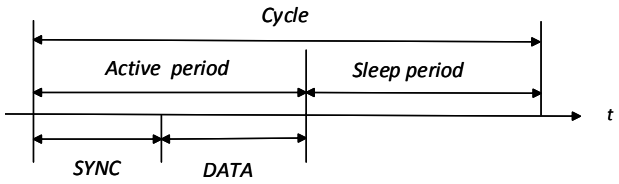


Fig. 1. Scheme of an S-MAC cycle, its subdivisions in *active* and *sleep* periods.

Scenario and assumptions

A heterogeneous WSN network with different classes of nodes forming the network is considered. Figure 2 shows a diagram of the heterogeneous network model, where all nodes are reached in one hop and send the packets to a *sink* node. The scenario shows the formation of a single cell cluster, but multiple clusters together can form a larger network. Two classes of nodes are considered, and it is assumed that class 1 nodes have priority in medium access over class 2 nodes. For convenience, one node of each class is selected as reference nodes RN_1 and RN_2. It is assumed that the *sink* node behaves like a packet absorption node, it only receives packets (never transmits $DATA$ packets).

In this study we also assume that: i) the 2D-DTMC model is based on infinite retransmissions, ii) all nodes contain the same initial energy; iii) the channel is

error-free; iv) packets arrive to the buffer of a node following a renewal process, and the number of packets that arrive per cycle is characterized by independent and identically distributed random variables. A node has a queue or buffer that can store at most Q packets, and it serves them according a FIFO discipline. For simplicity, we assume that the number of packets that arrive to a buffer follow a discrete Poisson distribution of mean λT, where λ is the packet arrival rate and T is the cycle duration.

On the other hand, for analysis and explanation purposes, we might refer a generic notation throughout the paper to identify any of both classes of parameters, unless otherwise specified. In that sense, the expressions that are developed in sections 3 and 4, are equally applicable to both classes of nodes

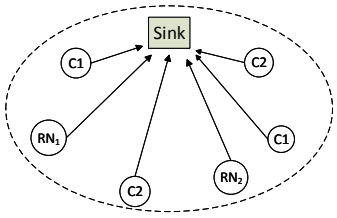


Fig. 2. Heterogeneous network model. There are two classes of devices conforming the WSN network. Class 1 (C1) nodes have priority in accessing the medium.

Assignment of medium access priorities

One way to prioritize class 1 devices over class 2 devices is to ensure that class 1 devices first perform the media access procedure. Figure 3 shows a diagram of the transmission process corresponding to the *date* period. In the figure, the synchronization period for both classes of nodes has been omitted.

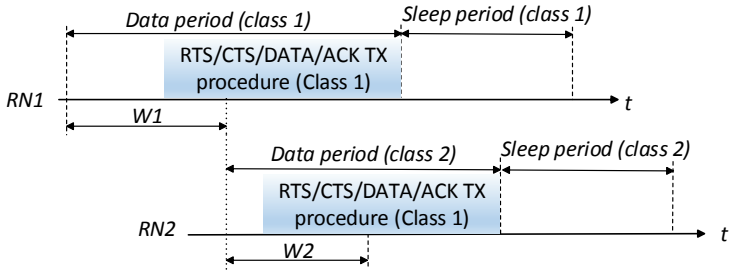


Fig. 3. Operation of MAC protocol during *data* period, for both classes of nodes.

In this model, the media access priority is granted to class 1 nodes considering the following process. At the beginning of a cycle, the class 1 nodes activate the media access mechanism, contending only between the same class 1 nodes, for accessing to the channel. The class 2 nodes wake up just after the class 1 contention window (W_1) has ended, and if they detect the medium unoccupied, they will try to transmit, activating their own contention procedure. If the class 2 nodes detect the medium occupied at that instant, they return to the *sleep* mode and they will wake up again in the next cycle.

3 Modelling of the system

Access to the medium

The *RN* is an arbitrarily chosen node. A node is considered active when it has at least one packet in the queue (non-empty queue). Active nodes generate a random back-off time selected from $\{0, ..., W-1\}$. When the RN is active, it transmits a packet successfully if the other contending nodes select back-off times larger than the one chosen by the RN. The packet transmitted by the RN will fail (collide) when the RN and one or more of the other contending nodes choose the same backoff time, and this backoff time is the smallest among all contending nodes. If the backoff time generated by the RN is not the smallest among those generated by the other contending nodes, two outcomes are possible: either another node is able to transmit successfully, or other nodes collide when transmitting. Nodes that loose the contention (because they hear a busy medium before their backoff time expires) or encounter an RTS collision, go to *sleep* until the *sync* period of the next cycle.

Consider a cycle where the RN is active and denote by $k, 0 \leq k \leq N-1$, the number of nodes that are also active in the same cycle in addition to the RN. Let $P_{s,k}, P_{sf,k}$ and $P_{f,k}$ be the probability that the RN transmits a packet successfully, transmits a packet (successfully or with collision), and it transmits with failure (collision), when it contends with other k nodes. Then,

$$P_{s,k} = \sum_{i=0}^{W-1} \frac{1}{W} \left(\frac{W-1-i}{W} \right)^k, \quad P_{sf,k} = \sum_{i=0}^{W-1} \frac{1}{W} \left(\frac{W-i}{W} \right)^k,$$

$$P_{f,k} = P_{sf,k} - P_{s,k} = \frac{1}{W} \,. \tag{1}$$

$P_{s,k}$ is the probability that the RN chooses a backoff value from $\{0, ..., W-1\}$ and the other k nodes choose a larger value. $P_{sf,k}$ and $P_{f,k}$ can be described in similar terms. Conditioned on a successful or unsuccessful packet transmission by the RN when contending with other k nodes, the average backoff times are given by,

$$BT_{s,k} = \frac{1}{P_{s,k}} \sum_{i=0}^{W-1} i \cdot \frac{1}{W} \left(\frac{W-1-i}{W} \right)^k,$$

$$BT_{f,k} = \sum_{i=0}^{W-1} i \cdot \left[\left(\frac{W-i}{W} \right)^k - \left(\frac{W-1-i}{W} \right)^k \right]. \quad (2)$$

System of two classes with priorities

Here we model the evolution of the number of packets in the queue of RN_1 and RN_2, and the number of active nodes in the cluster over the time, by a pair of 2D-DTMC, one chain for each class of nodes. The state in the 2D-DTMC is represented by (i, m), where $i \leq Q$ is the number of packets in queue of RN, and m is the number of active nodes in the network of the corresponding class other than the RN, $m \leq M = N - 1$. Then, $P_{(i,m),(j,n)}$ is the transition probability from state (i, m) to state (j, n).

The first 2D-DTMC is taken from [6] and designated to describe the evolution of the class 1 nodes in the network. A second 2D-DTMC is developed for the description of the evolution of the nodes of class 2. Before defining the transition probabilities of both 2D DTMCs, some useful expressions are defined.

Let $A_k = (\lambda T)^k e^{-\lambda T}/k!$ be the probability that k packets arrive, and $A_{\geq k} = 1 - \sum_{n=0}^{k-1} A_n$ be the probability that k or more packets arrive to a node in a cycle. Let $B_j(n) = \binom{n}{j} A_0^{n-j}, \hat{A} = 1 - A_0$, be the probability that j nodes out of n that have their queues empty receive packets in a cycle. Let $S_k = kP_{s,k-1}$ and $\hat{S}_k = 1 - S_k$ be the probabilities that an active node successfully transmits a packet in a cycle when k nodes compete, and its complementary. Then, $\hat{S}_k$ is the probability that packets from two or more nodes collide. Therefore, $S_1 = 1$, $\hat{S}_1 = 0$. Let E be the probability that the queue of a node becomes empty when it transmits successfully. Then, $E = P_s A_0 \pi_1 / P_s(1 - \pi_0)$, where π_k is the stationary probability of finding k packets in the queue of the RN and P_s is the (average) probability that the RN transmits a packet successfully in a random cycle, conditioned on the RN being active. Let $\hat{E}$ be the probability that the queue of a node remains non-empty when it transmits successfully. Then, $\hat{E} = 1 - E$. Let $\boldsymbol{P} = [P_{(i,m),(j,n)}]$ be the transition probability matrix of the 2D-DTMC, where $P_{(im),(jn)}$ is the probability that j packets are found in the queue of the RN and other n nodes are active in the cycle $m+1$, conditioned on finding i packets in the queue of the RN and m other nodes active at cycle m.

Solution of the 2D-DTMC

The state transition probabilities of the 2D-DTMC for class 1 nodes and for class 2 nodes, are given in Table 2 and Table 3, respectively (see the appendices A and B). In both tables, the first row defines transitions caused by new arrivals when the corresponding RN has an empty queue and no other node is active. The second row describes the transmissions made by other nodes while the RN has its queue empty. The third and fourth rows define RN's transmissions in cycles where there are no other active nodes and when there are other active nodes, respectively. Finally, the fifth row defines impossible transitions.

The solution of each of these 2D-DTMC can be obtained by solving the set of linear equations:

$$\boldsymbol{\pi P} = \boldsymbol{\pi}, \quad \boldsymbol{\pi e} = 1 \, . \tag{3}$$

Where $\boldsymbol{\pi} = [\pi(i,n)]$ is the stationary distribution, $\boldsymbol{P}$ is the transition probability matrix, whose elements are defined in Table 1 for class 1 nodes and in Table 2 for class 2 nodes, and $\boldsymbol{e}$ is a column vector of ones.

The average probability, P_s, that the corresponding the RN transmits a packet successfully, conditioned on the the RN being active, is given by,

$$P_s = \frac{1}{G} \sum_{i=1}^{Q} \sum_{k=0}^{K} \pi(i,k) \cdot P_{s,k}, \;\; G = \sum_{i=1}^{Q} \sum_{k=0}^{K} \pi(i,k), \;\; \pi_i = \sum_{0}^{K} \pi(i,k) \, . \tag{4}$$

By solving the set of equations (3), $\boldsymbol{\pi}(P_s)$ can be determined for a given P_s. Then, a new P_s can be obtained from (4) for a given $\boldsymbol{\pi}$. Denote by $\boldsymbol{\pi}$ the solution of this fixed-point equation, i.e., the stationary distribution at the fixed-point.

4 Performance Parameters

In this section we present the expressions used to determine the throughput, average packet delay and average energy consumed per cycle. The procedure to obtain the performance parameters is applied equally to both classes of nodes.

Throughput

The node throughput η is defined as the average number of packets successfully delivered by a node in a cycle. It is determined by,

$$\eta = \sum_{i=1}^{Q} \sum_{k=0}^{K} \pi(i,k) \cdot P_{s,k} \, . \tag{5}$$

In a network composed of N nodes, the aggregate throughput expressed in packets per cycle is obtained with,

$$Th = N \cdot \eta \, . \tag{6}$$

Average packet delay

Let (D) be the average delay (in cycles) that a packet experiences from its arrival until it is successfully transmitted. Then, D can be determined by applying Little's law,

$$D = \frac{N_{av}}{\gamma_a}, \quad N_{av} = \sum_{i=0}^{Q} i\pi_i, \quad \gamma_a = \eta \, . \tag{7}$$

Note that: i) π_i is the stationary probability of finding i packets in the queue of the correspondent RN, and is determined by the expression (4); ii) N_{av} is the average number of packets in queue; iii) γ_a is the average number of packets that entered the queue (*accepted*) per cycle, that it is equal to η.

Average energy consumption

As described in Section 2, the *active* period of a cycle is subdivided into *sync* and *date* periods. The energy consumed during the *active* period represents the most significant contribution to the total energy consumption. In this section we calculate the energy consumed by the reference node RN in the *data* period. It should be noted that only the energy consumed by the radio frequency transceiver is studied. The energy consumed by the sensor nodes due to events related to specific sensing or monitoring tasks depends on the application and is not included here.

Let $E_{d,k+1}$ be the average energy consumed by the corresponding RN when it contends with other $k \geq 1$ nodes during the *data* period of a cycle. It is given by,

$$\begin{aligned} E_{d,k+1} &= q_k[P_{s,k}E^{tx}_{s,k} + P_{f,k}E^{tx}_{f,k} + E_{oh,k}], \\ E^{tx}_{s,k} &= (t_{RTS} + t_{DATA})P_{tx} \\ &\quad + (BT_{s,k} + t_{CTS} + t_{ACK} + 4D_p)P_{rx}, \\ E^{tx}_{f,k} &= t_{RTS}P_{tx} + (BT_{f,k} + 2D_p)P_{rx}, \\ E_{oh,k} &= (kP_{s,k}BT_{s,k} + (1 - kP_{s,k} - P_{sf,k})BT_{f,k} + D_p)P_{rx} \ . \end{aligned} \tag{8}$$

Where $E^{tx}_{s,k}, E^{tx}_{f,k}, E_{oh,k}$ are energy consumption values of the RN, when it contends with other k nodes and it transmits successfully, it transmits with failure (collision), and it overhears other transmissions, respectively.

Let $q_{1,k} = (k+1)/N$ be the probability that the corresponding RN is active, conditioned on finding $k+1$ nodes active. When the RN is active, $1 - P_{sf,k}$ defines the probability that it does not transmit, but the other k do. In that case: i) one of them transmits successfully (with probability $kP_{s,k}$); or ii) two or more collide (with probability $1 - kP_{s,k}$). It should be noted that if the RN is not active, then it will not listen to the channel, since we assume that nodes transmit to the *sink*, and the *sink* only receives, never transmits.

In addition, $t_{RTS}, t_{DATA}, t_{CTS}$ and t_{ACK}, are the corresponding packet transmission times, P_{tx}, P_{rx} are the transmission and reception power levels respectively, and D_p is the one-way propagation delay. It should be noted that $P_{s,0} = 1$, $P_{f,0} = 0$, $BT_{s,0} = (W-1)/2$, $E_{d,1} = q_{1,0}E^{tx}_{s,0}$, y $E_{d,0} = 0$.

The average energy consumed by the RN during the *data* period in a cycle is given by,

$$E_d = \sum_{k=0}^{N} E_{d,k} \cdot R_k \ . \tag{9}$$

Where R_k is the stationary probability of finding k active nodes in a cycle, and is determined as:

$$R_k = \sum_{i=1}^{Q} \pi(i, k-1) + \pi(0, k),\ 1 \leq k \leq N-1, \quad R_N = \sum_{i=1}^{Q} \pi(i, N-1),$$

$$R_0 = \pi(0, 0)\ . \tag{10}$$

In addition to E_d, a node also consumes energy due to the exchange of signalling messages like $SYNK$. The *sleep* part of a cycle is not included, as we consider that the energy consumed is negligible compared to the energy consumed in the data period.

For class 2 nodes, the same expressions are applied to obtain the consumed energy in the *data* period E_{d_2}, to which must be added E_o, that is the energy consumed to sense the channel in the wake up instant after W_1. The final equations for determining the consumed energy for both classes of nodes, are given by,

$$E_1 = E_{d_1},$$

$$E_2 = (1 - R_{1,0})E_o + E_{d_2}\ . \tag{11}$$

Where $R_{1,0}$ is the stationary probability of finding none active nodes of class 1, in a cycle. E_1 and E_2 are the consumed energy for nodes of class 1 and class 2, respectively.

5 Numerical results

Scenarios and parameter configuration

The analytical results are obtained from the developed 2D-DTMC models. The simulation results are obtained by means of a custom-based discrete event simulator developed in C language, where the transmission scheme is implemented. The developed simulator mimics the physical behaviour of the system. That is, at each cycle a node receives packets according to a given discrete distribution, it contends for access to the channel with other nodes if it has packets in the buffer, and if it wins, it then transmits a packet according to the transmission scheme. The simulation results are completely independent of those obtained by the analytical model. That is, the calculation of the performance metrics in the simulations does not depend on the developed mathematical expressions.

Based on a WSN as the one illustrated in Figure 2, the network is configured considering two classes of nodes and two scenarios, with the following parameters: DATA packet size $S = 50$ bytes, transmission power level $P_{tx} = 52$ mW, reception power level $P_{rx} = 59$ mW, queue capacity of a node $Q = 5$ packets, packet arrival rate for class 1 nodes, $\lambda_1 = \{0.5, 1.0\}$ packets/s, packet arrival rate for class 2, $\lambda_2 = \{0.5, ..., 4.5\}$ packets/s. In scenario 1 (SC1), the number of sensor nodes of class 1 and 2 are $N_1 = 3$ and $N_2 = 6$. In scenario 2 (SC2), the

number of sensor nodes of class 1 and class 2 are $N_1 = 3$ and $N_2 = 9$. The time parameters are summarized in Table 1. The duration of a backoff time slot is set to 0.1 ms.

