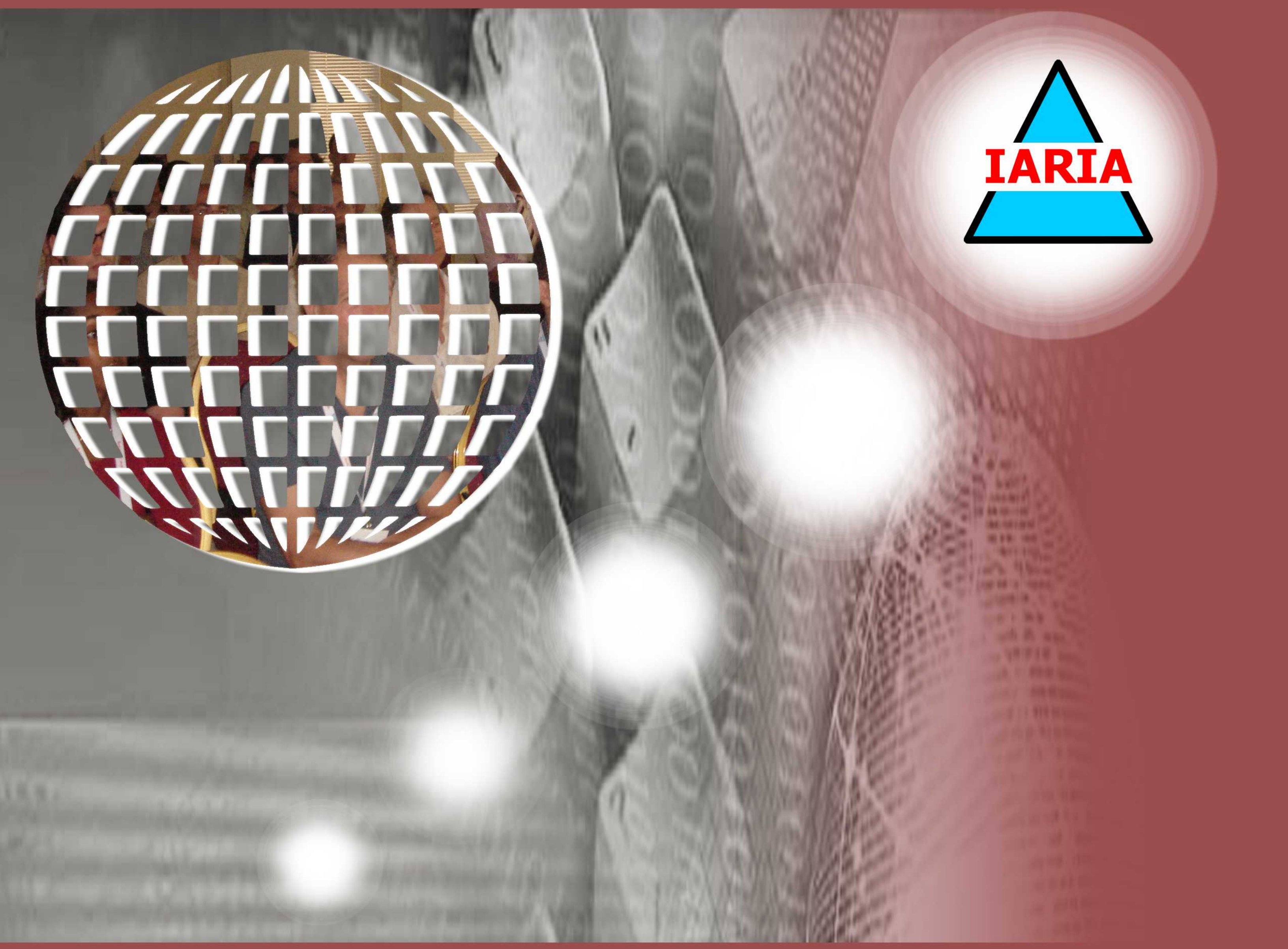
International Journal on

Advances in Telecommunications



2014 vol. 7 nr. 3&4

The *International Journal on Advances in Telecommunications* is published by IARIA.
ISSN: 1942-2601
journals site: http://www.iariajournals.org
contact: petre@iaria.org

Sponsored by IARIA
www.iaria.org

International Journal on Advances in Telecommunications
Volume 7, Number 3 & 4, 2014

Fernando Boronat Seguí, Universidad Politecnica de Valencia, Spain
Christos Bouras, University of Patras, Greece
Martin Brandl, Danube University Krems, Austria
Julien Broisin, IRIT, France
Dumitru Burdescu, University of Craiova, Romania
Andi Buzo, University "Politehnica" of Bucharest (UPB), Romania
Shkelzen Cakaj, Telecom of Kosovo / Prishtina University, Kosovo
Enzo Alberto Candreva, DEIS-University of Bologna, Italy
Rodrigo Capobianco Guido, São Paulo State University, Brazil
Hakima Chaouchi, Telecom SudParis, France
Silviu Ciochina, Universitatea Politehnica din Bucuresti, Romania
José Coimbra, Universidade do Algarve, Portugal
Hugo Coll Ferri, Polytechnic University of Valencia, Spain
Noel Crespi, Institut TELECOM SudParis-Evry, France
Leonardo Dagui de Oliveira, Escola Politécnica da Universidade de São Paulo, Brazil
Gerard Damm, Alcatel-Lucent, USA
Francescantonio Della Rosa, Tampere University of Technology, Finland
Chérif Diallo, Consultant Sécurité des Systèmes d'Information, France
Klaus Drechsler, Fraunhofer Institute for Computer Graphics Research IGD, Germany
Jawad Drissi, Cameron University , USA
António Manuel Duarte Nogueira, University of Aveiro / Institute of Telecommunications, Portugal
Alban Duverdier, CNES (French Space Agency) Paris, France
Nicholas Evans, EURECOM, France
Fabrizio Falchi, ISTI - CNR, Italy
Mário F. S. Ferreira, University of Aveiro, Portugal
Bruno Filipe Marques, Polytechnic Institute of Viseu, Portugal
Robert Forster, Edgemount Solutions, USA
John-Austen Francisco, Rutgers, the State University of New Jersey, USA
Kaori Fujinami, Tokyo University of Agriculture and Technology, Japan
Shauneen Furlong , University of Ottawa, Canada / Liverpool John Moores University, UK
Ana-Belén García-Hernando, Universidad Politécnica de Madrid, Spain
Bezalel Gavish, Southern Methodist University, USA
Christos K. Georgiadis, University of Macedonia, Greece
Mariusz Glabowski, Poznan University of Technology, Poland
Katie Goeman, Hogeschool-Universiteit Brussel, Belgium
Hock Guan Goh, Universiti Tunku Abdul Rahman, Malaysia
Pedro Gonçalves, ESTGA - Universidade de Aveiro, Portugal
Valerie Gouet-Brunet, Conservatoire National des Arts et Métiers (CNAM), Paris
Christos Grecos, University of West of Scotland, UK
Stefanos Gritzalis, University of the Aegean, Greece
William I. Grosky, University of Michigan-Dearborn, USA
Vic Grout, Glyndwr University, UK
Xiang Gui, Massey University, New Zealand
Huaqun Guo, Institute for Infocomm Research, A*STAR, Singapore
Song Guo, University of Aizu, Japan
Ching-Hsien (Robert) Hsu, Chung Hua University, Taiwan