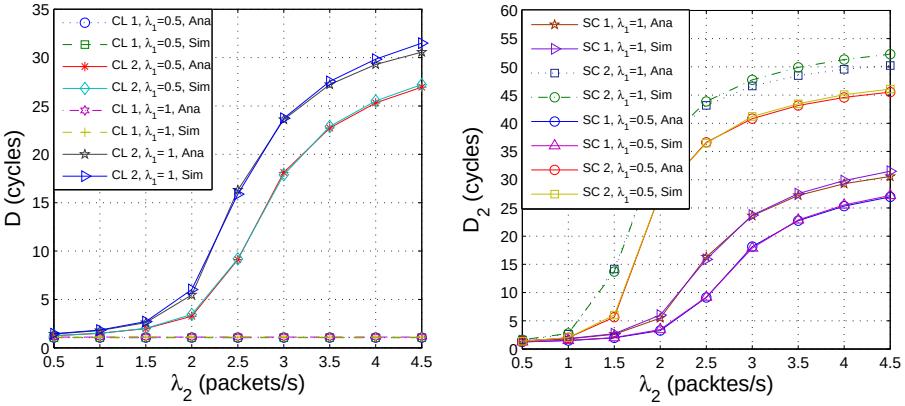
Under these scenarios, data derived from simulation and the solution of the analytical model are obtained for the performance parameters corresponding to the average packet delay, throughput and average energy consumption. In the following subsections the results of the performance parameters obtained from the analytical model and by simulation are shown.

Table 1. Temporary parameters (milliseconds)

Duration of cycle (T)	60	Propagation delay (D_p)	0.0001
t_{RTS}, t_{CTS} and t_{ACK}	0.18	t_{SYNC}	0.18
t_{DATA}	1.716	Contention window (W)	128

Average packet delay

Figure 4 (a) shows the average packet delay expressed in cycles, for both classes of nodes and two values of packet arrival rate λ_1. The average delay experienced by the packets of class 1 nodes, is a constant value. This is because λ_1 and the number of nodes N_1 are also constant values, and class 1 nodes have priority over class 2 nodes. However, for class 2 nodes the arrival rate of packets λ_2 is varied.



(a) Both classes, scenario 1.

(b) Class 2, both scenarios.

Fig. 4. Average packet delay in cycles.

From Figure 4 (a), note that the fraction of packets transmitted with failure (collisions) increases as λ_2 increases, which increases the average packet delay. Note also that by doubling the value of λ_1, the average packet delay for class 2 nodes, increases. This is explained by the load increase on the priority nodes, which now use the channel more often than before, therefore for the class 2 nodes, it takes longer to transmit their packets. In the delay of class 1 nodes, no effect is observed as a consequence of the λ_1 increment.

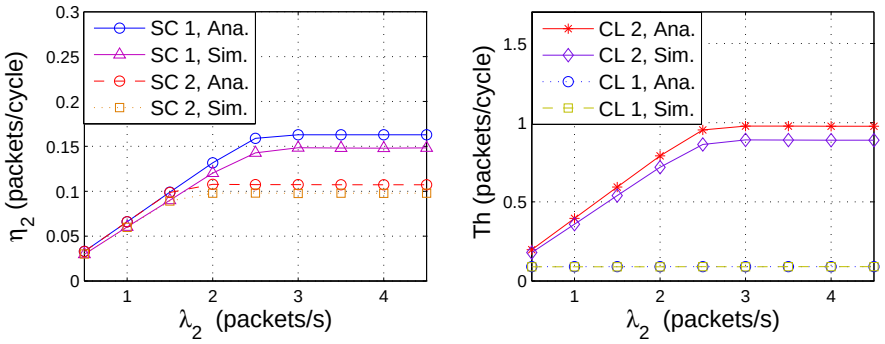
On the other hand, Figure 4 (b) shows the average packet delay for class 2, considering both scenarios and two λ_1 values. Note that the number of nodes of class 2 is double and triple, with respect to the number of nodes of class 1, for scenario 1 and scenario 2 respectively. For scenario 1, delay values are smaller than in scenario 2, especially when λ_2 is increased. The figures show relative stability up to the values of $\lambda_2 = \{1.5, 2\}$ packets/s, where there exist inflection points and a pronounced delay increase begins. The figures of scenario 2 present a similar behaviour, but the delay increase begins before, starting from $\lambda_2 = 1$ packets/s. For scenario 2, where the number of nodes is incremented, D_2 reaches higher values than in scenario 1. When the number of nodes increases, it takes longer to get channel access due to the increased contention. Then, packets have to wait longer time in the queue before being transmitted. In addition, more collisions occur which leads to more retransmissions. An increase in the average packet delay is observed when increasing λ_1, and can be explained in the same terms as in Figure 4 (a).

Throughput

Figure 5 (a) shows the throughput for class 2 nodes, considering both scenarios. There is a linear part which is similar in both scenarios. A maximum throughput value is reached, from which a constant behaviour with λ_2 is observed. A saturation level is reached.

In general, the throughput is higher for scenario 1 than for scenario 2. This is because in scenario 1, the number of nodes is smaller and therefore there is less contention. The nodes have more opportunities to access the channel and succeed in the transmission of packets,reaching higher values of throughput. On the other hand, since there are fewer nodes, it is possible to operate at higher arrival rates without saturation.

In general, the throughput is higher for scenario 1 than for scenario 2. This is because in scenario 1, the number of nodes is smaller and therefore there is less contention. The nodes have more opportunities to access the channel and succeed in the transmission of packets,reaching higher values of throughput. On the other hand, since there are fewer nodes, it is possible to operate at higher arrival rates without saturation.



(a) Throughput per node for class 2, both scenarios.

(b) Aggregate throughput for both classes, scenario 1.

Fig. 5. Throughput per node and aggregate throughput.

Figure 5 (b) shows the aggregate throughput for both classes, considering scenario 1. Given that the number of class 1 nodes is constant, as well as λ_1, their aggregate throughput also remain constant. For class 2, the aggregate throughput increases with λ_2, reaching a saturation level with a inflection point in $\lambda_2 = 2.5$. The aggregate throughput is the sum of the throughput of all nodes of the corresponding class.

Average energy consumption

Figure 6 shows the average energy consumption of class 2 node per cycle, considering both scenarios. Considering only the class 2 nodes, there is a higher energy consumption per node for the scenario 1 than for scenario 2. Recall that scenario 1 is composed with less number of nodes. This scenario of higher energy consumption corresponds to the scenario where a higher throughput is achieved (Figure 5 (a)). Greater values of throughput implies to carry out more transmissions and packet deliveries, leading to a higher activity of the nodes, and therefore a greater energy consumption.

Figure 7 shows the average energy consumption for both classes, considering both scenarios. For nodes of class 1, the energy consumed is constant with λ_2. This is because both, the packet arrival rate λ_1 and the number of nodes N_1 are constant. For class 2 nodes, the packet arrival rate is varied ($\lambda_2 = \{, ..., 0 - 4.5\}$ packets/s), and the number of nodes that constitute the network is different for each scenario.

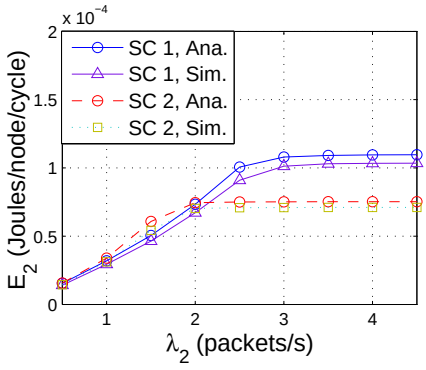
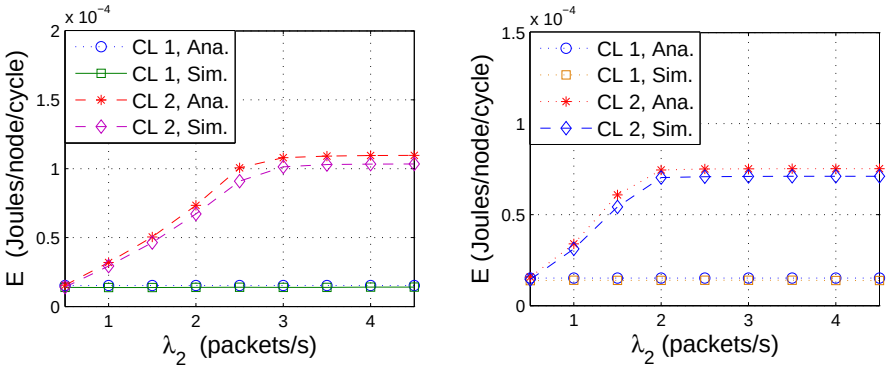


Fig. 6. Average energy consumption per cycle. Class 2. Both scenarios.

Figure 6 and Figure 7 show a constant increment of consumed energy for the first values of λ_2 (with practically the same slope). Then, after a certain value of λ_2, the energy consumption remains constant. This occurs when $\lambda_2 = 3$ for scenario 1 and $\lambda_2 = 2$ for scenario 2, in both scenarios for nodes of class 2.



(a) Both classes, scenario 1.

(b) Both classes, scenario 2.

Fig. 7. Average energy consumption per cycle.

The nodes eventually reach a limit of activity, that has associated a limit of energy consumption. Figure 7 (a) shows that for scenario 1, the limit of energy consumption is higher and is reached at a higher traffic load, than that of scenario 2 (figure 7(b)). As there are less nodes in scenario 1, the nodes can operate at higher packet arrival rates before reaching a saturation level of activity.

6 Conclusions

A performance study of a heterogeneous WSN network has been performed, considering different classes of nodes and the assignment of medium access priorities. A performance model was developed for a WSN MAC protocol that considers heterogeneity an priorities. It operates in WSNs with a synchronous duty-cycled MAC protocol like S-MAC. The model is based in two 2D-DTMC. Unlike existing models for duty-cycled MAC protocols, our model takes into account the different classes of nodes, and the assignment of medium access priorities.

The analytical model is solved for specific scenarios, obtaining values of performance parameters like average packet delay, throughput and average energy consumption. The analytical model is validated through discrete-event based simulations, showing accurate results.

An impact on the performance parameters of class 2 nodes is observed, due to the prioritization of the access to the medium for nodes of class 1. This occurs when the traffic and the number of nodes increase, although in principle there may be an acceptable traffic coexistence between both classes, considering that the priority class contributes with a low load.

Apendix A

Table 2. Transition probabilities of 2D DTMC for class 1 nodes

<table>
<tr><td colspan="2">There are no active nodes. Transitions are due to new arrivals</td></tr>
<tr><td>$P_{00,jn} = B_n(N_1)A_j; \quad 0 \le n \le N_1,\ 0 \le j \le (Q_1 - 1)$</td><td>$P_{00,Q_1 n} = B_n(N_1)A_{\ge Q_1}; \quad 0 \le n \le N_1$</td></tr>
<tr><td colspan="2">There are no packets in queue of RN_1. There are no transmissions due to RN_1. Transitions are caused by the other k active nodes</td></tr>
<tr><td>$P_{0m,jn} = S_m E B_{n-m+1}(N_1 - m)A_j$
$+ S_m \hat{E} B_{n-m}(N_1 - m)A_j$
$+ \hat{S}_m B_{n-m}(N_1 - m)A_j$
$0 \le j \le (Q_1 - 1),\ 1 \le m \le n < N_1$
$P_{0m,Q_1 n} = S_m E B_{n-m+1}(N_1 - m)A_{\ge Q_1}$
$+ S_m \hat{E} B_{n-m}(N_1 - m)A_{\ge Q_1}$
$+ \hat{S}_m B_{n-m}(N_1 - m)A_{\ge Q_1}$
$1 \le m \le n < N_1$
$P_{0m,Q_1 m-1} = S_m E B_0(N_1 - m)A_{\ge Q_1}$
$1 \le m \le N_1$</td><td>$P_{0m,Q_1 N_1} = S_m \hat{E} B_{N_1-m}(N_2 - m)A_{\ge Q_1}$
$+ \hat{S}_m B_{N_1-m}(N_1 - m)A_{\ge Q_1}$
$1 \le m \le N_1$
$P_{0m,jN_1} = S_m \hat{E} B_{N1-m}(N_1 - m)A_j$
$+ \hat{S}_m B_{N_1-m}(N_1 - m)A_j$
$0 \le j \le (Q_1 - 1),\ 1 \le m \le N_1$
$P_{0m,jm-1} = S_m E B_0(N_1 - m)A_j$
$0 \le j \le (Q_1 - 1),\ 1 \le m \le N_1$</td></tr>
<tr><td colspan="2">RN_1 is the only active node</td></tr>
<tr><td>$P_{i0,jn} = P_{s_1,0} B_n(N_1)A_{j-i+1}$
$+ (1 - P_{s_1,0})B_n(N_1)A_{j-i}$
$1 \le i \le j \le Q_1,\ 0 \le n \le N_1$</td><td>$P_{i0,Q_1 n} = P_{s_1,0} B_n(N_1)A_{\ge Q1-i+1}$
$+ (1 - P_{s_1,0})B_n(N_1)A_{\ge Q_1-i}$
$1 \le i \le Q_1,\ 0 \le n \le N_1$</td></tr>
<tr><td colspan="2">Transitions are caused by the $k+1$ active nodes including RN_1</td></tr>
<tr><td>$P_{im,jn} = P_{s_1,m} B_{n-m}(N_1 - m)A_{j-i+1}$
$+ mP_{s_1,m} E B_{n-m+1}(N_1 - m)A_{j-i}$
$+ mP_{s_1,m} \hat{E} B_{n-m}(N_1 - m)A_{j-i}$
$+ (1 - (m+1)P_{s_1,m})B_{n-m}(N_1 - m)A_{j-i}$
$1 \le i \le j \le (Q_1 - 1),\ 1 \le m \le n < N_1$
$P_{im,Q_1 n} = P_{s_1,m} B_{n-m}(N_1 - m)A_{\ge Q-i+1}$
$+ mP_{s_1,m} E B_{n-m+1}(N_1 - m)A_{\ge Q-i}$
$+ mP_{s_1,m} \hat{E} B_{n-m}(N_1 - m)A_{\ge Q-i}$
$+ (1 - (m+1)P_{s_1,m})B_{n-m}(N_1 - m)A_{\ge Q-i}$
$1 \le i \le Q,\ 1 \le m \le n < N_1$
$P_{im,Q_1 m-1} = mP_{s_1,m} E B_0(N_1 - m)A_{\ge Q_1-i}$
$1 \le i \le Q_1,\ 1 \le m \le N_1$</td><td>$P_{im,Q_1 N_1} = P_{s_1,m} B_{N_1-m}(N_1 - m)A_{\ge Q-i+1} +$
$mP_{s_1,m} \hat{E} B_{N_1-m}(N_1 - m)A_{\ge Q-i}$
$+ (1 - (m+1)P_{s_1,m})B_{N_1-m}(N_1 - m)A_{j-i}$
$1 \le i \le j \le Q_1,\ 1 \le m \le N_1$
$P_{im,jN_2} = P_{s_1,m} B_{N_1-m}(N_1 - m)A_{j-i+1}$
$+ mP_{s_1,m} \hat{E} B_{N_1-m}(N_1 - m)A_{j-i}$
$+ (1 - (m+1)P_{s_1,m})B_{N_1-m}(N_1 - m)A_{j-i}$
$1 \le i \le j \le (Q_1 - 1),\ 1 \le m \le n \le N_1$
$P_{im,jm-1} = mP_{s_1,m} E B_0(N_1 - m)A_{j-i}$
$1 \le i \le j \le (Q_1 - 1),\ 1 \le m \le N_1$
$P_{im,i-1,n} = P_{s_1,m} B_{n-m}(N_1 - m)A_0$
$1 \le i \le Q_1,\ 0 \le m \le n \le N_1$</td></tr>
<tr><td colspan="2">Impossible transitions</td></tr>
<tr><td>$P_{im,jm-1} = 0; \quad 1 \le i \le Q_2,\ j < i, \quad 1 \le m \le N_2$
$P_{im,i-1n} = 0; \quad 1 \le i \le Q_2,\ 1 \le m \le N_2, \quad n < m,$</td><td>$P_{im,j} = 0; \quad 2 \le i \le Q_2,\ j < i-1, \quad 0 \le m \le n \le N_2$
$P_{im,jn} = 0; \quad 0 \le i \le j \le Q_2,\ 2 \le m \le M_2, \quad n < m-1$</td></tr>
</table>