Peter Orosz, University of Debrecen, Hungary
Patrik Österberg, Mid Sweden University, Sweden
Harald Øverby, ITEM/NTNU, Norway
Tudor Palade, Technical University of Cluj-Napoca, Romania
Constantin Paleologu, University Politehnica of Bucharest, Romania
Stelios Papaharalabos, National Observatory of Athens, Greece
Gerard Parr, University of Ulster Coleraine, UK
Ling Pei, Finnish Geodetic Institute, Finland
Jun Peng, University of Texas - Pan American, USA
Cathryn Peoples, University of Ulster, UK
Dionysia Petraki, National Technical University of Athens, Greece
Dennis Pfisterer, University of Luebeck, Germany
Timothy Pham, Jet Propulsion Laboratory, California Institute of Technology, USA
Roger Pierre Fabris Hoefel, Federal University of Rio Grande do Sul (UFRGS), Brazil
Przemyslaw Pochec, University of New Brunswick, Canada
Anastasios Politis, Technological & Educational Institute of Serres, Greece
Adrian Popescu, Blekinge Institute of Technology, Sweden
Neeli R. Prasad, Aalborg University, Denmark
Dušan Radović, TES Electronic Solutions, Stuttgart, Germany
Victor Ramos, UAM Iztapalapa, Mexico
Gianluca Reali, Università degli Studi di Perugia, Italy
Eric Renault, Telecom SudParis, France
Leon Reznik, Rochester Institute of Technology, USA
Joel Rodrigues, Instituto de Telecomunicações / University of Beira Interior, Portugal
David Sánchez Rodríguez, University of Las Palmas de Gran Canaria (ULPGC), Spain
Panagiotis Sarigiannidis, University of Western Macedonia, Greece
Michael Sauer, Corning Incorporated, USA
Marialisa Scatà, University of Catania, Italy
Zary Segall, Chair Professor, Royal Institute of Technology, Sweden
Sergei Semenov, Broadcom, Finland
Sandra Sendra Compte, Polytechnic University of Valencia, Spain
Dimitrios Serpanos, University of Patras and ISI/RC Athena, Greece
Adão Silva, University of Aveiro / Institute of Telecommunications, Portugal
Pushpendra Bahadur Singh, MindTree Ltd, India
Mariusz Skrocki, Orange Labs Poland / Telekomunikacja Polska S.A., Poland
Leonel Sousa, INESC-ID/IST, TU-Lisbon, Portugal
Liana Stanescu, University of Craiova, Romania
Cosmin Stoica Spahiu, University of Craiova, Romania
Young-Joo Suh, POSTECH (Pohang University of Science and Technology), Korea
Hailong Sun, Beihang University, China
Jani Suomalainen, VTT Technical Research Centre of Finland, Finland
Fatma Tansu, Eastern Mediterranean University, Cyprus
Ioan Toma, STI Innsbruck/University Innsbruck, Austria
Božo Tomas, HT Mostar, Bosnia and Herzegovina
Piotr Tyczka, Poznan University of Technology, Poland
John Vardakas, University of Patras, Greece

International Journal on Advances in Telecommunications
Volume 7, Numbers 3 & 4, 2014

CONTENTS

Transmitter Based Look-Up Tables for Optical Wireless IR-UWB Systems

Mohammed Al-Olofi, Andreas Waadt, Guido H. Bruck, and Peter Jung
Department of Communication Technologies
University of Duisburg-Essen
Duisburg, Germany
Email: info@kommunikationstechnik.org

***Abstract*—This paper investigates a design for impulse radio ultra-wideband communication over optical wireless link using optical front-ends. The migration between IR-UWB and optical wireless communications introduces a solution for radio spectrum congestion and limited IR-UWB transmission power governed by spectrum mask. Nevertheless, the nonlinear effects of light emitting diodes (LED) limit the power dynamic range that modulates signal pulses. In this work, a transmitter based digital look-up table (LUT) is proposed. This transmitter allows of storing pulse samples values in form of binary codewords. The LUT's binary output switches LEDs groups and hence controls transmission power. This parallel switching enables power-efficient optical representation of pulses and avoids the nonlinear effects caused by LED operation characteristics. The utilization of LUT excludes the necessity of digital to analog converter (DAC) or transconductance amplifier (TCA) employed in common optical transmitters to perform the electrical-optical conversion. The selection of codeword length and LUT size are done based the BER performance generated by computer simulations. Moreover, the channel effects in presence of line of sight (LOS) and diffuse links are simulated and analyzed. The system simulations are described for a mobile user in indoor office environment. The BER performances for different mobility scenarios will be discussed.**

Keywords-Hybrid Optical/Radio systems, IR-UWB, LUT, OWC.

I. INTRODUCTION

The optical wireless communications (OWC) is an alternative technology to radio communications, which suffers from congested frequency bands as the number of mobile users increased significantly. The OWC offers a broad unlicensed free spectrum that enables high data rate, low cost, high speed, and ease of development systems. These advantages make the optical solution attractive for short range communications applications, such as smart homes, smart offices, wireless LANs, and sensor networks.

Currently, the ultra-wideband systems are the best technology choice for short range communication, since they offer a large bandwidth (3.1-10.6 GHz), high speed, immune to multipath fading, multi access capabilities, and low cost transceivers. The fractional bandwidth of UWB is defined by FCC as a signal with 20% of its center frequency or 500 MHz bandwidth, when the center frequency is above 6 GHz with a limited power levels regulated by authorities. The frequency and power regulations are different from region to region, which make the standardization of such technology a difficult task. Also, the different emission power levels increase the interference levels from UWB devices on other wireless devices operates in the same frequency band.

On other hand and a solution to these limitations, the optical link is a promising media for short range wireless communications since its offer an unlicensed free spectrum. The evaluations of using an optical media in IR-UWB systems will overcame the limitations in emission power and bandwidth and/or introduce a new radio/optical IR-UWB system that exploit the advantages of both configurations. In [1], we had proposed a design for optical IR-UWB system design to solve the problem of radio interference and enlarge the transmission power and frequency spectrum. However, the current optical sources introduce a low modulation bandwidth, which does not exceed 20 MHz in case of using LEDs and hence limit the data rate and available broadband spectrum. This limitation is caused by the characteristics of LEDs available in the market, which have a slow response to the feeding current. The 3dB frequency and hence the modulation bandwidth depends on minority carrier lifetime, which define the rise and fall time of LED. Nevertheless, a higher frequency modulation bandwidth LEDs in the range of 200–300 MHz was proposed in [2]. In this literature, a GaN-based blue LED with a higher optical output power of 1.6 mW at bias current of 35 mA and a high electrical-to-optical 3dB bandwidth of 225.4 MHz was reported. These LEDs have a smaller chip size and exhibited a lower optical output power. Commonly, a lot of solid state electronics researches focus on LED designed for OWC applications, which gives a sign that we will see high-speed low-cost LEDs in the next decade.

Another limitation factor in optical wireless transmitters based LED is the nonlinearity caused by thermal aspects. The LED has a threshold voltage, which is the minimum voltage to flow current across the LED and light emission. Below this voltage, the LED operate in cut-off region and above it, the current is increased exponentially as well as the light intensity toward saturation region. The LED current versus light power relation is nonlinear due to the thermal characteristic and the nonlinear transition into saturation region. This nonlinearity causes a drop in electrical to optical conversion and affects the modulated baseband signal represented by light intensity and hence, restricts the dynamic range and power transmission of optical transmitter. In conventional optical transmitter, a pre-distortion equalizer and post-equalization methods are used to decrease the signal distortion [3].

In RF communication system, a similar nonlinear effect was addressed. The power amplifier is a main source of nonlinearity when it operates near the saturation region causing a distortion in signal phase and amplitude [4]. Several linearization techniques are proposed to linearize the PA operation and increase the power efficiency such as, adaptive predistortion using LUT [5][6]. The LUT is an array of values, which are calculated based on linearization function needed to scale the amplitude and phase of the input signal and stored in memory. The creation of the LUT is done in a calibration mode. However, the updating could be done in real time to adjust to changing PA characteristics, such as a change in the temperature. For example, a cartesian LUT was proposed in [7] to eliminate the need for cartesian-to-polar conversion and modulator correction and reduces the power consumption of feedback path by allowing the use of low-speed low-SNR ADCs. In optical wireless systems, a look-up table coding method has been proposed to convert the information bits into the M-PPM codewords using a specially designed coding table. The utilization of M-PPM with forward correcting coding method increases system performance through constant power transmission. However, an exhaustive search using computer optimization method had to be done to find a subset of M-PPM codewords that satisfies the minimum hamming distance that defined according to system environment and the system purpose [8].

The OWC conventional transmitters employ DAC to transform the data bits into voltage and TCA to drive the optical source. The literature in [3] introduced a discrete power level stepping optical wireless OFDM transmitter, which consists of several LEDs groups that emit specific stepped optical intensities using switches driven by input data. This transmitter improves the system performance and reduces power consumption since DAC and TCA are no longer needs. This idea can be extended in our system to transmit an optical version of IR-UWB pulses.

In this paper, the employment of optical media to transmit a signal designed for IR-UWB radio system is investigated. To reach this goal, the electrical to optical conversion of pulse shapes using LUT is proposed. The proposed transmitter omits the DAC and quantization process needed for each pulse transmission as proposed in [1]. For a given modulation scheme, the signal pulse is stored in LUT and hence there is no need to repeat the same steps to convert the electrical Gaussian pulse to optical power. Also, the BER performance related to the size of LUT and quantization bit depth are presented. The IR-UWB structure of transmitter and receiver are discussed to elaborate system modification in seeks of optical link utilization. Moreover, the simulation of system is carried out to show the BER performance.

The reminder of the paper is organized as follows. In Section II, the system design describes the optical transmitter operation and quantization effects are presented. In Section III, the design of optical transmitter based LUT is introduced and in Section IV the system model including the optical channel impulse response regarding the environment and SNR are explained and analyzed. The system simulation is introduced and the results are discussed in Section V. Finally, Section VI concludes the paper.

II. SYSTEM DESIGN

Optical wireless systems consist of transmitter and receiver that employ optical parts mounted in front-ends as shown in Fig. 1. The transmitter modulate data information using one of single modulation schemes such as, (OOK, PAM, PPM, etc.) or multi modulation schemes such as, (OFDM, MPPM, M-PAM). These modulated symbols convert to analog signal that control the intensity of light emitted from an optical source such as LED or LD. The LD is preferred than LED, since it provides higher power, larger bandwidth, and linear electrical to optical characteristics at expense of eye safety limitation and higher cost. In the same manner, LEDs are low cost, higher lamination but limited bandwidth and suffer from nonlinearities of optical to electrical conversion characteristics. Table I summarizes the comparison of the two sources. Consequently, the receiver detects the optical light impinging on the surface and converts it to electrical current. Normally, the detector is p-intrinsic-n (PIN) photodiode or avalanche photodiode (APD). After the detector, low pass filter and amplifier are used to remove the higher frequency components and amplify the signal.

In the same manner, the IR-UWB transmitter sends the modulated symbol on a stream of short pulses shaped by pulse shape filter without carrier. One of modulation schemes (OOK, PPM, PAM, PSK, etc.) can be performed. After that, a Gaussian pulse or one of its derivatives is used to represent the modulated symbol. Also, multiple access methods such as, time hopping (TH) and direct sequence (DS) might be used along with modulation to increase the symbol power and range as well as enables multiple users. At the receiver, a correlation receiver or rake receiver employs to correlate signal shape with a local generated replica of pulse and hence obtain the detected bit using a decision rule.

To gain the benefits of these two systems, a modified IR-UWB over optical wireless links was proposed in [1]. This section illustrates the structure of system and introduces a solution for LEDs nonlinear effects.

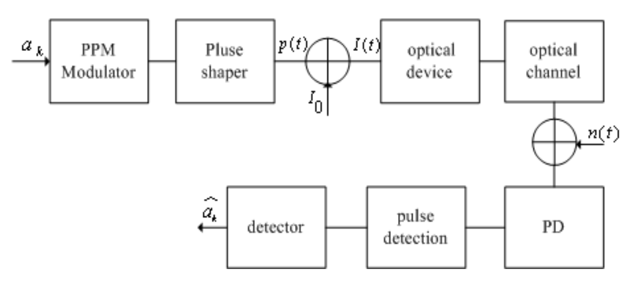


Figure 1. Optical IR-UWB system

TABLE I. COMPARISON OF LED AND LD

Characteristics	LED	LD
Output power	Low	High
Modulation bandwidth	10x-100x MHz	10x-100x GHz
Beam	Broad	Narrow
Coherence	Noncoherent	Coherent
Cost	Low	High
Harmonic distortion	High	Less

A. *Transmitter Design*

In a symbol time T_s, every data symbol is transmitted in form of N_f frames with frame duration T_f. In each frame, the information bits are modulated on short pulses $p(t)$ with pulse width T_p and $T_p \ll T_f$. The transmitter in Fig. 1 uses Pulse Position Modulation (PPM) with Gaussian shape pulse. The transmitted signal is represented as:

$$x(t) = p(t - kT_s - a_k \varepsilon) \tag{1}$$

In this modulation scheme, the transmitted pulse $p(t)$ is delayed for ε time shift, when the transmitted bit is one. In case of zero bits, the transmitter applies no delay to pulses. After signal modulation, the pulses enter pulse shaper that converts UWB pulses to be adequate for optical transmission. To transmit IR-UWB pulses over optical wireless link, the feed signal operates LED should be positive and above bias level. This means that any negative amplitude will bias light source reversely. Unfortunately, most of IR-UWB pulses are bipolar with negative and positive amplitudes. Nevertheless, the bipolar signal can be converted to unipolar by adding certain DC-bias that provides signal positivity as proposed in [1]. Although, large DC-bias increases signal power for a given BER while small DC-bias increases clipping noise causing signal distortion. Therefore, selection of DC-bias is tradeoff between signal power and clipping noise.

B. *Optical pulse shaping*

The signal pulse $p(t)$ is used to drive the LED by convert it to a set of quantized current levels as shown in Fig. 2. The pulse is quantized to L current level and every level is mapped to a defined brightness level. The current levels represent the intensity power should be non-negative to ensure that LED is not reversely biased.

$$I(t) = I_0 + p(t) \tag{2}$$

A dc-bias is chosen to boost the negative part of pulse in order to keep the LED in the 'ON' state and illuminance a 10% of full brightness. The constant forward current (I_0) will keep LED operating even if the pulse time end and hence the LED remains working in the active region. Consequently, the modulation bandwidth should be increased since the LED switching times is minimized.

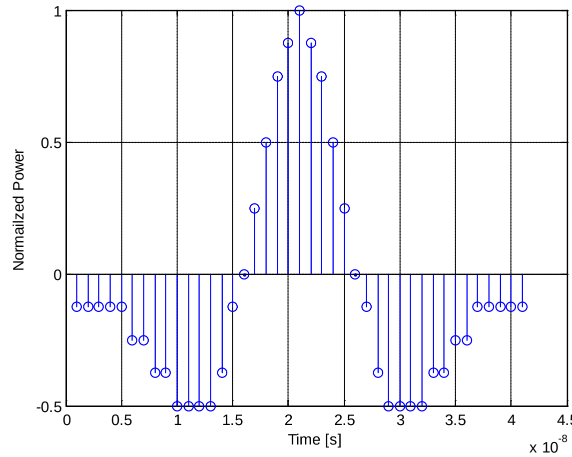


Figure 2. Quantization of the monocycle pulse.