Apendix B

Table 3. Transition probabilities of 2D DTMC for class 2 nodes

There are no active nodes. Transitions are due to new arrivals

$$P_{00,jn} = B_n(N_2)A_j; \quad 0 \le n \le N_2,\ 0 \le j \le (Q_2-1) \qquad P_{00,Q_2n} = B_n(N_2)A_{\ge Q_2}; \quad 0 \le n \le N_2$$

There are no packets in queue of RN_2. There are no transmissions due to RN_2. Transitions are caused by the other k active nodes

$$\begin{aligned}P_{0m,jn} &= [S_m E B_{n-m+1}(N_2-m)A_j \\ &+ S_m \hat{E} B_{n-m}(N_2-m)A_j + \\ &\hat{S}_m B_{n-m}(N_2-m)A_j]\ R_{1,0} \\ &+ B_{n-m}(N_2-m)A_j\ (1-R_{1,0}) \\ &0 \le j \le (Q_2-1),\ 1 \le m \le n < N_2\end{aligned}$$

$$\begin{aligned}P_{0m,Q_2n} &= [S_m E B_{n-m+1}(N_2-m)A_{\ge Q_2} \\ &+ S_m \hat{E} B_{n-m}(N_2-m)A_{\ge Q_2} \\ &+ \hat{S}_m B_{n-m}(N_2-m)A_{\ge Q_2}]\ R_{1,0} \\ &+ B_{n-m}(N_2-m)A_{\ge Q_2}\ (1-R_{1,0}) \\ &1 \le m \le n < N_2\end{aligned}$$

$$P_{0m,Q_2m-1} = S_m E B_0(N_2-m)A_{\ge Q_2}\ R_{1,0} + 0\ (1-R_{1,0}) \qquad 1 \le m \le N_2$$

$$\begin{aligned}P_{0m,Q_2N_2} &= [S_m \hat{E} B_{N_2-m}(N_2-m)A_{\ge Q_2} \\ &+ \hat{S}_m B_{N_2-m}(N_2-m)A_{\ge Q_2}] \\ &R_{1,0} + B_{N_2-m}(N_2-m)A_{\ge Q_2}\ (1-R_{1,0}) \\ &1 \le m \le N_2\end{aligned}$$

$$\begin{aligned}P_{0m,jN_2} &= [S_m \hat{E} B_{N2-m}(N_2-m)A_j \\ &+ \hat{S}_m B_{N_2-m}(N_2-m)A_j]\ R_{1,0} \\ &+ B_{N_2-m}(N_2-m)A_j\ (1-R_{1,0}) \\ &0 \le j \le (Q_2-1),\ 1 \le m \le N_2\end{aligned}$$

$$P_{0m,jm-1} = S_m E B_0(N_2-m)A_j\ R_{1,0} + 0\ (1-R_{1,0}) \qquad 0 \le j \le (Q_2-1),\ 1 \le m \le N_2$$

RN_2 is the only active node

$$\begin{aligned}P_{i0,jn} &= [P_{s_2,0} B_n(N_2)A_{j-i+1} \\ &+ (1-P_{s_2,0})B_n(N_2)A_{j-i}]\ R_{1,0} \\ &+ B_n(N_2)A_{j-i}\ (1-R_{1,0}) \\ &1 \le i \le j \le Q_2,\ 0 \le n \le N_2\end{aligned}$$

$$\begin{aligned}P_{i0,Q_2n} &= [P_{s_2,0} B_n(N_2)A_{\ge Q2-i+1} \\ &+ (1-P_{s_2,0})B_n(N_2)A_{\ge Q_2-i}]\ R_{1,0} \\ &+ B_n(N_2)A_{\ge Q_2-i}\ (1-R_{1,0}) \\ &1 \le i \le Q_2,\ 0 \le n \le N_2\end{aligned}$$

Transitions are caused by the $k+1$ active nodes including RN_2

$$\begin{aligned}P_{im,jn} &= [P_{s_2,m} B_{n-m}(N_2-m)A_{j-i+1} \\ &+ mP_{s_2,m} E B_{n-m+1}(N_2-m)A_{j-i} \\ &+ mP_{s_2,m} \hat{E} B_{n-m}(N_2-m)A_{j-i} \\ &+ (1-(m+1)P_{s_2,m})B_{n-m}(N_2-m)A_{j-i}]\ R_{1,0} \\ &+ B_{n-m}(N_2-m)A_{j-i}\ (1-R_{1,0}) \\ &1 \le i \le j \le (Q_2-1),\ 1 \le m \le n < N_2\end{aligned}$$

$$\begin{aligned}P_{im,Q_2n} &= [P_{s_2,m} B_{n-m}(N_2-m)A_{\ge Q-i+1} \\ &+ mP_{s_2,m} E B_{n-m+1}(N_2-m)A_{\ge Q-i} \\ &+ mP_{s_2,m} \hat{E} B_{n-m}(N_2-m)A_{\ge Q-i} \\ &+ (1-(m+1)P_{s_2,m})B_{n-m}(N_2-m)A_{\ge Q-i}]R_{1,0} \\ &+ B_{n-m}(N_2-m)A_{\ge Q-i}\ (1-R_{1,0}) \\ &1 \le i \le Q,\ 1 \le m \le n < N_2\end{aligned}$$

$$P_{im,Q_2m-1} = mP_{s_2,m} E B_0(N_2-m)A_{\ge Q_2-i}\ R_{1,0} + 0\ (1-R_{1,0}); \quad 1 \le i \le Q_2,\ 1 \le m \le N_2$$

$$\begin{aligned}P_{im,Q_2N_2} &= [P_{s_2,m} B_{N_2-m}(N_2-m)A_{\ge Q-i+1} + \\ &mP_{s_2,m} \hat{E} B_{N_2-m}(N_2-m)A_{\ge Q-i} \\ &+ (1-(m+1)P_{s_2,m})B_{N_2-m}(N_2-m)A_{j-i}] \\ &R_{1,0} + B_{N_2-m}(N_2-m)A_{j-i}\ (1-R_{1,0}) \\ &1 \le i \le j \le Q_2,\ 1 \le m \le N_2\end{aligned}$$

$$\begin{aligned}P_{im,jN_2} &= [P_{s_2,m} B_{N_2-m}(N_2-m)A_{j-i+1} \\ &+ mP_{s_2,m} \hat{E} B_{N_2-m}(N_2-m)A_{j-i} \\ &+ (1-(m+1)P_{s_2,m})B_{N_2-m}(N_2-m)A_{j-i}]\ R_{1,0} \\ &+ B_{N_2-m}(N_2-m)A_{j-i}\ (1-R_{1,0}) \\ &1 \le i \le j \le (Q_2-1),\ 1 \le m \le n \le N_2\end{aligned}$$

$$P_{im,jm-1} = mP_{s_2,m} E B_0(N_2-m)A_{j-i}\ R_{1,0} + 0\ (1-R_{1,0}); \quad 1 \le i \le j \le (Q_2-1),\ 1 \le m \le N_2$$

$$P_{im,i-1,n} = P_{s_2,m} B_{n-m}(N_2-m)A_0\ R_{1,0} + 0\ (1-R_{1,0}); \quad 1 \le i \le Q_2,\ 0 \le m \le n \le N_2$$

Impossible transitions

$$P_{im,jm-1} = 0; \quad 1 \le i \le Q_2,\ j < i, \quad 1 \le m \le N_2 \qquad P_{im,j} = 0; \quad 2 \le i \le Q_2,\ j < i-1, \quad 0 \le m \le n \le N_2$$

$$P_{im,i-1n} = 0; \quad 1 \le i \le Q_2,\ 1 \le m \le N_2, \quad n < m, \qquad P_{im,jn} = 0; \quad 0 \le i \le j \le Q_2,\ 2 \le m \le M_2, \quad n < m-1$$

Acknowledgement

This work has been supported by the Ministry of Economy and Competitiveness of Spain through projects TIN2013- 47272-C2-1-R and TEC2015-71932-REDT. Portillo-Jimenez thanks the support given by SEP-SES (DSA/103.5/15/6629) and EuroinkaNet project (EN15DF0205). Work done as part of the Program Financed by the European Union - Action 2- Erasmus Mundus Partnerships, GRANT AGREEMENT NUMBER - 2014 − 0870/001/001.

References

[1] Shu Yinbiao, Kang Lee, Peter Lanctot, F Jianbin, H Hao, B Chow, and JP Desbenoit. Internet of things: wireless sensor networks. *White Paper, International Electrotechnical Commission, http://www. iec. ch*, 2014.

[2] Ajinkya Rajandekar and Biplab Sikdar. A survey of MAC layer issues and protocols for machine-to-machine communications. *IEEE Internet of Things Journal*, 2(2):175–186, 2015.

[3] Wei Ye, John Heidemann, and Deborah Estrin. An energy-efficient MAC protocol for wireless sensor networks. In *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 3, pages 1567–1576. IEEE, 2002.

[4] Ou Yang and Wendi Heinzelman. Modeling and throughput analysis for SMAC with a finite queue capacity. In *Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2009 5th International Conference on*, pages 409–414. IEEE, 2009.

[5] Ou Yang and Wendi Heinzelman. Modeling and performance analysis for duty-cycled MAC protocols with applications to S-MAC and X-MAC. *IEEE Transactions on Mobile Computing*, 11(6):905–921, 2012.

[6] Jorge Martinez-Bauset, Lakshmikanth Guntupalli, and Frank Y Li. Performance analysis of synchronous duty-cycled MAC protocols. *IEEE Wireless Communications Letters*, 4(5):469–472, 2015.

[7] Lakshmikanth Guntupalli, Jorge Martinez-Bauset, Frank Y Li, and Mary Ann Weitnauer. Aggregated Packet Transmission in Duty-Cycled WSNs: Modeling and Performance Evaluation. *IEEE Transactions on Vehicular Technology*, 66(1):563–579, 2017.

A low-power IoT architecture proposal applied to preventive conservation of cultural heritage

Angel Perles[1], Ricardo Mercado[1], Eva Pérez-Marín[2], Manuel Zarzo[3], J. Damian Segrelles[4], Raúl Sáez[1], Antonio Martí[1], Ramón Ferrer[1], and Fernando J. Garcia-Diego[5]

[1] Instituto ITACA. Universitat Politècnica de València. 46022-Valencia, Spain
[2] Instituto Universitario de Restauración del Patrimonio. Universitat Politècnica de València. 46022-Valencia, Spain
[3] Department of Applied Statistics. Universitat Politècnica de València. 46022-Valencia, Spain
[4] Instituto de Instrumentación para Imagen Molecular (I3M), Centro Mixto CSIC - Universitat Politècnica de València. 46022-Valencia, Spain
[5] Department of Applied Physics. Centro de Investigación Acuicultura y Medio Ambiente ACUMA. Universitat Politècnica de València. 46022-Valencia, Spain

Abstract. Cultural Heritage has an important value that must be promoted and safeguarded for future generations. Preventive conservation of cultural heritage is a key aspect that could be enhanced thanks to Internet of Things (IoT) technologies. This work proposes and develops an specific IoT architecture solution focused on the most critical aspect of the architecture, which are the sensor nodes. The proposed solution is based on LoRa and Sigfox LP-WAN technologies and warrants a minimum impact in the artwork, achieving a lifespan of more than 10 years.

Keywords: cultural heritage, preventive conservation, Internet of things, cloud computing

1 Introduction

Cultural Heritage (CH) has an important value that must be promoted and safeguarded for future generations, because the deterioration or destruction of artworks is a serious loss that cannot be recovered in many cases. The protection and conservation of CH is an issue of particular interest because every artwork undergoes certain deterioration with time. Such degradation depends on the type of material, the action of external weather conditions and human factors. For example, in outdoor archaeological sites, the effect of wind, rain, solar radiation, water condensation and other factors can cause enormous damage.

The effects of ambient conditions on a given artefact are cumulative, so it is os special interest to record real-time measurements in order to detect as soon as possible unexpected harmful events or dangerous conditions, particularly in those places with risk of vandalism, robberies or accidents (e.g. water leaks), the on-line control of artworks would improve it safety and preventive conservation.

C.Fernandez-Llatas and M.Guillen (Ed.):
Workshop on Innovation on
Information and Communication Technologies
(ITACA-WIICT 2017)

ISBN 978-84-697-7327-7

Remote monitoring of factors affecting the conservation state of artworks may improve their long-term preservation. Nowadays, remote monitoring is mainly based on the Internet of Things (IoT) paradigm, where a "thing" could be any type of artefact. An IoT approach for art conservation would involve the installation of small sensor nodes and gateways for data transfer to the cloud. The application of this IoT approach would allow on-line monitoring and continuous supervision of individual pieces, given the easy access from the cloud to data recorded from electronic sensors, improving its safety and preventive conservation.

In most cases, given the particular requirements for the preventive conservation of pieces, it is necessary to design specific nodes for data transmission that should be able to efficiently monitor climatic conditions without for short-term maintenance. Therefore, the most critical issue for an IoT approach in CH is the appropriate design of nodes regarding power consumption and communication distances.

As the ideal technical solution does not yet exist, one of the targets of the present work was to apply our custom design based on a previous ITACA node and the convenience of using two leading options: LoRa [1] and Sigfox [2]. Both LoRa and Sigfox are LPWAN (Low-power Wide-area networks) commercial designs for wireless sensorization which are being extensively used for the paradigm of smart cities. All these technologies use the European ISM bands of 433 MHz and 868 MHz, and they have provided the best results in previous projects of our research group.

Given the technological evolution undergone in recent years in the field of IoT and cloud computing, new perspectives appear with respect to wireless microclimate monitoring of artefacts. The approach proposed in this work is based on microelectronic wireless systems of great autonomy able to record sensor data and to transfer them wirelessly, so that measurements can be processed or downloaded on-line by means of a computer connected to Internet or a mobile telephone. Apart from avoiding wiring, the main advantage is the possibility to access the data in real time, which allows the detection of abnormal situations that require urgent action.

In this work in progress, we describe briefly the proposed architecture and shows preliminary results of cultural heritage monitoring implemented in the Church of Santo Tomàs y San Felipe Neri in Valencia (Spain), an extended information about that work can be found at [3]. Preliminary tests on the nodes' wireless operations were carried out at this Church and proved to be satisfactory. Also, we evaluated the power requirements of the nodes for a standard monitoring problem of air temperature and humidity. The results suggest that it is possible to create a node with a life-span of more than 10 years using standard batteries.

The article is structured as follows; the second section presents the related work of previous studies. In the third section, we describes the proposed IoT solution for CH implementation. The fourth section is focused on testing the performance of the proposed node from the point of view of radio communica-

tions and energy requirements. Finally, the main conclusions and future work are presented in Section 5.

2 Related work

There are different types of installation systems for microclimatic monitoring. One approach is based on the use of wired sensors, which are typical in industry and domotic, but not in CH monitoring.

The present research group developed in 2007 [4, 5] a wired installation for the microclimate monitoring of the valuable fresco paintings discovered on the ceiling over the main chapel at the Cathedral of Valencia.

Another approach for artwork monitoring is the use of commercial dataloggers [6, 7], which are autonomous devices powered by batteries with the advantages of being easy to install, wire-free, and small. The main disadvantage is their limited data storage, the use of batteries with a certain lifespan, and the recorded data need to be manually downloaded from each datalogger. With a sampling rate of one datum per hour, which is adequate in many cases [8], dataloggers can work autonomously between 6-12 months [9, 10].

In a previous work performed in 2010, 10 autonomous hygrometric probes with temperature and humidity sensors were installed at the Roman ruins in Plaza de L'Almoina (Valencia, Spain) [11]. In this context it would be of interest to have real-time wireless sensors able to give warning signals in case of abnormal conditions.The use of wireless sensors for on-line monitoring would have allowed a direct access to the information from anywhere in the world, which would be of special interest in world-famous sites.