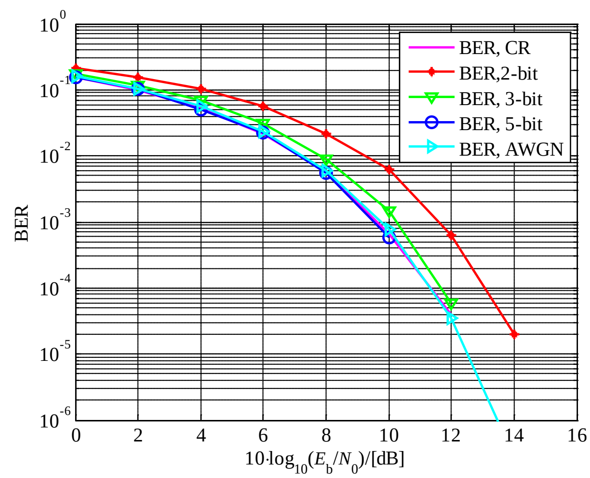


Figure 3. Efffects of quaintazition on BER performance.

To create a different emission power according to current levels, we suggest the white-LED lnGaN/GaN to transmit the optical pulses. This white-LED as proposed in [9] shows that emission power in the blue spectrum portion is increased relative to the injected current. Nevertheless, a blue filter at the receiver front-end should be used to gain the blue spectrum power.

C. *Quantization effects*

The quantization of pulse plays important factor in system performance. The resolution of quantization defined by number of bits used to represent the quantization levels. Fig. 3 shows the BER performance of system employs a quantized pulse $p_o(t)$ at different bit resolution. The effect of channel is neglected to distingue the quantization effects and therefore, only optical noise is considered as additive white Gaussian noise (AWGN). The simulation shows BER of Optical system at resolution of 2-bit, 3-bit, 5-bit, and analog signal (computer resolution). The result in Fig. 3 shows that every incremental bit reduces BER. After 5-bit, the error rate is close to that simulated with analog signal.

In the proposed transmitter in [1], the modulation is performed on bit level using PPM and IM is used to optically represent the signal shape in order to control the optical power derived by input current. However, this operation has to be carried out each time, which is power dissipative and time consume. Thus, we introduce a solution by storing the pulse samples values as binary codewords using LUT.

III. LOOK-UP TABLE TRANSMITTER

In this section, a look-up table is proposed to digitally store the quantized pulse in form of binary codewords that represents samples values. This design aims to increase the LED dynamic range and decrease the nonlinearities caused by transition to saturation region. Moreover, the transmitter complexity and electrical power consumption are expected to be minimized.

A. Transmitter structure

The proposed transmitter based LUT in Fig. 4 consists of PPM modulator, counter, and LUT connected to LED groups via resistors and switches. The LUT contains a number of codewords equal to pulse samples. The LUT input is connected to counter output, which address the LUT index. The counter is activated upon receiving modulated unit pulse from modulator. The modulator modulates zero bits as unit pulse with zero delay. Otherwise, the modulator shifts unit pulse by delay time ε. The counter feeds LUT with address of codeword that controls the total optical power of LEDs groups. The number of groups is defined by bit resolution specified in LUT design phase as will be discussed in the next subsection.

The output of LUT is digital binary vector that switch each group of LEDs (On/Off) based on the state of vector element (0/1). Also, each of these groups input current is scaled by a resistor R_N to emit different power levels P_N. Each group consists of number of LEDs N_{LED} to increase the optical power. The first group D_1 connected to most significant bit (MSB) and emits a maximum intensity power P_0 while D_1 emit half of P_0 and so on. The total emitted power is summation of power emitted by all active groups.

$$P_{total} = N_{LED} \sum_{n=1}^{N} P_n \tag{3}$$

$$P_n = \frac{P_{max}}{2^{n-1}} \tag{4}$$

For example, the codeword (10100) activates the first and third LED groups, while second; fourth and fifth remain in off state. The total optical power calculated using (3) and (4).

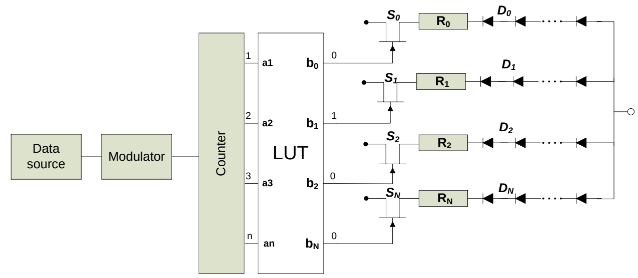


Figure 4. Optical IR-UWB transmitter based LUT.

B. Design of LUT

The LUT is used to store samples values as binary codewords. Fortunately, most of pulse shapes are symmetric, which minimize the LUT size to be adequate for only half of the pulse. This LUT design process is done in offline mode using computer simulations. However, a recalibration mechanism can be adopted to mitigate optical power nonlinearities and fluctuations as explained in next subsection. A digital representation of 5-bit quantized Gaussian pulse consists of 41 samples is sufficient to maintain a good performance as shown in Fig. 3. Table II illustrates a 5-bit binary codewords for half Gaussian pulse used in IR-UWB systems. The codeword length is optimized based on quantization bit depth that gives minimum SNR for a given BER. For example, a 5-bit codeword requires slightly less SNR compared to a 4-bit for BER of 10^{-4} at expense of large codeword length and hence LUT size that is determined by number of bits and number of samples. Although, the 4-bit codeword LUT transmitter needs four groups of LEDs and hence draws less power than 5-bit transmitter. The tradeoff between power consumption, size, and BER performance should be considered based on application. Also, several parameters should be taken in account in order to gain best performance such as, switching speed, resistors values, and maximum allowable LED input current, which are not discussed in this work.

C. LED Nonlinear effects

As mentioned in Section I, the LED suffers from limited dynamic range and nonlinear effects. For most of LEDs, the V-I curve has a limited linear range that can be used to modulate the signal power. This range can be extended further for a small range using predistortion techniques. However, the gain from this extension is not worthwhile, since only a small portion is achieved before the LED reaches the maximum allowable AC/pulsed current. In the proposed transmitter, the LED input current is fixed and scaled by resistors to keep LED operating in the linear region and controlling the total optical power by parallel switching of LEDs groups rather than direct current-power relation. However, the LED thermal behavior decreases the power conversion efficiency causing a drop in light intensity. To mitigate the thermal effects, the LUT should be recalibrated in the real-time to maintain constant optical power.

TABLE II. FIVE-BIT BINARY CODE REPRESENTATION OF GAUSSIAN PULSE.

Counter	Codeword
1	01111
2	01111
3	01111
4	01110
5	01110
6	01101
7	01100
8	01011
9	01010
10	01001
11	01000
12	01001
13	01001
14	01011
15	01110
16	10001
17	10101
18	11001
19	11101
20	11111

The emitted optical power is generated according to digital codewords, and hence can be digitally adjusted or corrected to emit more or less power. This can be done by adding one more bin to the output of LUT connected to an extra LED group for intensity correction. This additive group compensates the optical power in case of power fluctuations using real-time measurements for the emitted power via optical sensor. The data is sent back through feedback as a digital codeword to compare it with the current values. Moreover, the LUT values can be also updated based on the measurements to obtain more power efficiency. As a result, the nonlinear effect from LED whether thermal or electrical can be mitigated and constant optical power will be achieved.

IV. SYSTEM MODEL

For any wireless communication system, the multipath channel differentiates the transmitted and received signal. The path loss and dispersion determines the system design parameters. This section describes the modeling of optical wireless system in indoor environment. A simulation of channel impulse response and SNR in an office room will be discussed. Also, the receiver design for the proposed system including the optical front-end will be presented.

A. Optical Multipath Channel

The optical multipath channel is characterized by an impulse response $h(t)$, which describes the propagation of optical signal between the transmitter and receiver. The propagation pattern is approximated by lambertian radiation pattern, which state that the light intensity emitted from a source has a cosine dependence on the angle of emission with respect to the surface normal [10][11]. The luminous intensity in angle ϕ is given by

$$I(\phi) = I(0)cos^m(\phi) \tag{5}$$

where $I(0)$ is the center luminous intensity of the LED and ϕ is the angle of irradiance, m is the order of lambertian emission and is given by the semi angle at half illuminance of the LED $\phi_{1/2}$ as

$$m = \frac{-\ln(2)}{\ln(\cos\phi_{1/2})} \tag{6}$$

and the horizontal illuminance I_{hor} at a point (x, y, z) on the working plane is defined as

$$I_{hor}(x, y, z) = \frac{I(0)cos^m(\phi)}{d^2\cos(\psi)} \tag{7}$$

where d is the distance between the transmitter and receiver and ψ is the angle of incidence.

In an office room environment, the light arrive receiver directly (LOS link) or after number of reflections (diffuse link). The impulse response at zero reflection is given as:

$$\begin{aligned} h_{los} = \frac{A_r(m+1)}{2\pi d^2} cos^m(\phi)T_s(\psi)g(\psi) \\ \cdot\cos(\psi)\delta\left(t - \frac{d}{c}\right) . 0 \le \psi \le \psi_{con} \end{aligned} \tag{8}$$

where $T_s(\psi)$ is the filter transmission, $g(\psi)$and $\psi_{con}$are the concentrator gain and field of view (FOV), respectively. The gain of the optical concentrator at the receiver is defined by:

$$g(\psi) = \begin{cases} \frac{n^2}{\sin^2\psi_{con}}, & 0 \le \psi \le \psi_{con} \\ 0, & 0 \ge \psi_{con} \end{cases} \tag{9}$$

where n is the refractive index. To model the reflections, every wall is partitioned to a number of small areas. Each one acts as new lambertian source when light incident on it. The impulse for the first reflection is given by:

$$h_{ref}(t) = \begin{cases} \frac{A_r(m+1)}{2(\pi d_1 d_2)^2}\ \rho\ A_{wall}\ cos^m(\phi_r)\ \cos(\psi_r) \\ \cos(\alpha_{ir})\cos(\beta_{ir})\ T_s(\psi)\ g(\psi_r) \\ \delta\left(t - \frac{d_1+d_2}{c}\right), \qquad 0 \le \psi_r \le \psi_{con} \\ 0, \qquad\qquad \psi_r \ge \psi_{con} \end{cases} \tag{10}$$

where d_1 and d_2 are the distances between the LED and a reflective point, and between a reflective point and a receiver surface, ρ is the reflectance factor, A_{wall} is a reflective area. The angles α_{ir} and β_{ir} represent angle of incidence to a reflective point and angle of irradiance to a receiver, respectively, ϕ_r and ψ_r are the angle of irradiance from LED to a reflective point and angle of incidence from reflective point to a receiver.

The optical channel is characterized by the room dimensions, reflectance indices of walls, and transmitter and receiver orientation. Table III describes the parameters used to simulate the channel impulse response. The optical impulse response is used to calculate the channel gain, which is important to estimate the influence of channel on the received power. The power contained in a LOS component h_{los} is larger than the power contained in first reflection components h_{ref} as shown in Fig. 5. The long distance and reflection from the surfaces introduce power loss and longer delay. The received power in NLOS case can be calculated as sum of LOS and NLOS pathlosses multiplied by transmission power. The received power can be expressed as

$$p_r = \left(p_t H_{los}(0) + \int p_t H_{ref}(0)\right) \tag{11}$$

where p_t represents transmitted power, and $H_{los}(0)$, $H_{ref}(0)$ represent power in direct and reflected paths, respectively.

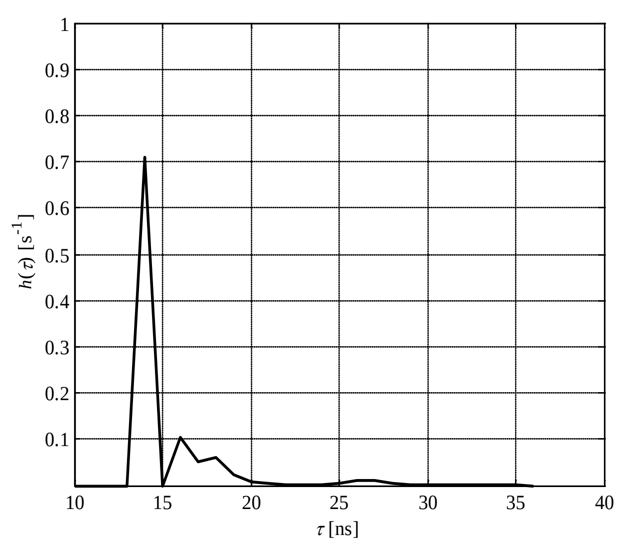


Figure 5. Optical channel impulse response.

Another important future is the root mean squared (RMS) delay, which describes how much delay added by the channel. A large delay led to ISI, which make the detection of transmitted signal complicated. The RMS delay calculated from the channel impulse response as

$$\tau_{rms} = \sqrt{\frac{(t-\tau_0)^2 h^2(t)dt}{\int h^2(t)dt}} \tag{12}$$

where τ_0 is mean delay time defined by:

$$\tau_0 = \frac{\int t\ h^2(t)dt}{\int\ h^2(t)dt} \tag{13}$$

Using equations (12), (13) and parameters in Table III, RMS delay time can be calculated in case of LOS and NLOS channel as shown in Fig. 6 and Fig. 7. This parameter determines the upper bound of transmission rate. For the proposed system, the symbol time is more than the simulated RMS delay time in case of LOS scenario. Hence, equalization stage is not necessary.

TABLE III. THE CHANNEL PARAMETERS.

	Parameter	Value
Room	Room size	5×5×3 m^3
	ρ_{wall}	0.8
Transmitter	Location	(2.5, 2.5, 3)
	M	1
	Elevation	-90°
	Azimuth	0°
	Power	1
Receiver	Location	(0.5,1,0)
	A_r	1 cm^2
	FOV	60°
	Elevation	90°
	Azimuth	0°

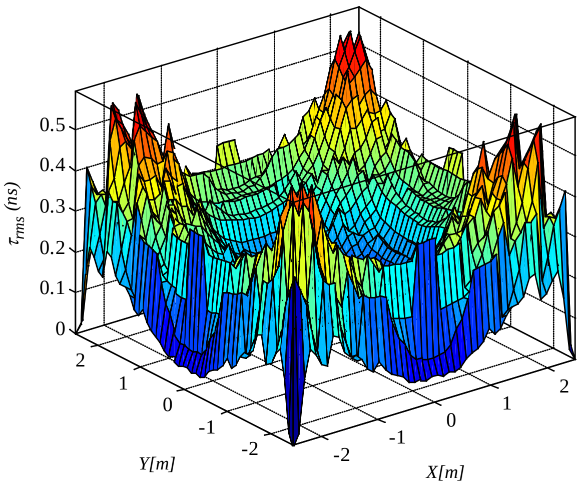


Figure 6. RMS delay for LOS configuration.