Some authors [8] have proposed both wired and wireless systems given their advantages and disadvantages. In any case, the complexity of the installation and the inconvenience of data downloading makes it difficult to achieve the real objective, which is to preserve the cultural heritage by appropriately splitting the work between restorers and those responsible for the installation.

Given the current enormous competition in the market that supplies products for the so-called IoT [12, 13] and the development of cloud computing, their application for preventive conservation of artworks is justified.

Our group has experience in the development of wireless sensor networks which have been applied satisfactorily for the monitoring of air conditions [14] and the wireless monitoring of wooden structures [15]. Such systems have turned out to be efficient for assessing the conservation conditions of wooden pieces such as altarpieces or carvings. Thanks to the autonomy achieved, the devices can work continuously for more than a decade without battery replacement, which notably reduces maintenance and the impact on artworks.

The deployment of efficient sensor nodes in archaeological sites, cathedrals or open areas is not straightforward. Sensor nodes use RF to transmit information that has to travel relatively long distances and/or pass through thick walls, moreover, the energy requirements must be kept low to give the nodes a long life-span.

Our practical experience in developing ultra-low power embedded systems for wireless sensor networks showed us that subGHz bands are the best for these scenarios. In a previous reasearch [15] we designed an ultra-low power sensor node appropriate for heritage building monitoring. This node was initially designed for monitoring the equilibrium moisture content in wood and the detection of termites [16]. The node uses the 868 MHz band in order to deal with thick solid materials, and the communication range turned out to be adequate. From our point of view, the main drawback of our implementation is the lack of standard communication protocols, which restricts interoperability with other available products on the market, reducing its applicability.

3 CH monitoring sytem

3.1 Architecture

Figure 1 represents the different types of scenarios related to CH monitoring. IoT deployments for CH has different needs compared with common applications as shown in Fig. 1, which tries to summarize the different scenarios. The key difference between common IoT deployment and one focused on CH preventive conservation is the node design (not unique) and the requirement for different types of gateways.

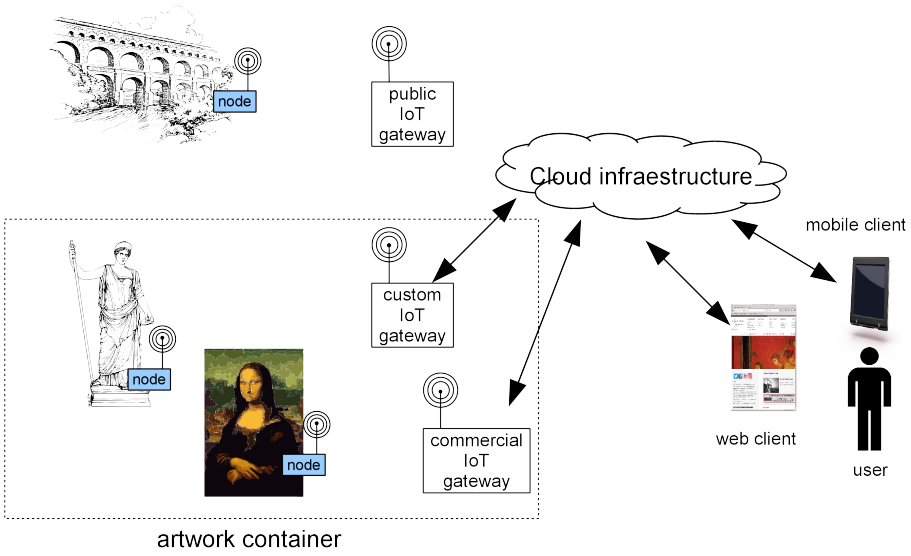


Fig. 1. IoT for CH diagram with different types of scenarios and gateways

In a typical IoT deployment [13] there are sensing/actuating nodes that communicate with the cloud infrastructure through gateways. These gateways are

responsible to collect the data provided by nodes in custom format or using typical IoT formats such as MQTT [17], to adapt these data formats and transport them to the cloud infrastructure using an Internet connection which may be wired-based or GSM-based.

In an art display (or archaeological excavation, church, museum, etc.) provided with an electric power supply and Internet connection, the best solution is to use an IoT commercial gateway. This gateway can be maintained active to collect data from the local nodes installed in the artworks. If connection-less protocols are utilized, the nodes do not require to spend too much energy for their communication, thanks to the relative proximity of the gateway and given the simplicity of the communication algorithm. Redundant gateways can be utilized to better cover the area of interest, the cloud infrastructure being responsible for eliminating data duplicates. In the case of large artworks (e.g. an altarpiece) or specific needs (e.g. a gas sensor), more than one sensor can be attached to the artwork.

For isolated pieces, for example a statue in the middle of a city, the best solution is to try to use a public IoT operator, provided that this service is available in the target country and area. For example, in Valencia (Spain), the Sigfox IoT provider has good coverage. Using Sigfox, it is feasible to develop an ultra-low power sensing node that transfers information directly to the cloud.

These two gateway types cannot cover open archaeological sites, or isolated churches, etc. In these cases, the main limiting factor is, often, the unavailability of power supply. To overdue this problem, we developed a prototype of open gateway architecture focused on dealing with energy-constrained environments and/or complex Internet connectivity that is described in this work.

3.2 Wireless node

For the node design, we fell back on our previous experience of building ultra-low power wireless nodes, updating the requirements and the lessons learned. The initial node design was based on a previous wireless node designed for Equilibrium Moisture Content (EMC) measurement and termite detection [15, 16] shown in Fig. 3 left, which can be inserted into wood on artworks. This node has a lifespan of more than 10 years using an 1 Ah battery.

The same design principles were applied to the CH node. Fig. 2 shows a block diagram in which the main principle is the use of loosely coupled blocks to facilitate node upgrading or adaptation (sensor types, battery technology and radio protocols). This is not an ideal design for massive market deployment, but it fits perfectly into the specification/evaluation phase.

Due to the highly constrained energy requirements, we opted to use an available ultra-low power microcontroller capable enough to support a wide variety of interfaces. The actual node is based on Intel 8051 architecture using a Silabs C8051F920 microcontroller. The node is designed for sesning temperature and relative humidity. It uses a Sensirion SHT1X sensor.

The radio-frequency (RF) part is uses the ISM unlicensed band of 868 MHz. This is a good compromise between range and antenna size.

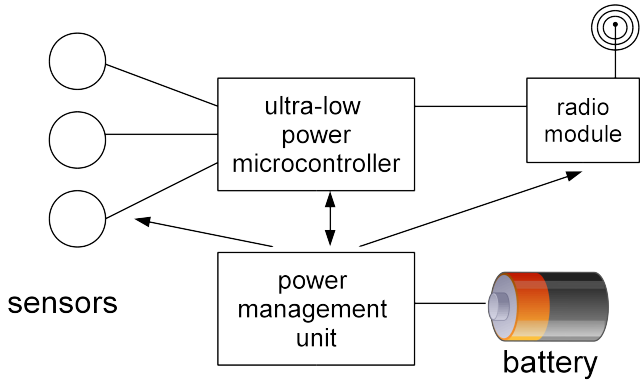


Fig. 2. Sensor node architecture

Finally, the Power Management Unit of the node is responsible for providing electric power to those components of the node that require energy at any given instant of time. This unit is more a concept than an independent block; for example, our custom node uses a simple output pin for providing energy to the RH and temperature sensor when a measure is requested. The idea is to completely eliminate the "stand-by" energy of all subsystems when not in use.

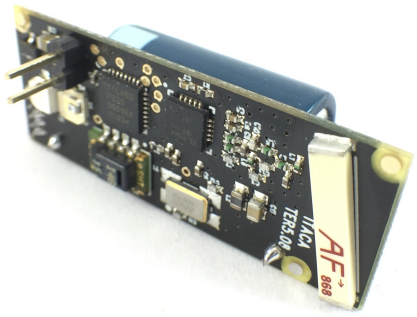


Fig. 3. ITACA's custom node for 868 MHz ISM band.

3.3 Gateway

Three different types of gateway were considered for the CH IoT platform: the Sigfox IoT operator gateway, a LoRa commercial gateway and a custom gateway for energy-constrained scenarios. In this paper we describe only our custom gateway.

When monitoring open archaeological sites in Cuenca (Spain) with no power and poor Internet connectivity using GSM services, we decided to design a low-power gateway architecture that was very rich from the point of view of openness, programming options and flexibility. This gateway was designed to deal with energy-constrained environments and/or complex Internet connectivity (satellite-based or private point-to-point antennas).

Our gateway proposal is to mix the benefits of general purpose microprocessors and ultra-low power microcontroller to keep energy requirements to a minimum. The only element that requires to be kept switched on is the radio module. The microcontroller is awakened every time a new packet arrives to the radio module, which is processed, stored in a local database and then returned to sleep mode. The general-purpose microprocessor is kept off until it is decided to recollect data, analyse them and, if desired, transfer them to the cloud. Fig. 4 shows a block diagram of this proposal.

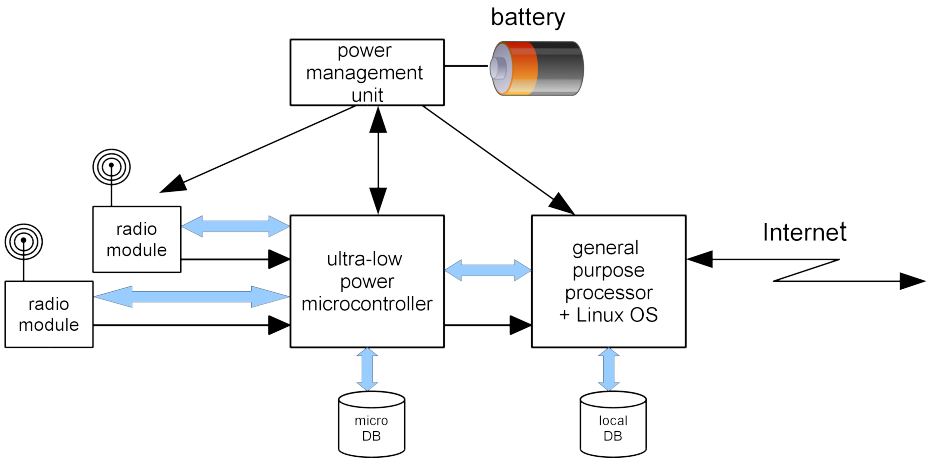


Fig. 4. Block diagram of the low-power CH gateway

To prove this concept we decided to build a demonstration prototype using widely available and well documented components. Fig. 5 shows a Raspberry Pi 2 with a ARM Cortex-A based general-purpose processor with Linux OS. To control the on/off cycle of the main processor, a Spell Foundry's Sleepy Pi board was selected; for the ultra-low power microcontroller we chose the same model as for the nodes in the form of an StMicroelectronics STM32L476-Nucleo board, and for the radio we used a Texas Instrument's CC1101EMK868-915 evaluation module, that pairs with our custom wireless node.

Fig. 5. Low-power gateway prototype

3.4 Cloud

The IoT for CH framework deployment proposed in this work uses a cloud infrastructure for hosting the server-part of the architecture using Amazon Web Services (AWS) public cloud provider. It hosts the databases, processing services inthe form of a Linux virtual machine and a web application to provide end users with a user interface including the real-time visualization of the data stored in the database.

The component called Infrastructure Manager (IM) [18, 19] was employed in this work, it allows the management of infrastructure deployments on a cloud provider and an open-source solution that can deploy customized virtual computational environments on a wide range of Cloud IaaS. The evolution of this platform includes all components of the framework, such as web interfaces (i.e Meteor), database back-ends (i.e. MongoDB) and processing components (i.e. SPARK) that are showed in Fig. 6, being the actual implementation a simplified proof-of-concept.

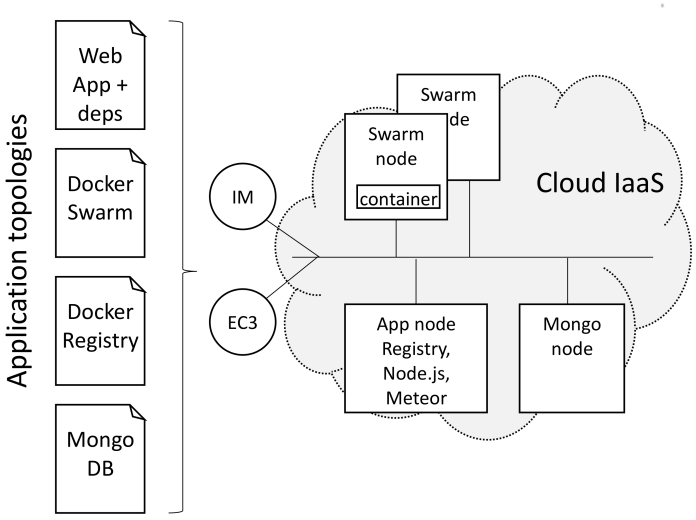


Fig. 6. Architecture of the cloud infrastructure

4 Example deployment and results

4.1 Historical building test site

For the case study we selected the baroque 17th century Church of Santo Tomás y San Felipe Neri in Valencia (Spain) (see Fig. 7), declared as National Monument

in 1982. Apart from the inherent value of the building itself, it contains an important collection of diverse artifacts, including works by renowned painters such as Juan de Juanes (1507 - 1579), Jerónimo Jacinto de Espinosa (1600 - 1667), José Vergara (1726 - 1799), and Vicente López (1772 - 1850), as well as murals, panels and several wooden altarpieces made to replace those destroyed in the Spanish Civil War.

Fig. 7. Aerial view of Santo Tomás y San Felipe Church at the centre of Valencia (Spain). View taken from Google Maps. Long: 39.47, Lat: -0.37

The target was to monitor the environment inside the church by considering different zones and a range of objects in order to determine possible risks. For example, we deployed a set of wireless nodes in the major altarpiece at the apse with the aim of studying the correlation between readings of environmental conditions with the moisture content in different parts of the wood. Fig. 8 shows this piece of art.

4.2 User interface

To demonstrate the suitability of the proposal from the point-of-view of user interaction, data stored in the AWS cloud is processed and showed to the user through a web page available at http://lpwan4itaca.blogs.upv.es/. Fig. 9 shows temperature and relative humidity rendered in this web service.

Fig. 8. Major altarpiece of Santo Tomás y San Felipe Church

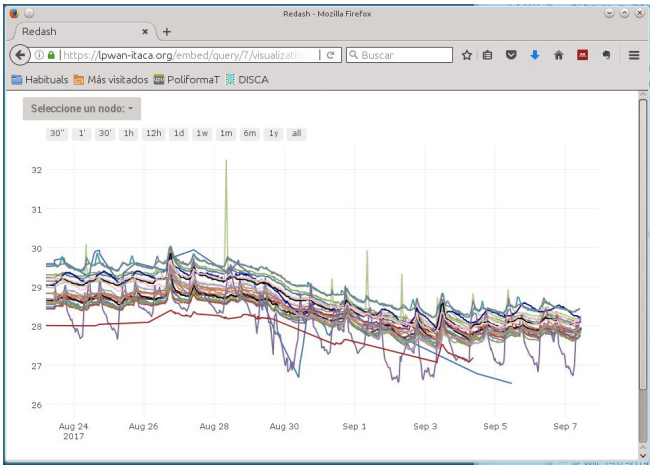


Fig. 9. Web-based view of temperature and relative humidity colleted information

4.3 Node's energy requirement

The lifespam of the sensor node is critical aspect of the proposal. In this subsection, we calculate the energy requirements of the node.

Total energy requirements are the is the sum of the energy required by all components. Each component will require different amount of energy depending on its state (e.g. turned on, sleeping, transmitting, receiving, etc.). A lower quantity of energy is required when the components are sleeping o, better, completely turned off so these states should be maximized to get long life nodes. Table 1 summarizes these energy requirements in different modes, were energy requirements has been obtained from manufacturer's datasheet and laboratory measurements.