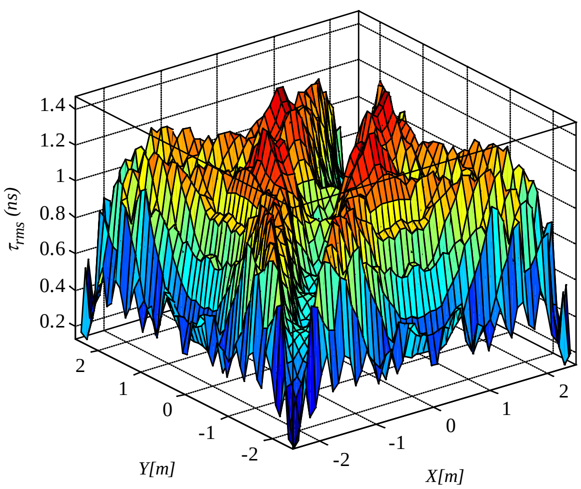


Figure 7. RMS delay for NLOS configuration

B. SNR

In optical systems, SNR is defined by received power p_r, photodiode responsivity [A/W], and noise variances of the shot noise σ_{sh}^2 and thermal noise σ_{th}^2 as [12]

$$SNR = \frac{(Rp_r)^2}{(\sigma_{sh}^2 + \sigma_{th}^2)} \tag{14}$$

The noise sources are classified into two types namely shot noise and thermal noise. The shot noise is a time-varying process generated by external light sources like background noise and quantum noise or internal source as intensity radiation, dark noise and excess noise. These sources are independent Poisson random variables and their photoelectron emission follows the distribution of Poisson distribution with mean equal to the sum of the individual processes. The variance of any shot noise process associated with photodetection is represented as [12]:

$$\sigma_{sh}^2 = 2qB\langle i \rangle \tag{15}$$

where q is the electronic charge, and B is the equivalent bandwidth, and $\langle i \rangle$ is the mean current generated by $\langle n \rangle$ electron. However, if the photoelectron count is large, the generated signal current probability distribution can be approximated as Gaussian process [12].

$$p(i) = \frac{1}{\sqrt{2\pi\sigma_{sh}^2}} exp\left(\frac{[i - \langle i \rangle]^2}{2\sigma_{sh}^2}\right) \tag{16}$$

In addition to the shot noise, the thermal noise caused by thermal fluctuation of electrons in receiver circuit generates a random current, which is modeled as Gaussian process has a zero mean and its variance described as

$$\sigma_{th}^2 = \frac{4\kappa T_k B}{R_L} \tag{17}$$

where κ is Boltzmann's constant, T_k is absolute temperature, and R_L is the equivalent resistance. The total generated current probability distribution of thermal noise and shot noise can be represented as in [12]

$$p(i) = \frac{1}{\sqrt{2\pi[\sigma_{sh}^2 + \sigma_{th}^2]}} exp\left(\frac{[i - \langle i \rangle]^2}{2[\sigma_{sh}^2 + \sigma_{th}^2]}\right) \tag{18}$$

To evaluate channel variations, the room is partitioned to small areas equal to 225. The channel power is simulated using parameters in Table III. The received power in any place in room of size (5m×5m×3m) is estimated as shown in Fig. 8. The minimum received power is -2.3dBm at corners while maximum received power reaches 2.6dBm in room center. The utilization of large number of LEDs provides high received power that is required for good quality communications. However, multiple sources results in path difference that leads to ISI at the receiver and degrade the performance [13].

The received power needed to achieve a BER of 10^{-6} in PPM modulation is to be calculated as [13]:

$$BER = Q\left(\sqrt{SNR}\right) \tag{19}$$

As shown in Fig. 9, the range of SNR is between 11dB and 40dB for a total power summation of LOS path and NLOS paths. The noise parameters required to calculate shot noise and thermal noise are illustrated in [13]. The SNR is obtained using received power shown in Fig. 8 and calculated noise power. The proposed system can achieved the required low data rate at minimum data errors. In absence of LOS power, the received power is reduced dramatically.

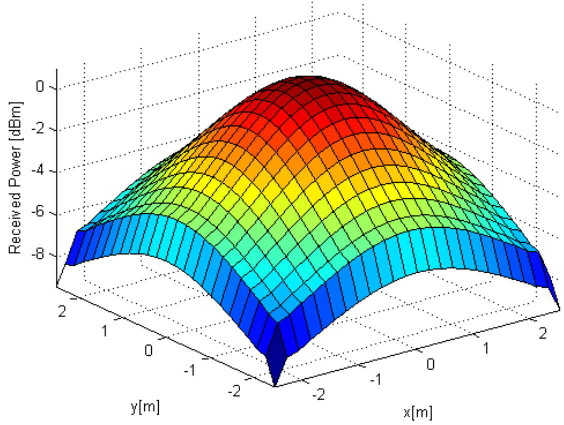


Figure 8. Received optical power.

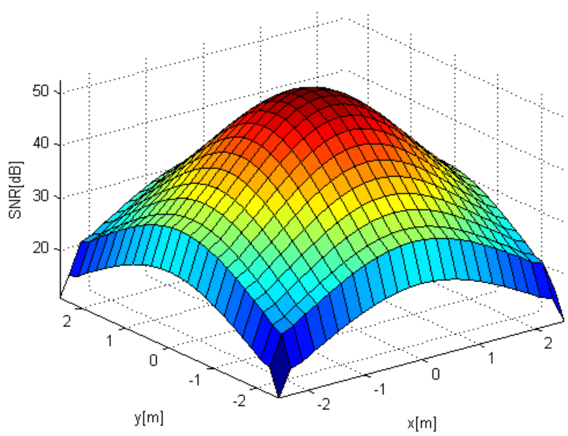


Figure 9. Simulation of SNR.

C. Receiver Design

The receiver of optical wireless system shown in Fig. 10 is based mainly on photodiode (PD) employing the direct detection. The detector area and the orientation play important role in the receiver design and performance. The PD generates an output photocurrent relative to the incident light power pinging on the surface, i.e., the changes produced in intensity modulation at the transmitter are detected by direct detection at the receiver. The photocurrents induced by PD form a replica of the transmitted pulse. At the receiver front-end, the received signal is defined as

$$y(t) = R\, I(t) * h(t) + n(t) \tag{20}$$

The transmitted current signal $I(t)$ is convoluted with optical multipath channel $h(t)$, and added to AWGN $n(t)$. After conversion of optical signal to electrical signal in receiver, the correlation between the mask of transmitted pulse and the

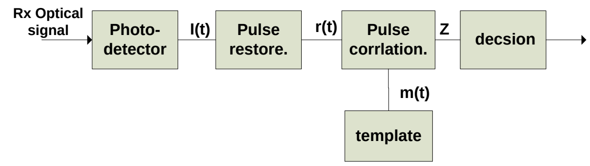


Figure 10. Optical IR-UWB Receiver

received signal is performed as:

$$m(t) = I(t-\tau-kT_s) - I(t-\tau-kT_s-\varepsilon)$$
$$Z = \int_{\tau}^{T_s+\tau} y(t)m(t)dt \tag{21}$$

The detector compares the power of correlation to a threshold and decides whether the received bit is '0' or '1' [14].

$$\hat{a} = \begin{cases} Z > 0, & \hat{a} = 0 \\ Z < 0, & \hat{a} = 1 \end{cases} \tag{22}$$

V. Simulation and Results

In this section, we simulate system design presented in [1] and the proposed system using transmitter based LUT. Both systems are simulated using Monte Carlo simulation in Matlab program.

The system design represented in Fig. 1 is simulated using MATLAB program. In this design, the transmitter based LUT method substitutes the former one presented in [1]. The Monte Carlo simulations were carried out to generate the bit-error-ratio (BER) versus E_b/N_0 results. The information bits are modulated by 2-PPM modulator with symbol time $T_s = 240ns$. The sampling frequency $f_c = 1GHz$ and shift time $\varepsilon = 120ns$, $N_f = 1$. The monocycle Gaussian pulse is used with width $T_p = 41ns$. The pulse samples are quantized to $L = 32$ current levels and stored in LUT as binary codeword. These code words drive five LED groups each consists of eight LEDs. Each of these groups emits a different power level that contributes in evaluation of the intended pulse sample value as explained earlier. The optical channel impulse response in Fig. 5 is simulated in a room with dimensions of (5m×5m×3m) and a fixed transmitter and receiver are assumed. The optical pulse is convolved with the channel and added to the noise.

On the receiver side, the photodiode is percepts the incident light and converts it into current. We assume that the detector responsivity R equal to one. The received pulse constructed from current levels is correlated with the mask of transmitted pulse and the peak power is compared to the threshold in time window. This system is assumed to be synchronized and no equalization stage is performed.

In OWC systems, a unit rectangular pulse is used to transmit the power of modulated binary bits with duration

T_s as proposed in [15][16]. In Fig. 11, the same system design was simulated using rectangular pulse and monocycle pulse to compare the BER performances of using the shaped pulse used for wireless system and rectangular pulse used in OWC systems. Fig. 11 shows that the BER for both signals are equal because of power direct detection, which is evident that other pulses shape could be used without loss of performances. Although rectangular pulse evaluation is simpler than Gaussian pulse, the later introduces capability for utilizing advantages of a designed UWB radio wireless system communicating on optical link in sensitive environments. Nevertheless, effects of LEDs nonlinearities and shot noise are expect to disfigurement the transmitted Gaussian pulse at transmitter and receiver front-ends. These effects will be studied experimentally in the future work to find the performance degradation for the proposed design. Fig. 12 compares the BER performance of the proposed systems operate on LOS optical wireless channel with that on diffuse channel in the absence of LOS link. The direct path between transmitter and receiver delivers higher power than paths reach the PD after reflections. This explains why the BER in the presence of LOS channel is lower than that in diffuse channel by ~2dB. Also, the influences of optical wireless channel gain and delay on the proposed system increase the BER by ~4dB compared to the AWGN bound.

The second simulation determines BER performances for the proposed design in case of channel variation during data transmission. In this case, the mobile user changes her/her location inside office environment with dimensions listed in Table III. Therefore, we simulate channel impulse responses for 2500 points that represent possible receiver's location in X-Y plane. The channel is simulated between fixed transmitter and receiver locations.

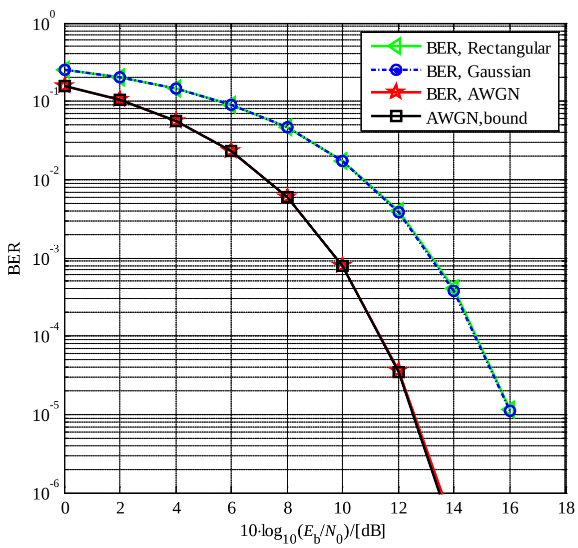


Figure 11. BER of the system with rectangular and Gaussian Pulses.

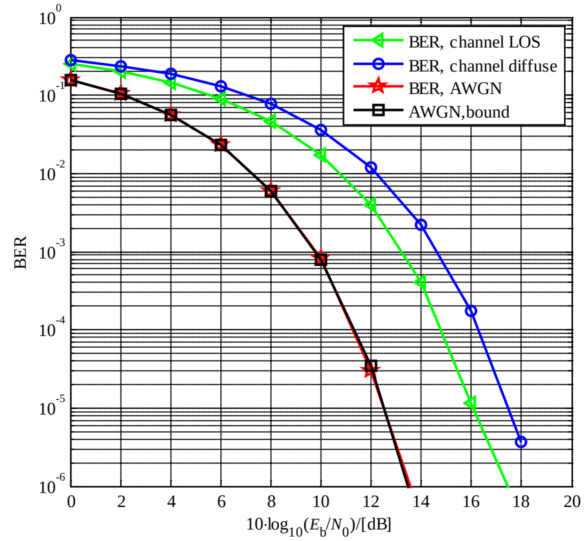


Figure 12. BER of the system with LOS and diffuse link configurations

Moreover, we have distinguished two scenarios for each observation. First, the LOS link is not blocked by obstacles and both of direct radiation and reflected radiations are present. In the second scenario, NLOS link was established and only reflected radiation is received regardless receiver's position. In Fig. 13 and Fig. 14, the channel impulse responses are depicted. As seen from figures, the channels in second scenario experiences a larger pathloss than in first scenario due to diffusion phenomena in NLOS case and absence of LOS power contribution, which contain largest portion of received signal. Nevertheless, the system simulation for first scenario is done over 2500 channels. We assume that a channel response changes once within frame time duration $T_s = 240ns$. The user's location number is determined as uniform random variable.

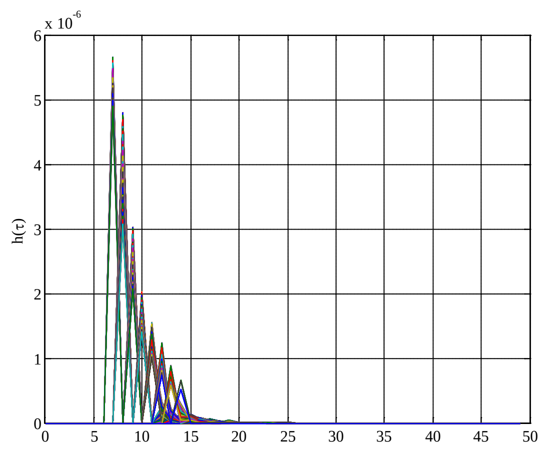


Figure 13. Channel impulse responses for LOS and NLOS links.

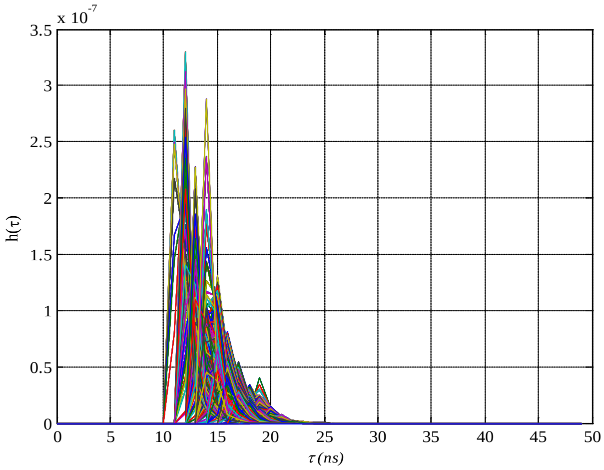


Figure 14. Channel impulse responses for NLOS links

.