Table 1. Current requirements for the different states of the main electronic components @ 3 VDC

Description	**Current (μA)**
Microcontroller Silabs C8051F920 sleep + RTC	1.00
Microcontroller Silabs C8051F920 active	$3.00 \cdot 10^3$
Sensor Sensirion SHT1x sleep	0.30
Sensor Sensirion SHT1x active	$0.90 \cdot 10^3$
RF Texas Inst. CC1101 sleep	1.00
RF Texas Inst. CC1101 act.+trans.	$35.00 \cdot 10^3$

In order to evaluate the energy requirements, we established one sampling per hour of the RH and temperature parameters. Given this sample rate, table 2 indicates the energy requirements in a per day basis and the annual amount of energy required by each node

Table 2. Energy required for each component assuming one sampling per hour

Description	**Daily time (s)**	**Annual ener.(mAh)**
Micro. Silabs C8051F920 sleep + RTC	$86.39 \cdot 10^3$	8.76
Micro. Silabs C8051F920 active	2.00	0.61
Sensor Sensirion SHT1x sleep	0.00	0.00
Sensor Sensirion SHT1x active	12.00	1.10
RF Texas Inst. CC1101 sleep	$86.39 \cdot 10^3$	8.76
RF Texas Inst. CC1101 act.+trans.	9.00	31.94

Adding together the energy required for each component at each state, we obtained the total amount of energy required during a year of operation. Battery lifespan can be estimated from these values; for example, Table 3 summarizes the estimated life of a typical coin-cell battery (CR2032) and two high-density energy batteries assuming 75% of usable energy. We choose lithium technology because they exhibit very low self-discharge figure, so they are for long-life systems, but as it is not negligible.

Table 3. Estimated battery life

Battery model	**Capaity**	**Lifespam**
Battery model	**(mAh)**	**(years)**
Panasonic CR2032	220.00	3.0
Maxell Lithium-tionil 2/3 AA 3.6 V	1,700.00	23.2
Maxell Lithium-tionil AA 3.6 V	3,100.00	42.3

As shown in table 3, more than 10 years of battery lifespam can be achieved using one sample per hour of data.

5 Conclusions and future work

The present work in progress studies the viability of applying IoT technologies to CH by analysing the requirements and proposing the appropriate cloud and field systems for these requirements. The results and a first deployment demonstrates that it is feasible to develop a preventive conservation campaign of artworks by evaluating the environmental conditions.

The node's design based on 868 MHz ISM band showed adequate to cope with the need of reaching relative long distances and pass through thick walls that are typical scenarios in artwork containers or open archaeological sites.

The energy requirements were then calculated based on the typical RH and temperature measurement requirements at a rate of one sample per hour. Using a small 1.7 mAh lithium-thionyl battery it is possible to achieve a node life-span better than 10 years with no maintenance.

Since the results achieved were excellent we are developing new nodes based on LoRa and Sigfox technology and plan to introduce automatic analysis algorithms than can detect harmful conditions of the artefacts.

Acknowledgements This work was partially funded by the Generalitat Valenciana's project AICO/2016/058 and by the Plan Nacional de I+D, Comisión Interministerial de Ciencia y TecnologíA (FEDER-CICYT) under the project HAR2013-47895-C2-1-P.

The authors wish to express their gratitude to the members of the Parish of Santo Tomás and San Felipe Neri for their collaboration in this project.

References

References

1. LoRa: Lora alliance (2017)
2. Sigfox: Sigfox - the global communications service provider for the internet of things (iot) (2017)
3. Perles, A., Prez-Marn, E., Mercado, R., Segrelles, J.D., Blanquer, I., Zarzo, M., Garcia-Diego, F.J.: An energy-efficient internet of things (iot) architecture for preventive conservation of cultural heritage. Future Generation Computer Systems (2017)
4. Garca-Diego, F.J., Zarzo, M.: Microclimate monitoring by multivariate statistical control: The renaissance frescoes of the cathedral of valencia (spain). J. Cultural Heritage **11** (2010) 339–344
5. Zarzo, M., Fernndez-Navajas, A., Garca-Diego, F.J.: Long-term monitoring of fresco paintings in the cathedral of valencia (spain) through humidity and temperature sensors in various locations for preventive conservation. Sensors **11**(9) (2011) 8685–8710
6. Visco, G., Plattner, S.H., Fortini, P., Sammartino, M.: A multivariate approach for a comparison of big data matrices. case study: thermo-hygrometric monitoring inside the carcer tullianum (rome) in the absence and in the presence of visitors. Environmental Science and Pollution Research (2017) 1–15
7. Carcangiu, G., Casti, M., Desogus, G., Meloni, P., Ricciu, R.: Microclimatic monitoring of a semi-confined archaeological site affected by salt crystallisation. Journal of Cultural Heritage **16**(1) (2015) 113 – 118
8. García-Diego, F.J., Verticchio, E., Beltrán, P., Siani, A.M.: Assessment of the minimum sampling frequency to avoid measurement redundancy in microclimate field surveys in museum buildings. Sensors **16**(8) (2016)
9. Merello, P., Garcia-Diego, F.J., Zarzo, M.: Microclimate monitoring of ariadnes house (pompeii, italy) for preventive conservation of fresco paintings. Chemistry Central Journal **6** (2012) 145
10. Merello, P., Garcia-Diego, F.J., Zarzo, M.: Evaluation of corrective measures implemented for the preventive conservation of fresco paintings in ariadne's house (pompeii, italy). Chemistry Central Journal **7** (2013) 87
11. Fernández-Navajas, A., Merello, P., Beltrán, P., García-Diego, F.J.: Multivariate thermo-hygrometric characterisation of the archaeological site of plaza de l'almoina (valencia, spain) for preventive conservation. Sensors **13**(8) (2013) 9729–9746
12. Gubbi, J., Buyya, R., Marusic, S., Palaniswami, M.: Internet of things (iot): A vision, architectural elements, and future directions. Future Generation Computer Systems **29**(7) (2013) 1645 – 1660 Including Special sections: Cyber-enabled Distributed Computing for Ubiquitous Cloud and Network Services. Cloud Computing and Scientific Applications Big Data, Scalable Analytics, and Beyond.
13. Gershenfeld, N., Krikorian, R., Cohen, D.: Internet of things. Scientific American (2004)
14. Calvet, S., Campelo, J.C., Estells, F., Perles, A., Mercado, R., Serrano, J.J.: Suitability evaluation of multipoint simultaneous co2 sampling wireless sensors for livestock buildings. Sensors **14**(6) (2014) 10479–10496
15. Capella, J.V., Perles, A., Bonastre, A., Serrano, J.J.: Historical building monitoring using an energy-efficient scalable wireless sensor network architecture. Sensors **11**(11) (2011) 10074–10093

16. Perles, A., Mercado, R., Capella, J.V., Serrano, J.J.: Ultra-low power optical sensor for xylophagous insect detection in wood. Sensors **16**(11) (2016)
17. MQTT: Mq telemetry transport (2017)
18. Caballer, M., Blanquer, I., Moltó, G., de Alfonso, C.: Dynamic Management of Virtual Infrastructures. Journal of Grid Computing **13**(1) (mar 2015) 53–70
19. Caballer, M., Segrelles, D., Moltó, G., Blanquer, I.: A platform to deploy customized scientific virtual infrastructures on the cloud. Concurrency and Computation: Practice and Experience **27**(16) (nov 2015) 4318–4329

New configurations to improve reliability and redundancy in high performance memory systems

J. Gracia-Morán[1], P.J. Gil-Vicente[1], D. Gil-Tomás[1], L.J. Saiz-Adalid[1],

[1] Instituto ITACA, Universitat Politècnica de València,
Camino de Vera s/n, 46022, Spain
{jgracia, pgil, dgil, ljsaiz}@itaca.upv.es

Abstract. New fault tolerant methods are needed to cope with the fault rate augment in memory systems. Traditionally, Error Correction Codes (ECCs) have been used. This Fault-Tolerance method works well with single faults. Nevertheless, the increase of the integration density in current deep submicron chips, as well as the decrease of the energy needed to provoke Single Event Upsets (SEUs) in storage, has provoked the augment in the occurrence of Multiple Cell Upsets (MCUs). In this way, new ECCs able to tolerate MCUs are needed. In this work, we summarize the different ECCs that we have proposed in order to tolerate MCUs.

1 Introduction

As memory area grows, also its probability of suffering faults [1]. In this way, the reduction of the energy needed to provoke a Single Event Upset (SEU) in storage has been reduced, as it can be seen in Fig. 1 (extracted from [2]). Conversely, this energy reduction provokes an augment of Multiple Cell Upsets (MCU) [3][4].

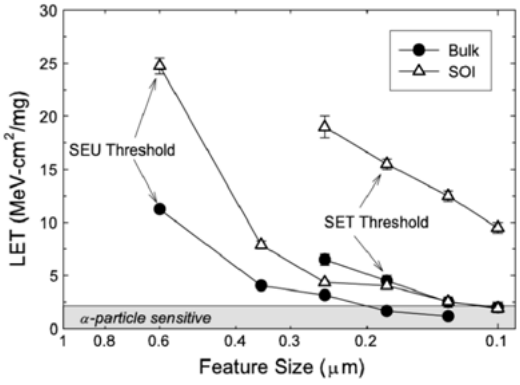


Fig. 1. Simulated critical LET (Linear Energy Transfer) for unattenuated transient propagation and SEU threshold LET as a function of scaling for bulk and SOI CMOS technologies (extracted from [2]).

C.Fernandez-Llatas and M.Guillen (Ed.):
Workshop on Innovation on
Information and Communication Technologies
(ITACA-WIICT 2017)

ISBN 978-84-697-7327-7

Traditionally, Error Correction Codes (ECCs) were used to protect data stored in memories, specially Single Error Correction (SEC) codes or Single Error Correction-Double Error Detection (SEC-DED) codes [5][6][7]. SEC codes are able to correct an error in one single memory cell, while SEC-DED codes can correct an error in one single memory cell, as well as detect two errors in two independent cells.

In any case, it is necessary to add extra bits in order to carry on fault detection and correction. Currently, ECC memories add 8 redundant bits (also called code bits) for 64-bit data words, that is, a 12.5% of redundancy [8]. The problem arises when MCUs are present, as this number of code bits may not be sufficient to correct them.

In this work, we summarize the different ECCs that we have proposed in order to tolerate MCUs. The common characteristic of these ECCs, together with special memory configurations, is that they improves reliability on memory systems preserving, or even decreasing, the redundancy of actual memory protection methods.

This paper is organized as follows. Section 2 summarizes how to improve reliability in DRAM storage systems. Section 3 introduces the basic properties of the new ECCs that improve reliability in DRAM memories, and finally, Section 4 concludes this work and presents some future work.

2 Improving reliability in DRAM devices

As just commented, Single Error Correction (SEC) codes or Single Error Correction-Double Error Detection (SEC-DED) codes were traditionally used to protect memories. An important characteristic of these codes is that the redundancy they introduce decreases with the length of the data word, as it can be seen in Table 1 (extracted from [9]). Another fact we can see in Table 1 is that for the same data word length, a greater error coverage needs a greater redundancy.

Table 1. Redundancy and coverage for common data word lengths (extracted from [9])

ECC	Redundancy	Coverage
SEC (12, 8)[1]	50%	Single Error Correction
SEC (21, 16)	31.25%	
SEC (38, 32)	18.75%	
SEC (71, 64)	10.94%	
SEC-DED (13, 8)	62.50%	Single Error Correction – Double Error Detection
SEC-DED (22, 16)	37.50%	
SEC-DED (39, 32)	21.88%	
SEC-DED (72, 64)	12.50%	

As commented previously, modern ECC memories introduce a 12.5% of redundancy (that is, 8 redundant bits for 64-bit data words). These memories are usually built with 4-bit or 8-bit DRAM chips. In this way, to construct a 64-bit data word plus 8 redundant bits, eighteen 4-bit or nine 8-bit DRAM chips are used. Nevertheless, this

[1] An (n, k) binary ECC encodes a k-bit input word in an n-bit output word. In this way, the number of redundant bits introduced can be calculated as the result of ($n - k$).

12.5% of redundancy is not enough to cope with MCUs. Some changes have to be introduced. For instance, some advanced methods able to tolerate MCUs, such as IBM's Chipkill Error Protection scheme [10], Hewlett Packard's Advanced ECC [11] or Intel's Single Device Data Correction (SDDC) [12], interleave data and parity bits through different memory chips in order to spread errors in different ECC words.

Another change is the data word length. Currently, memory channels can be joined in order to feed with enough data multi-threaded applications running in multi-core processors. This action can be implemented with the lockstep mode [13][14]. This method runs the same memory command in various channels at a time. In this way, it is possible to combine data from these various channels to form a data word with a bigger length, and as it has just seen in Table 1, longer data word length allows the reduction of the redundancy used by the ECCs.

In a previous work, we have defined the FUEC (Flexible Unequal Error Control) methodology [16]. By using an algorithm based in this methodology, we are able to design ECCs with a very low redundancy. In this way, when designing an ECC using the FUEC methodology, four parameters must be set:

- the data length (k);
- the encoded word length (n);
- the set of error vectors to be corrected ($\boldsymbol{E}_+$); and
- the set of error vectors to be detected ($\boldsymbol{E}_\Delta$).

The number of redundant bits needed by an ECC depends on the size of the set of error vectors to be corrected ($\boldsymbol{E}_+$), as well as the size of the set of error vectors to be detected ($\boldsymbol{E}_\Delta$). In this way, by reducing these sets, we can reduce the number of extra bits needed. More information about this reduction can be seen in [15].

$\boldsymbol{E}_\Delta$ and $\boldsymbol{E}_+$depends on the type of errors to detect and correct. The term *random error* commonly refers to one or more bits in error, distributed randomly in the encoded word (data bits plus parity bits generated by the ECC). *Random errors* can be *single* (only one bit affected) or *multiple*.

Single errors are the simplest ones. They are commonly produced by single event upsets (SEU) in random access memories, or single event transients (SET) in combinational logic [17].

Multiple errors usually manifest as *burst errors*, rather than *randomly* [18]. A *burst error* is a multiple error that spans l bits in a word [19], i.e. a group of contiguous bits where, at least, the first and the last bits are in error. The separation l is known as burst length. The main physical causes of a burst error in the context of DRAM memories are diverse: high energy cosmic particles that hit some neighbor cells, crosstalk between neighbor cells, etc. [20].

3 New ECCs to improve memory reliability

We have designed a series of new ECCs that improve memory reliability. To do this, we have used:

- Interleaving. By using this method, it is possible to spread a multibit error in different ECC words. In this way, the corresponding ECC must correct a lower

number of bits in error. Thus, the combination of all ECCs included in a memory system allows supporting greater multibit errors.

- Lockstep multichannel to enlarge the data word size. As we have seen in Section 2, longer data words length allows a greater coverage and reduces redundancy.
- Efficient reduction of the space of error vectors. It is possible to design efficient and low redundant ECCs by reducing the number of error vectors. This is possible by taking into account the data storage physical distribution. We have decreased the number of error vectors by eliminating those error vectors that represents very uncommon errors.
- FUEC methodology with an algorithm developed by the authors. Once the different parameters to design an ECC have been established, our algorithm is able to find a parity check matrix that defines this ECC. By using the FUEC methodology, this ECC will be very efficient in terms of area and/or speed.

Table 2 and Table 3 show a summary of these new ECCs. Next paragraphs summarize the main characteristics of the different ECCs. A more detailed description can be found in [9][15].

Table 2. ECCs' Fault Tolerance Capabilities Summary (I)

	ECC1	***ECC2***	***ECC3***
Data Word Size	64	128	128
Parity bits per ECC	8	8	12
Redundancy	12.5%	6.25%	9.38%
Interleaving	Yes	Yes	Yes
Memory Channels	2	4	4
Correction capabilities	Complete 4-bit DRAM device Single bit errors 2- and 3-bit burst errors	Complete 4-bit DRAM device Single bit errors 2- and 3-bit burst errors	Complete 8-bit DRAM device Single bit errors 2- to 5-bit burst errors

3.1 Error Correction Code 1 (ECC1)

The ECC1 uses 2 ECCs with 64 data bits and 8 code bits each. In this way, by using 2 memory channels, a 128-bit data word and 16 code bits are used in this memory configuration. Thus, redundancy is 12.5%.

This ECC has been designed for memory DIMMs built with 4-bit DRAM devices. By interleaving data and code bits in pairs, this memory configuration is able to correct single random errors, as well as 2- and 3-bit burst errors. Also, this memory configuration can recover the error of a complete 4-bit DRAM device. As far as we know, this memory configuration presents the highest fault tolerance capabilities with the lowest redundancy for 128-bit data words.

Table 3. ECCs' Fault Tolerance Capabilities Summary (II)

	ECC4	*ECC5*	*ECC6*
Data Word Size	128	256	256
Parity bits per ECC	8	12	10
Redundancy	6.25%	4.7%	3.9%
Interleaving	Yes	Yes	Yes
Memory Channels	8	8	8
Correction capabilities	Two adjacent 4-bit DRAM devices or a single 8-bit DRAM device 1 bit in error 2- to 7-bit burst errors	Single 8-bit DRAM device 1 bit in error 2- to 5-bit burst errors	Single 4-bit DRAM device 1 bit in error 2- and 3-bit burst errors

3.2 Error Correction Code 2 (ECC2)

The second ECC introduced (ECC2) uses also 2 ECCs, but now each ECC generates 8 code bits from a 128-bit data word. That is, the complete scheme uses now a 256-bit data word and 16 code bits. To do this, 4 memory channels are needed. In this way, we have introduced a 6.25% of redundancy, half ECC1 code.

This ECC has been also designed for memory DIMMs built with 4-bit DRAM devices. In this way, by interleaving data and code bits in pairs, it is possible to correct single random errors, as well as 2- and 3-bit burst errors. As in the case of the ECC1, this memory configuration can also recover the error of a complete 4-bit DRAM device.

Another property of this code is that if we use "standard" memory DIMMs (64-bit data word and 8 code bits), we have now 16 unused code bits. These bits can be used as spare bits, increasing the fault tolerance capabilities of the complete memory system. The use of spare bits is a known method utilized by different memory protection schemes, such as the IBM's Memory ProteXion method [21].

3.3 Error Correction Code 3 (ECC3)

In this case, we have augmented the redundancy used by the ECC in order to be able to use 8-bit DRAM devices. Specifically, we have now two ECCs with a 128-bit data word and 12 code bits each one, getting a 9.38% of redundancy. As in the case of ECC2, we have used also 4 memory channels, but now, we have interleaved data and code bits in groups of 4 bits. With this configuration, this memory scheme is able to correct single bit random errors, as well as 2- to 5-bit burst errors. Also, the failure of a complete 8-bit DRAM device is supported.

As in the previous case, such a lower redundancy has provoked the presence of some unused bits. Specifically, we have now 8 bits that can be used as spare bits, increasing the fault tolerance capabilities of the complete system.

3.4 Error Correction Code 4 (ECC4)

This code can be used by 4-bit or 8-bit DRAM devices. In this case, we have used 8 memory channels. Combining 4 ECCs (with 128-bit data word and 8 code bits each ECC), we can built a 512-bit data word with 32 code bits, provoking a 6.25% of redundancy.

By interleaving data and code bits in pairs, this memory configuration can correct one random bit in error, as well as from 2- to 7-bit burst errors. Also, this memory scheme is able to correct the failure of two adjacent 4-bit DRAM devices or a single 8-bit DRAM device. Lastly, there are also 32 unused bits, which can be employed as spare bits.

3.5 Error Correction Code 5 (ECC5)

This code has been designed for 8-bit DRAM memories. In this case, the redundancy is 4.7%, as this code uses 12 parity bits for 256 data bits. By combining 2 of these ECCs, and interleaving bits in groups of 4, this memory protection format can correct all single errors, as well as 2- to 5-bit burst errors. Also, a complete single 8-bit DRAM device can be recovered. On the other hand, we also have 40 spare bits.

3.5 Error Correction Code 6 (ECC6)

This code, designed for 4-bit DRAM devices, uses two ECCs with 256-bit data word and 10 code bits each. As in the previous case, we have used 8 memory channels. Combining 2 ECCs, we can build a 512-bit data word with 20 code bits, provoking a 3.9 % of redundancy, that is, the lowest redundancy of the ECCs presented in this work.

This so low redundancy provokes that this design presents the greatest quantity of spare bits available respect of the ECCs introduced in this work. Specifically, with this ECC, there are 44 spare bits.

In this way, with this memory scheme, we are able to correct one bit in error, 2- and 3-bit burst errors and the complete failure of a single 4-bit DRAM device.

4 Conclusions and Future work

As memory capacity increases, also its fault rate augments, provoking an increment in the number of SEUs and MCUs. As traditional ECCs cannot cope with this fault rate increment, new ECCs are needed.

By designing efficient ECCs, using various memory channels in lockstep mode, interleaving bits and applying our algorithm based in the FUEC methodology, it is possible to design advanced memory systems able to work with the new fault rates. These new codes are able to correct MCUs and improve the redundancy and reliability of actual commercial solutions. In addition, they are able to tolerate the failure of a whole memory chip.

In the future, we want to continue developing very low redundant ECCs, as well as to study how to tolerate the failure of several DRAM devices.

References

1. J. Meza, Q. Wu, S. Kumar, and O. Mutlu, "Revisiting Memory Errors in Large-Scale Production Data Centers: Analysis and Modeling of New Trends from the Field", 45th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN 2015), pp. 415-426, June/July 2015.
2. P.E. Dodd, M.R. Shaneyfelt, J.A. Felix, and J.R. Schwank, "Production and Propagation of Single-Event Transients in High-Speed Digital Logic ICs", IEEE Trans. on Nuclear Science, vol. 51, no. 6, pp. 3278-3284, December 2004.
3. R. Baumann, "Soft errors in advanced computer systems," IEEE Design & Test of Computers, vol. 22, no. 3, pp. 258–266, May/Jun. 2005.
4. E. Ibe, H. Taniguchi, Y. Yahagi, K. Shimbo, and T. Toba, "Impact of scaling on neutron-induced soft error in SRAMs from a 250 nm to a 22 nm design rule," IEEE Trans. Electron Devices, vol. 57, no. 7, pp. 1527–1538, July 2010.
5. E. Fujiwara, Code Design for Dependable Systems: Theory and Practical Application, Ed. Wiley-Interscience, 2006.
6. R. W. Hamming, "Error detecting and error correcting codes," Bell System Technical Journal, vol. 29, pp. 147 – 160, 1950.
7. C.L. Chen and M.Y. Hsiao, "Error-correcting codes for semiconductor memory applications: a state-of-the-art review", IBM Journal of Research and Development, vol. 58, no. 2, pp. 124-134, March 1984.
8. Hewlett-Packard Development Company, L.P. "Memory technology evolution: an overview of system memory technologies", Technology brief, 9th edition, Dec. 2010.
9. J. Gracia-Morán, P. J. Gil-Vicente, D. Gil-Tomás, L. J. Saiz-Adalid, "Improving Error Correction Capabilities on Advanced ECC Memories by using Multichannel Configurations", Under Review, IEEE Transactions on Very Large Scale Integration (VLSI) Systems, May 2017.
10. T.J. Dell, "A white paper on the benefits of Chipkill-correct ECC for PC server main memory", *IBM Microelectronics Division*, 1997.
11. Hewlett-Packard Development Company, L.P. "Memory technology evolution: an overview of system memory technologies", Technology brief, 9th edition, Dec. 2010.
12. Intel, "Intel® E7500 Chipset MCH Intel® x4 Single Device Data Correction (x4 SDDC) Implementation and Validation", Application Note (AP-726), Aug. 2002, available online at http://www.ece.umd.edu/courses/enee759h.S2003/references/29227401.pdf.
13. T. Willhalm, "Independent Channel vs. Lockstep Mode – Drive your Memory Faster or Safer", Intel, July 2014, available online at: https://software.intel.com/en-us/blogs/2014/07/11/independent-channel-vs-lockstep-mode-drive-you-memory-faster-or-safer
14. J. Hoskins and B. Wilson, "Exploring IBM Eserver Xseries and PCs: An Instant Insider's Guide to Intel-Based Servers and Systems", Ed. Maximum Press, 2002.
15. J. Gracia-Morán, L. J. Saiz-Adalid, D. Gil-Tomás, P. J. Gil-Vicente, "Improving Reliability in High Performance DRAM Memories using new Error Correction Codes for 8 Memory Channels", Under Review, IEEE Transactions on Emerging Topics in Computing, March 2017.
16. L.J. Saiz-Adalid, P.J. Gil-Vicente, J.C. Ruiz, D. Gil-Tomás, J.C. Baraza, and J. Gracia-Morán, "Flexible Unequal Error Control Codes with Selectable Error Detection and Correc-

tion Levels", 32th International Conference on Computer Safety, Reliability and Security (SAFECOMP 2013), pp. 178-189, September 2013.
17. K.A. LaBel, "Proton single event effects (SEE) guideline" submitted for publication on the NASA Electronic Parts and Packaging (NEPP) Program web site, August 2009. Available online at https://nepp.nasa.gov/files/18365/Proton_RHAGuide_NASAAug09.pdf.
18. S. Shamshiri and K.T. Cheng, "Error-Locality-Aware Linear Coding to Correct Multi-bit Upsets in SRAMs," IEEE International Test Conference, pp. 1-10, November 2010.
19. E. Fujiwara, Code Design for Dependable Systems: Theory and Practical Application, Ed. Wiley-Interscience, 2006.
20. M. Greenberg, "Reliability, availability, and serviceability (ras) for ddr dram interfaces," in memcon.com. Available online at: http://www.memcon.com/pdfs/proceedings2014/NET105.pdf, 2014.
21. D. Watts, R. Moon, "IBM System x3950 M2 and x3850 M2. Technical Introduction", IBM Redbooks, 2008, available online at: http://www.gebruikteservers.nl/upload/1304416621.pdf.

New technologies applied as alternative approaches in averting depression

Harold Kazungu Woods[1], Sabina Asensio-Cuesta[1], Juan Miguel García-Gómez[1]

[1] ITACA Institute - Universitat Politècnica de València, Camino de Vera s/n. 46022 Valencia, Spain

hawooka@doctor.upv.es, sasensio@dpi.upv.es, juanmig@ibime.upv.es

Abstract. The World Health organization (WHO) anticipates depression to be a disorder with the leading disease burden in developed countries by 2030. Today, depression is a global health concern rated third behind respiratory and cardia diseases among the major global causes of disability. Thus, a noteworthy disease burden on both individuals and healthcare systems globally. However, only an estimated small number of people get treatment and many go untreated, for example; an estimated 14 million people in Europe and 20 million in North and South America (combined) within a 12-month period. With the widespread adoption of e-mental health and rapid advances in technology, a unique opportunity arises for tackling mental health problems. Though evidence of rates of effectiveness across diverse settings suggests varied results, through E-mental health programs exits an inherent opportunity for early prevention depression and in-stage treatment of depression. This has sparked a growing interest in viable technology-based treatments for depression that can reach a vast number of end-users, including computer-based self-help solutions. Technological advancements in recent years paved way for a rapid growth in the use of computer-based technology approaches as alternatives to traditional approaches in the assessment, evaluation and treatment of depression. The scope of application of these new technological interventions, methods, contexts and their effectiveness as an alternative applied to depression is presented.

1 Introduction

Notably, depression is synonymous with occupational and interpersonal impairment; and a significant influence on health and productivity. And alone accounts for 2.5% of the global disease burden ranking as the prime single root cause of disability globally with an estimated 8.2% of all years lived with disability globally. A study on the Epidemiology of Major Depressive Episode (MDE) in a Southern European Countries by the European Study of the Epidemiology of Mental Disorders (ESEMeD) concluded that in Spain alone, the lifetime prevalence of depression in Spain reached an estimated 10.6% with a 12-month prevalence of 3.9% [1,2]. In comparison, the yearly prevalence of significant depressive disorder reached 6.6% in the USA, and 3.9% in Canada. The basic characteristic of depression is a loss of positive affect which manifests itself in a range of symptoms, including sleep disturbance, lack of self-interest,

C.Fernandez-Llatas and M.Guillen (Ed.):
Workshop on Innovation on
Information and Communication Technologies
(ITACA-WIICT 2017)

ISBN 978-84-697-7327-7

motivation, thoughts among others [3]. Depression has been of equal concern among adolescents, adults and the elderly.

Depression is prevalent with a reported up to 50% prevalence and symptoms of depression among primary care patients generally though often undiagnosed [4]. Technological progress can potentially transform mental healthcare, for example healthcare services, data and connect patients. Contrary to traditional approaches of healthcare for depression prevention, increased use of computer-based technological approaches as alternatives in evaluation, prevention and treatment of depression spans beyond the realm of technology, but to also represent a healthcare transformation that empowers patients with greater accessibility to information and services, choice of healthcare services, and control; and enhance integrated patient management and real-time patient data for healthcare services providers and professionals.

Historically, though face-to-face patient consultations in clinic settings have been the norm of mental health intervention from professionals, substantial evidence exists to support use of cognitive behaviour therapy (CBT) as an alternative which has inspired Computerised cognitive behaviour therapy (CCBT) - a technology based self-help CBT option with minimal involvement from healthcare professionals. This may include some non-specific features of the therapy relationship such as empathy, alliance and motivation that have manifested to be effective [5]. As a result, some programs are offered as an alternative depression and anxiety treatment within healthcare, for example "Beating the Blues" in the UK [6] which coupled with several studies of web-based self-help interventions have proven effectiveness in reduction of symptoms. Thus, there is considerable support for the use of the ICT-based approaches for delivering evidence-based psychotherapy for depression [7]. Whereas depression patients may ignore seeking help for many reasons ranging from reluctance to stigma among other reasons, technological-based interventions offer alternatives to such patients that uncompliant with the current modes of psychological therapy delivery and treatment.

Effective treatment programs for depression in E-mental health [8] for both depression and anxiety are based on interpersonal therapy (IPT) or cognitive behavioural therapy (CBT) models [9]. To date, numerous technological approaches have been applied to depression including; mobiles, 2D & 3D gaming, virtual reality, virtual agents and mobile phones that can be categorised between systems designed for use in prevention of mental illness, computer-based treatment and self-help frameworks, and systems to compliment face-to-face psychotherapy. Given the variety in contexts of use, systems can be aligned with treatment objectives for patients, for example; virtual reality (content interaction), self-help for patient self-monitoring, communication and patient-specific content delivery, and delivery of content medium. Thus, providing insights into additional insight into patient behavioural patterns, adaptability and progress that later inform the when (early) and how an intervention for depression should be offered [10], and collective decision-making.

E-mental health has the potential to have far reaching impacts beyond addressing only the needs of depression patients and those at or above the depression threshold by

advancing the secondary prevention – compatible with current global public health directives to reduce the burden of depression [11]. Digital technology favours self-help approaches (12,13]; averts stigma; to offer faster, easier, and more (cost-) effective access to help [12-16] and to provide a more neutral space in which service users can speak more freely [12, 17]. Nevertheless, patient and clinician involvement is still crucial to the evaluation of depression treatment technologies for reasons of public trust, effectiveness and patient satisfaction because significant doubts still exist in the evidence-base underlying these technologies.

In view of the significance, new technologies developed for the prevention, evaluation and treatment of depression are presented. The review explores (1) the kind of technologies and programs that have been developed for depression (2) how these technologies and programs are being applied (3) contents and modes of delivery (4) their context of use (5) user profiles definition and (6) clinical relevance of the technologies. The focus is profoundly on the programs and technologies designed for the outpatient treatment; including CCBT interventions and programs in the form of web-based systems, virtual reality, IVR systems, and dialogue platforms. The questions; (1) what types of e-mental health interventions have been developed and evaluated for depression? and (2) how these technologies are being used in identified interventions?

2 Methods

2.1 Identification of studies and Programs

A comprehensive literature review of PubMed, PsycINFO, BioMed Central, CINAHL and MEDLINE databases were conducted to identify programs and published peer reviewed studies on technologies applied to the prevention, evaluation and treatment of depression. Initially, no limitations were placed on years included in the search but later the search was restricted to 2000-2016. Relevant publications were first checked by abstract and later scrutinised for method and content(s), pivoted on the inclusion criteria of the technologies. Additional articles were identified through the reference lists of the retrieved articles and previous review studies. Additionally, a state-of-the-art (scoping) review was conducted of programs, projects and software applied to the treatment of depression in Cordis, NHS and the internet search engines (Google and Yahoo).

2.1.1 Search strategy

The interdisciplinary nature of this research demanded a comprehensive search of various databases including PubMed, PsycINFO, BioMed Central, CINAHL, MEDLINE, EBSCO, Essential Evidence Plus, Evidence-Based Medicine Reviews,

and Health Reference Centre. Keywords included a combinations of MeSH terms and keywords often synonymous with research in depression and e-Mental Health (see Table 1). These keywords were used to search titles or abstracts of potentially eligible articles published from 2000 to 2016. Reference lists of existing reviews and meta-analyses were further investigated for any eligible studies. The initial search returned 1072 eligible citations of which included a few systematic reviews and meta-analyses that further led to an additional 37 citations.

Table 1. Academic databases and search engines used and key words used.

Database	Keywords
PubMed	Depression prevention and control [MeSH] OR depression/therapy [MeSH] AND internet OR wireless technology [MeSH], AND intervention AND Internet AND depression, depression AND internet/cellular phone AND intervention, internet AND depression AND cognitive behavior therapy AND CCBT AND psychoeducation
CINAHL	depression AND intervention, depression intervention AND computer assisted/cellular phone, depression AND psychotherapy/ psychoeducation
PyschINFO	Internet-based intervention AND depressive disorder, depression intervention AND phone/text message/computer/internet-based, internet/phone/mobile phone/computer AND cognitive behavior therapy AND depression
Health Reference Center	depression intervention AND computer/internet/technology/mobile/mobile application
Evidence-Based Medicine Reviews	computer-assisted depressive, disorder intervention, internet depressive, disorder intervention, and major depressive disorder intervention
BioMed Central, MEDLINE, EBSCO	Internet-based intervention AND depressive disorder, depression intervention AND cell phone computer-assisted depressive, disorder intervention, internet depressive, disorder intervention, and major depressive disorder, intervention, depression, depressive symptoms, depressive disorders, e-mental health, new technologies, CCBT, self-help, intervention, treatment and prevention, empirical study, and peer reviewed journal.
Cordis/NHS/ Google/Yahoo	depression, depressive symptoms, depressive disorders, e-mental health, new technologies, CCBT, self-help, intervention, treatment and prevention, empirical study, and peer reviewed journal.

Inclusion criteria were that: (1) studies that exclusively included technologies applied to depression; (2) or with the flexibility of studies that combined both depression and anxiety; (3) studies with technologies that solely or were jointly/complimentary used with other methods of treatment and (4) studies that were psychoeducational. The exclusion criteria were; programs that did not meet the above criteria - (1) that were primarily usability studies, proposals, protocols or feasibility studies; (2) focused predominantly on professional and clinician-training; (3) delivered technology-based interventions were less than 50%. Unavailable abstracts were counted. After analysing titles, abstracts and the reference lists for related articles, 317 articles were retrieved. After careful reading of these articles, an additional 70 articles were excluded because the study or technology was to a lesser extent intended for depression. The remaining 58 programs and articles were included and were evaluated.

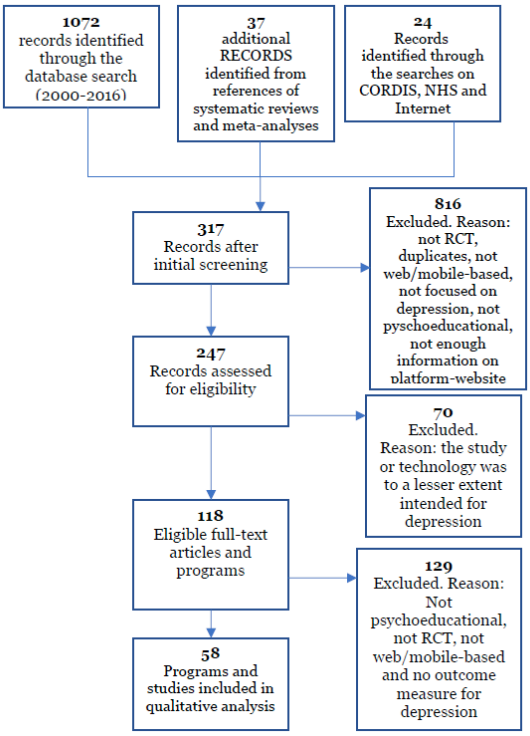


Fig. 1. Selection process flow diagram

3 Results

3.1 Summary

The objective was to conduct a literature and a scoping review on new technologies applied to the treatments for depression. A total of 41 studies in 47 published papers and scoping search identified collectively 58 different interventions programs, both EBP and NBP that met inclusion criteria. However, most of the studies and programs (46/58, 79.3%) employed interventions targeting symptoms both depression and anxiety. Previous knowledge of Help4Mood to the authors; thus, one of the program satisfying the inclusion criteria not identified through search.

Generally, interventions were offered in individual format with one exception of a group format program. 1 or more positive outcomes were identified in approximately

half (31/58, 53.4%) in comparison to control at post-intervention. However, no significant effect in at least 31% (18/58). CBT was the method most deployed, being employed in 26/41 (63%) of the studies; and in half of the 58 identified interventions (32/58, 55%). Although alternative content was identified, online delivery was preferred the majority programs with 4 delivered by CD-ROM, for example Beating the Blues (later transferred to online format). Whereas the inclusion criteria focused mostly on programs in English due to language restrictions, other foreign language programs that have been reviewed and published in English before; 3 programs - Internetpsykiatri, Kleur Je Leven, and Interapy were included. The review included both evidence-based or none (EBP vs NBP) programs whilst it was also identified that certain programs like This Way Up and e-Centre Clinic (Wellbeing Course, Uni-Wellbeing Course, Mood Mechanic Course, etc..), among others offered more depression treatments than one.

3.2 Programs and their content

In total, we identified 58 different interventions for treating depression with Beating the Blues (BTB) [18] being the most studied (with 3 RCTs and 10 open trials) followed by MoodGym. BtB composed of 8 CBT sessions and a series of filmed case studies modelling individual depressive symptoms [19]. MoodGym integrated workbook and graphic characters, interactive exercises, and relaxation audios, among others [20, 21]. Originally developed for an over 50-years target group [22], the Colour your Life program comprised of sessions on psychoeducation, cognitive restructuring, behaviour change, and relapse prevention; and has overtime been adapted to an adult population (18–65) [23, 24]. Deprexis, a CBT -based 10-module program that tailored content to the users' responses to given options, such as childhood experiences, and positive psychology [25]. Overcoming Depression on the Internet [26; ODIN] is a program applied to young adults (18-24 years) that includes modules on cognitive restructuring skills and behavioural activation [27]. However, other interventions too deviated from the standard CBT content. For example; Overcoming Depression [28] offered CBT concepts in a multimodal format including problem-solving therapy (PST) [29] a structured writing intervention (SWI) [30], a combination of face-to-face and cognitive therapy [31], or and format online [32]. In addition to CBT approaches, Recovery Road also integrated a clinician side for the management of client cases [33]. Most interventions were targeted at specific population, for example, those with remitted depression [34], and diabetes patients with comorbid depression [35]. Treatments varied in features reviewed. For example; technologies employed, mode of content delivery, language of interaction, country of origin, registration and referrals requirements, therapeutic approach, level and degree of human contact, target audience, length of intervention, additional tools and clinical impact. The majority of the programs used CBT content and delivered by web-based platforms, multimedia, and interaction; with a few exceptions [30, 25, 29, 24] - text format [36, 31], CD-ROM technology [37, 28], and online chat-based technology [38, 32].

3.2.1 Origin

Numerous studies were conducted in the United States [39-41,42,43] and Australia [44, 45], with a few other studies scattered in other countries with EBP studies from Netherlands, Germany, USA, Australia and the United Kingdom; whilst NBP originated from international collaboration projects, Australia, Sweden, New Zealand, Canada the United Kingdom, and United States.

3.2.2 Language

Reflective of the scope of countries of origin, a majority of the programs were offered in English. 50/58 (86%) exclusively in English whereas 2 programs; Kleur je Leven and Interapy were exclusively in Dutch. 4 EBP delivered in English but with alternative languages options - eCentre Clinic's Wellbeing Course in Arabic; MoodGYM in Norwegian Dutch, and Chinese; MoodHelper: Spanish and Deprexis offered in Swedish and German.

3.2.3 Technologies applied and features

CBT interventions were offered through websites to compliment therapist support or via email. Monitoring was predominantly via email. For interventions that involved exposure, relaxation and stress inoculation training used audio, and a combination of both audio and video, virtual reality and mobile. Many technologies were applied; audio (n=11), virtual reality (n=6), video (n=4), stand-alone computer programs (n=1), internet (n=21), and/or a combination of these; internet plus computer program, n=5; computer plus audio, n=1; internet plus audio, n=1; and audio plus video, n=5 among others.

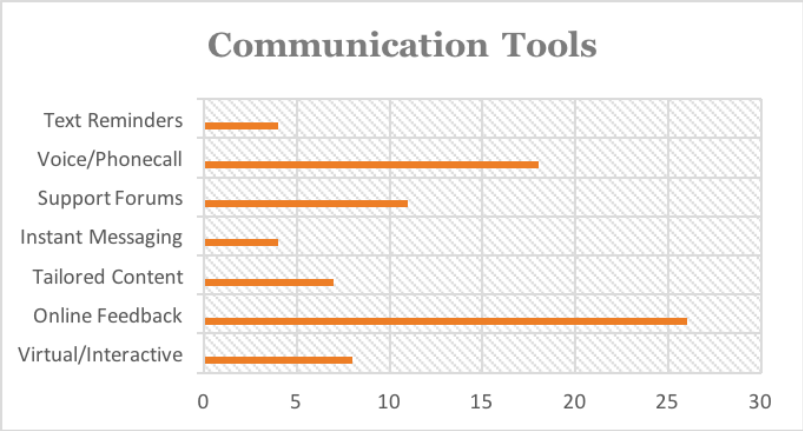


Fig. 2. Main communication tools in intervention

3.2.4 Mode of delivery

Only two studies used a native mobile application to deliver psychoeducation for depression [46, 47]. In both EBP and NBP, the treatment interventions varied in mode of delivery. In NBP offered 14 in multimodal format and 6 in text-only format. In EBP, 9 offered a multimodal-format that integrated a combination of audio, video, and text; and 1 in text-only format; mode of delivery unevaluated in other EBP. Regardless of whether EBP or NBP, its noteworthy that Watts et al. [46] identified no significant differences in clinical outcomes between mode of delivery within the same treatment program. However, Proudfoot et al. identified that a combination of both web-based platforms and in delivery of CBT permitted significant influence on the reduction of depression symptoms with an up to 87% program satisfaction [47]. These findings suggest that delivery through technology platforms is more favourable than standard psychotherapies.

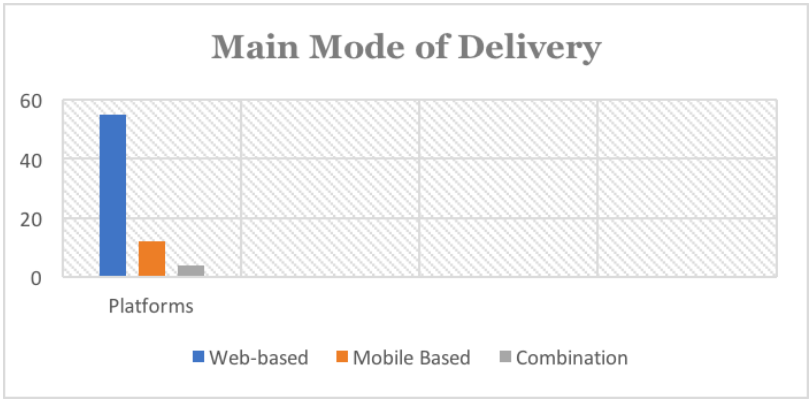


Fig. 3. Main modes of delivery in intervention

3.2.5 Therapy and intervention approach

Whereas a variety of therapeutic techniques were administered, a majority delivered CBT-focused treatments; 5 offered integrated therapies including CBT, IPT (Intrapersonal Therapy), physical activity, psychoeducation, and relaxation therapy; 11 NBP offered CBT-based content, 8 offered a of combination therapy models, and 4 NBP were either unknown or never defined therapeutic approach. Some programs required registration of a personal account while others did not. Whereas some programs required no registration in respect to anonymity, other for example; Deprexis, required a professional referral to be approved for registration.

3.2.6 Target audience

2 NBP targeted specific audiences - After Deployment for discharged veterans and e-Centre Clinic Wellbeing Plus Course of older adults. Adults were the target audience in the majority 32/58 (55%) of programs; 5/58 for adolescents, young adults or students, 6/58 for combined adolescents and adults, 3/58 for specific special population and 4/58 for unknown audiences. Programs targeting students and young adults focused on academic life (the stressors of an academic environment) and changes during adolescence.

3.2.7 Length of intervention

The length and degree of intervention ranged between 15 minutes to 12 weeks depending each program's objectives. In programs where post-intervention was offered, the duration of follow-ups was between immediate post-intervention up to 12 months post-intervention.

3.2.8 Level and degree of human contact

Mostly, self-administered interventions were identified (42/58, 72%), and other 5/58 (8.6%) were found to involve minimal contact with therapists; and 12/58 (20%) were predominantly self-help with email monitoring and discussion forums as the only forms of contact. 15/58 (25.8%) interventions can be categorised as therapist administered that often involved face-to-face interventions and/or integrated a technology-based component [48, 49].

3.2.9 Level of clinical intervention and relevance

Studies with greater levels of clinician involvement were those of greater compliance with adherence measures. 7 EBP and 11 NBP offered clinician support whereas 4 and 11 programs didn't, respectively. Correspondingly, greater communication and interaction tools favoured clinical engagement and thus showed greater compliance to treatment. Previous studies and meta-analyses support the conclusion that greater effect was always achieved from human-supported web-based psychoeducational interventions as opposed to unsupported interventions. Thus web-based delivery is one way to provider wider accessibility whilst relieving professional of previous job-related burdens [50, 51].

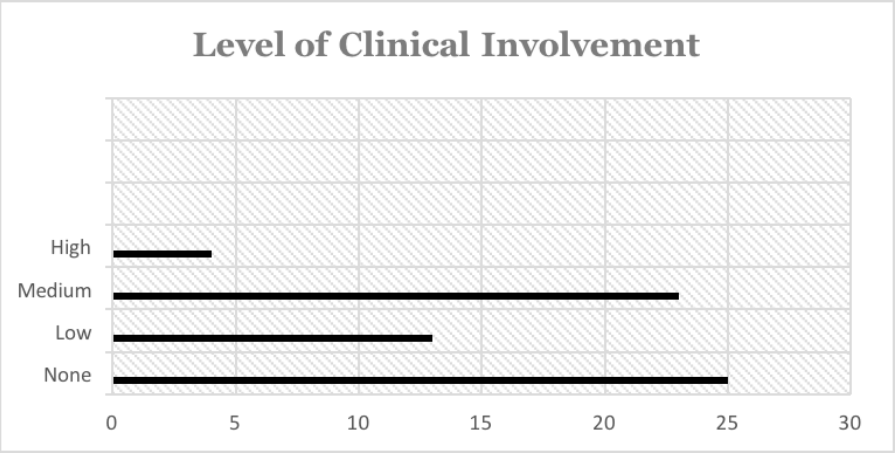


Fig. 4. Degree of ICT-tools in clinical involvement

4 Discussion

The depression treatment programs are evidence of the wide spread usage of web-based mental health care interventions for depression. However, despite the growing body of evidence in their favour over conventional psychotherapy, some people still perceive that web-based interventions comparable with conventional psychotherapy in their potential to enhance wellbeing outcomes. Thus, potential barriers exit, ranging from; negative attitudes of clinician and patient [52,53], regulations (both ethical and legal) [54], reluctance by clinicians to recommend such interventions and clinician job-insecurity [55]. We acknowledge the possibility that some interventions reviewed may have been applied in contexts of use for which they were not designed as for example some findings identified that CCBT may be more effective in averting anxiety than depression in contexts that involved both depression and anxiety. Therefore, it is critical to focus the target audience, and context than to generalise the mental disorder to be treated. In instances where a population group (targeted audiences) was identified, the intervention content remained generic in nature with little to no evidence that individualised treatment plans being offered.

While web-based programs interventions offer greater accessibility and evidence supporting efficacy, varying degrees of these aspects exist. For example, the availability of some programs in only one language suggests the need for alternative languages. Other programs had accessibility restrictions that can perhaps be attributed to national policies, country of residence and availability of professional support. Certain programs like MoodGYM indicated positive results effectiveness partly because they included additional integrated support forums, and crisis links to assist patients [56]. Additional therapist support was identified as effective too [56, 57], further

research should attempt to understand the coalition between therapeutic approach, and patient experience; and how the two impact the outcomes of treatment.

The efficacy and validity of most programs (31/58, 53%) isn't evaluated using RCTs. Thus, EBP used guided approaches, and were subjected to validated assessment tools, integrated safeguards and additional features and provide user statistics as opposed to NBP. Though some NBP programs like MoodGYM, and This Way Up were freely available, easily accessible, and with supporting evidence for their efficacy, most required a higher fee than EBP. The web-based self-help programme, could be an effective stand-in option for patients on a waiting list especially with an overloaded current healthcare system globally. The ICT-based interventions are a great alternative and have demonstrated positive results in 6-month follow-up [58]. Such approaches (1) accessibility, (2) cost-effectiveness, (3) anonymity and privacy, and (4) self-guided help and management by the user. However, it is noteworthy that several challenges may come with web-based approaches. For example; ethical concerns, digital privacy, among others. Further investigation could study how to mirror such face-to -face interactions of conventional psychotherapy into ICT-based interventions to traditional therapeutic techniques while answering the challenges that come with the technological transformation in e-mental health.

5 Limitations

The review searched five (6) databases but there is a possibility that relevant journals not indexed by these databases could have been missed. However, we also searched the Beacon portal [59] and previous reviews in references to remedy this problem**.** Additional programs listed in Beacon may not be up-to-date because Beacon has not been updated since 2016. There was a language restriction set for only English studies and programs with an exception of 2 programs that were studied in English though not originally in English language. The Spain version of Google and Yahoo was deployed; however, there is a possibility that due to geolocation and targeting, searches conducted elsewhere (using the above-mentioned search engines) may offer different results due to geolocation ad local advertising.

6 Conclusions

The benefits of ICT-based treatment in depression in addition to reduced depression symptoms poses substantial opportunities to generate clinically significant improvements and service in depression treatment. Further research should validate the efficiency of ICT-based depression treatment because a lot of doubts and insufficient scientific evidence remains regarding the effectiveness in comparison to conventional therapy. Though some technologies have been co-developed with inclusion and involved the input of both clinicians and patients to integrate mental healthcare evi-

dence-based principles, further research is inevitable to illustrate the such principles in routine clinical settings. The clear potential of at the intersection of health and digital technology in depression is yet to be realized. This coupled with publications of results would help expand the evidence base for the relevance and efficacy of ICT-based tools on treating depression and to inform the design of better programs and technologies in the future. There are numerous key challenges; patient-centric approaches in technology implementation, expanding the evidence base of clinical effectiveness and guidelines regarding safety, usability, regulations etc.

References

1. The ESEMeD/MHEDEA 2000 Investigators, Alonso, J., et. al.: The European Study of the Epidemiology of Mental Disorders (ESEMeD/MHEDEA 2000) project: rationale and methods. Int J Methods Psychiatry Res. (2002); 11(2):55–67
2. The ESEMeD/MHEDEA 2000 investigators, Alonso, J., et. al.: Sampling and methods of the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatrica Scandinavica (2004) 109: 8–20. doi:10.1111/j.1600-0047.2004.00326
3. NICE: Depression. The treatment and management of depression in adults, NICE clinical guideline 90. NICE clinical guideline 23. London: National Institute for Health and Clinical Excellence (2009)
4. Yamamoto K., Shiota S., et.al.: A diagnosis of depression should be considered in patients with multiple physical symptoms in primary care clinics. J Exp Med. (2013) 279-85
5. Cavanagh, K., Shapiro, D. A.: Computer treatment for common mental health problems. Journal of Clinical Psychology (2004) 60, 239–251
6. National Institute for Health and Clinical Excellence (NICE): Computerised Cognitive Behaviour Therapy for Depression and Anxiety: review of technology appraisal 51. NICE Technology Appraisal 97. London: NICE (2006)
7. Richards D., Richardson T.: Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev. (2012) 32(4):329–342. doi: 10.1016/j.cpr.2012.02.004.
8. NHS Confederation: E-Mental Health: What's all the Fuss About? (2014) (accessed 2017-06-25)http://www.nhsconfed.org/~/media/Confederation/Files/Publications/Documents/E-mental-health.pdf
9. Cuijpers P., Donker T., van SA., Li J., Andersson G.: Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. Psychol Med. (2010) 1943–57
10. Prociow PA., Crowe JA.: Towards personalised ambient monitoring of mental health via mobile technologies. Technol Health Care (2010) 275–84
11. Andrews G.: Implications for intervention and prevention from the New Zealand and Australian mental health surveys. Aust NZJ Psychiatry (2006) 827–9
12. Marks IM., Cavanagh K., Gega L.: Hands On Help: Computer-Aided Psychotherapy. New York, Psychology Press, (2007)
13. Kenwright M., Liness S., Marks I.: Reducing demands on clinicians by offering computer-aided self-help for phobia/panic: feasibility study. British Journal of Psychiatry (2001) 179: 456–459
14. McGorry PD., Yung AR., Pantelis C., et al: A clinical trials agenda for testing interventions in earlier stages of psychotic disorders. Medical Journal of Australia 190 (suppl); (2009) S33–S36

15. McCrone P., Knapp M., Proudfoot J., et al: Cost-effectiveness of computerised cognitive behavioural therapy for anxiety and depression in primary care: randomised controlled trial. British Journal of Psychiatry (2004)185:55– 62
16. Kilbourne AM.: E-health and the transformation of mental health care. Psychiatric Services (2012) 63:1059
17. Ainsworth M.: My life as an e-patient; in ETherapy: Case Studies, Guiding Principles, and the Clinical Potential of the Internet. Edited by Hsiung RC. New York, Norton (2002)
18. Proudfoot, J., Ryden, C., Everitt, B., Goldberg, D., Tylee, A., Gray, J. A., et al.: Clinical efficacy of computerised cognitive-behavioural therapy for anxiety and de-pression in primary care: Randomised controlled trial. The British Journal of Psychiatry, (2004) 185(1), 46–54. doi:10.1192/bjp. 185.1.46.
19. Cavanagh, K., Shapiro, D.A., Van Den Berg, S., Swain, S., Barkham, M., Proudfoot,J.: The effectiveness of computerized cognitive behavioural therapy in routine care. British Journal of Clinical Psychology(2006) 45(4), 499–514. doi:10.1348/014466505x84782.
20. Christensen, H., Griffiths, K. M., & Korten, A.: Web-based cognitive-behaviour therapy: Analysis of site usage and changes in depression and anxiety scores. Journal of Medical Internet Research (2002) 1–8. doi:10.2196/jmir.4.1.e3.
21. Perini, S., Titov, N., & Andrews, G.: The Climate Sadness program: An open trial of Internet-based treatment for depression. E-journal of Applied Psychology (2008) 18–24
22. Spek, V., Nyklicek, I., Smits, N., Cuijpers, P., Riper, H., Keyzer, J., et al.: Internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years old: A randomized controlled clinical trial. Psychological Medicine (2007) 37,1797–1806
23. de Graaf, L. E., Gerhards, S. A. H., Arntz, A., Riper, H., Metsemakers, J. F. M., Evers, S. M. A.A., et al.: Clinical effectiveness of online computerised cognitive-behavioural therapy without support for depression in primary care: Randomised trial. The British Journal of Psychiatry (2009) 73–80
24. Warmerdam, L., van Straten, A., Twisk, J., Riper, H., & Cuijpers, P.: Internet-based treatment for adults with depressive symptoms: Randomized controlled trial. Journal of Medical Internet Research (2008) 10(4), e44. doi:10.2196/jmir.1094.
25. Meyer, B., Berger, T., Caspar, F., Beevers, G. C., Andersson, G., & Weiss, M.: Effectiveness of a novel integrative online treatment for depression (Deprexis): Randomized controlled trial. Journal of Medical Internet Research (2009)11(2),e15. doi:10.2196/jmir.1151.
26. Clarke, G., Reid, E., Eubanks, D., O'Connor, E., DeBar, L. L., Kelleher, C., et al.: Overcoming depression on the internet (ODIN): A randomized controlled trial of an internet depression skills intervention program. Journal of Medical Internet Research (2002) 4(3), e14. doi:10.2196/jmir.4.3.e14
27. Clarke, G., Kelleher, C., Hornbrook, M., DeBar, L., Dickerson, J., & Gullion, C.: Randomized effectiveness trial of an internet, pure self-help, cognitive behavioral intervention for depressive symptoms in young adults. Cognitive Behaviour Therapy(2009) 38(4), 222–234. doi:10.1080/16506070802675353
28. Whitfield, G., Hinshelwood, R., Pashely, A., Campsie, L., & Williams, C.: The Impact of a novel computerized CBT CD-ROM (Overcoming Depression) offered to patients referred to clinical psychology. Behavioural and Cognitive Psychotherapy (2006) 1–11
29. van Straten, A., Cuijpers, P., & Smits, N.: Effectiveness of a web-based self-help intervention for symptoms of depression, anxiety, and stress: Randomized con-trolled trial. Journal of Medical Internet Research (2008) 1–10
30. Kraaij, V., Van Emmerik, A., Garnefski, N., Schroevers, M. J., Lo-Fo-Wong, d., Van Empelen,P., et al.: Effects of a cognitive behavioral self-help program and a computerized structured writing intervention on depressed mood for HIV-infected people: A pilot randomized controlled trial. Patient Education and Counseling (2010) 80, 200–204

31. Wright, J. H., Wright, A. S., Albano, A., Basco, M. R., Goldsmith, L. J., Raffield, T., et al.: Computer-assisted cognitive therapy for depression: Maintaining efficacy while reducing therapist time. The American Journal of Psychiatry (2005) 1158–1164
32. Thompson, N. J., Walker, E. R., Obolensky, N., Winning, A., Barmon, C., DiIorio, C., et al.: Distance delivery of mindfulness-based cognitive therapy for depression: Project UPLIFT. Epilepsy & Behavior (2010) 247–254
33. Robertson, L., Smith, M., Castle, D., & Tannenbaum, D.: Using the internet to enhance the treatment of depression. Australasian Psychiatry (2006) 413–417
34. Holländare, F., Johnsson, S., Randestad, M., Tillfors, M., Carlbring, P., Andersson, G., et al.: Randomized trial of internet-based relapse prevention for partially remitted depression. Acta Psychiatrica Scandinavica, Early View.(2011) doi:10.1111/j.1600-0447.2011.01698.x.
35. van Bastelaar, K. M. P., Pouwer, F., Cuijpers, P., Riper, H., & Snoek, F. J.: Web-based depression treatment for Type 1 and Type 2 diabetic patients. DiabetesCare (2011) 320–325
36. Andersson, G., & Cuijpers, P.: Internet-based and other computerized psychological treatments for adult depression: A meta-analysis. Cognitive Behaviour Therapy (2009) 196–205
37. Purves, D. G., Bennett, M., & Wellman, N.: An open trial in the NHS of Blues Be-gone®: A new home-based computerized CBT program. Behavioural and Cognitive Psychotherapy (2009) 541–551
38. Kessler, D., Lewis, G., Kaur, S., Wiles, N., King, M., Weich, S., et al.: Therapist-delivered internet psychotherapy for depression in primary care: A randomised con-trolled trial. The Lancet (2009) 628–634
39. Braithwaite SR., Fincham FD., ePREP: computer based prevention of relationship dysfunction, depression and anxiety. Journal of Social and Clinical Psychology (2007) 609-622
40. Braithwaite SR., Fincham FD.: A randomized clinical trial of a computer based preventive intervention: replication and extension of ePREP. J Fam Psychol (2009) 32-38
41. Cukrowicz KC., Joiner TE.: Computer-based intervention for anxious and depressive symptoms in a non-clinical population. Cogn Ther Res. (2007) 677-693
42. Mailey EL., Wójcicki TR., Motl RW., Hu L., Strauser DR., Collins KD., et al.: Internet-delivered physical activity intervention for college students with mental health disorders: a randomized pilot trial. Psychol Health Med (2010) 646-659
43. Seligman ME., Schulman P., Tryon AM.: Group prevention of depression and anxiety symptoms. Behav Res Ther (2007) 1111-1126
44. Sethi S., Campbell AJ., Ellis LA.: The use of computerized self-help packages to treat adolescent depression and anxiety. Journal of Technology in Human Services (2010) 144-160
45. Ellis LA., Campbell AJ., Sethi S., O'Dea BM.: Comparative randomized trial of an online cognitive-behavioral therapy program and an online support group for depression and anxiety. J Cyber Ther Rehabil (2011) 461-467
46. Watts, S., Mackenzie, A., Thomas, C., Griskaitis, A., Mewton, L., Williams, A., et al.: CBT for depression: a pilot RCT comparing mobile phone vs. computer, BMC Psychiatry 13 (2013) 49
47. Proudfoot, J., Clarke, J., Birch, M.R., Whitton, E.A., Parker, G., Manicavasagar, V., et al.: Impact of a mobile phone and web program on symptom and functional outcomes for people with mild-to-moderate depression, anxiety and stress: a randomised controlled trial, BMC Psychiatry 13 (2013)
48. Mailey EL., Wójcicki TR., Motl RW., Hu L., Strauser DR., Collins KD., et al.: Internet-delivered physical activity intervention for college students with mental health disorders: a randomized pilot trial. Psychol Health Med (2010) 646-659.

49. Sethi S., Campbell AJ., Ellis LA.: The use of computerized self-help packages to treat adolescent depression and anxiety. Journal of Technology in Human Services (2010) 144-160
50. Cowpertwait, L., Clarke D.: Effectiveness of web-based psychological interventions for depression: a meta-analysis, Int. J. Ment. Health Addict. 11 (2013) 247–268
51. Barak, A., Klein, B., Proudfoot, J.: Defining internet-supported therapeutic interventions, Ann. Behav. Med. 38 (2009) 4–17
52. Stallard P., Richardson T., Velleman S.: Clinicians' attitudes towards the use of computerized cognitive behaviour therapy (cCBT) with children and adolescents. Behav Cogn Psychother (2010) 545-560
53. Mohr DC., Siddique J., Ho J., Duffecy J., Jin L., Fokuo JK.: Interest in behavioral and psychological treatments delivered face-to-face, by telephone, and by internet. Ann Behav Med (2010) 89-98
54. Dever Fitzgerald T., Hunter PV., Hadjistavropoulos T., Koocher GP.: Ethical and legal considerations for internet-based psychotherapy. Cogn Behav Ther (2010) 173-187.
55. Andersson G., Titov N.: Advantages and limitations of Internet-based interventions for common mental disorders. World Psychiatry (2014) 4-11
56. Høifødt RS., Lillevoll KR., Griffiths KM., Wilsgaard T., Eisemann M., Waterloo K., et al.: The clinical effectiveness of web-based cognitive behavioral therapy with face-to-face therapist support for depressed primary care patients: randomized controlled trial. J Med Internet Res (2013);15(8):e153
57. Donker T., Bennett K., Bennett A., Mackinnon A., van Straten A., Cuijpers P., et al.: Internet-delivered interpersonal psychotherapy versus internet-delivered cognitive behavioral therapy for adults with depressive symptoms: randomized controlled noninferiority trial. J Med Internet Res (2013);15(5):e82
58. Andersson, G., Bergstorm, J., Holländare, F., Varlbring, P., Kaldo, V., & Ekselius, L.: Internet-based self-help for depression: Randomised controlled trial. The British Journal of Psychiatry (2005) 187, 456–461
59. Hopewell S., Clarke M., Lefebvre C., Scherer R.: Handsearching versus electronic searching to identify reports of randomized trials. Cochrane Database Syst Rev 2007(2):MR000001

The ITACA-WIICT'17 is a meeting forum for scientifics, technicians and other professionals who are dedicated to Information and communication technologies study and research. Its fundamental scope is to promote the contact among scientific and professionals, improving the cooperation as well as the technological transfer among professionals.

Institute ITACA
Information and Communication Technologies

978-84-697-7327-7