Fig. 15 shows system's BER under channel response composed of LOS and NLOS radiations. Likewise, the system is simulated for second scenario where LOS link is shadowed over transmission time. Fig. 15 shows larger BER than presented by first scenario due to larger channel pathloss and delay. A third case is simulated, when channel changes over 5000 channel observations simulated in first and second scenarios. The channel configuration and user's location is determined as uniformly-distributed random variable. This case is more realistic where channel response varieties between two scenarios presented above. During user mobility, the user moves in an office where some positions are shadowed by obstacles or sudden obstacle is taken place. The system BER is closely better to that presented in second scenario as shown Fig. 15. These scenarios summarized system's BER performance for mobile user situation. It worthwhile to recall that, the user's location is randomized as uniform-distributed random variable.

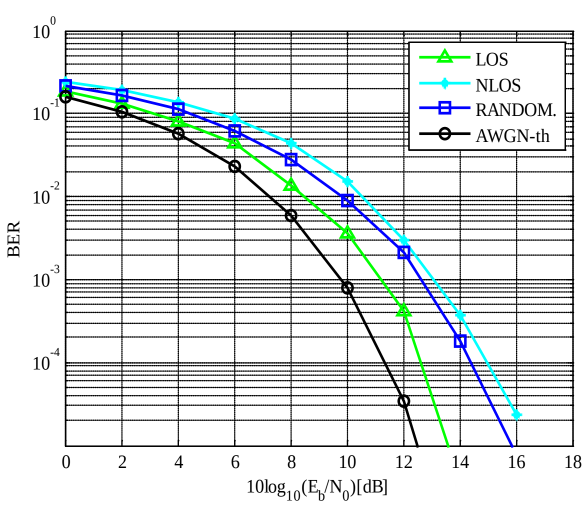


Figure 15. BER of proposed system in mobility scinarios

VI. CONCLUSIONS AND FUTURE WORKS

In this paper, an optical wireless transmitter based LUT was introduced. The radio system pulse was translated to binary codewords stored in LUT. The needs for ADC and TCA were excluded as well as pulse reformation for each symbol time. The utilization of LUT enables more linear characteristics of LEDs. The transmitter complexity and power consumption were reduced as compared with transmitter employing DC-bias and direct quantization. The optical multipath channel was investigated and the system performances using Monte-Carlo simulations have been obtained and analyzed. Also, the BER performances for fixed and mobile user in office environment were simulated and analyzed. This design will be suitable for non-radio environment like hospitals or for systems that operate in both optical/radio configurations. This work will be continued to evaluate better performance using predistortion techniques to adapt the LUT and hence more linearity for LEDs power.

REFERENCES

[1] M. Al-Olofi, A. Waadt, G. H. Bruck, and P. Jung, "Design of optical wireless IR-UWB wystems for Low Data Rate applications," Proc. of the Ninth Advanced International Conference on Telecommunications AICT 2013, pp. 7-12, Rome 2013, ISBN: 978-1-61208-279-0

[2] C. L. Liao, Y. F. Chang, C. L. Ho, and M. C. Wu, "High-Speed GaN-Based Blue Light-Emitting Diodes with Gallium-Doped ZnO current spreading layer," Electron Device Letters, IEEE , vol. 34, no. 5, pp. 611-613, May 2013 doi: 10.1109/LED.2013.2252457.

[3] T. Fath, C. Heller, and H. Haas, "Optical wireless transmitter employing discrete power level stepping," Journal of Lightwave Technology, vol. 31, no. 11, pp. 1734,1743, June, 2013 doi: 10.1109/JLT.2013.2257984

[4] Z. Zhu and X. Huang, "Design of an LUT-Based predistorter for power amplifier lnearization," Advances in Satellite and Space Communications, 2009, pp. 113-116, July 2009, doi: 10.1109/SPACOMM.2009.13

[5] J. K. Cavers, "Amplifier linearization using a digital predistorter with fast adaptation and low memory requirements," IEEE Trans. Veh. Technol., vol. 39, no. 4, pp. 374-382, Nov. 1990.

[6] K. J. Muhonen, M. Kavehrad, and R. Krishnamoorthy, "Look-up table techniques for adaptive digital predistortion: A development and comparison," IEEE Trans. Veh. Technol., vol. 49, no. 9, pp. 1995-2002, Sep. 2000.

[7] S. W. Chung, J. W. Holloway, and J. L. Dawson, "Energy-Efficient Digital Predistortion With Lookup Table Training Using Analog Cartesian Feedback," IEEE Transactions on Microwave Theory and Techniques, vol. 56, no. 10, pp. 2248-2258, Oct. 2008, doi: 10.1109/TMTT.2008.2003139

[8] H. F. Rashvand, Y. Zeng, R. J. Green, and M. S. Leeson, "Look-up table error correcting multiple pulse PPM codes for wireless optical communication channels," IET , vol. 2, no. 1, January 2008, pp. 27-34, doi: 10.1049/iet-com:20060386

[9] J. Grubor, S. C. J. Lee, K. D. Langer, T. Koonen, and J. W. Walewski, "Wireless high-speed data transmission with phosphorescent white-light LEDs," in Proc. Eur. Conf. Optical Communications (ECOC 2007), Berlin, Germany, Sept. 2007, pp. 1-2.

[10] J. R. Barry, J. M. Kahn, W. J. Krause, E. A. Lee, and D. G. Messerschmitt, "Simulation of multipath impulse response for

wireless optical channels," IEEE J. Select. Areas in Communications, vol. 11, no. 3, Apr. 1993, pp. 367–379, doi: 10.1109/49.219552.

[11] H.Q. Nguyen, J. Choi, M. Kang, Z. Ghassemlooy, D. H. Kim, S. Lim, T. Kang, and C. G. Lee, "A MATLAB-based simulation program for indoor visible light communication system," CSNDSP 2010 Proceedings, July 2010, pp. 537-541.

[12] Z. Ghassemlooy, W. Popoola, and S. Rajbhandari, Optical wireless communications: system and channel modelling, CRC Press, 2012, pp. 66-74.

[13] T. Komine and M. Nakagawa, "Fundamental analysis for Visible-Light Communication system using LED light," IEEE Transaction on Consumer Electronics, Vol. 50, No. 1, February 2004, pp. 100-107, doi: 10.1109/TCE.2004.1277847.

[14] M. G. Di Benedetto, Understanding Ultra Wide Band radio fundamentals, Pearson Education, 2008, pp. 241-252.

[15] M.D Audeh, J.M Kahn, and J.R. Barry, "Performance of pulse-position modulation on measured non-directed indoor infrared channels," IEEE Transactions on Communications, vol. 44, no. 6, Jun. 1996, pp. 654-659, doi: 10.1109/26.506380.

[16] J. Zhang, "Modulation analysis for outdoors applications of optical wireless communications," International Conference on Communication Technology Proceedings, 2000. WCC - ICCT 2000, vol. 2, no. 2000, pp. 1483-1487, doi: 10.1109/ICCT.2000.890940.

LEOCast: An Optical Multicast Protocol for LEO Satellites based on Optical Codewords

Maha Sliti *, Walid Abdallah †, and Noureddine Boudriga ‡
Communication Networks and Security Research Lab, University of Carthage
Ariana, Tunisia
Email: *slitimaha@gmail.com, †ab.walid@gmail.com, ‡nab@supcom.rnu.tn

***Abstract*—Satellite networks provide worldwide coverage and support a large variety of services. Since Low Earth Orbit (LEO) satellites offer short round-trip delays, they are progressively becoming more and more important for real-time services, such as voice, data and video traffic. The majority of these services need a mechanism to deliver data to several receivers. In the present work, we propose the LEOCast protocol, which is an Optical Multicast Protocol for LEO Satellites based on Optical Codewords. This protocol can function in two modes: the first mode is based on the shortest paths and the second mode is based on the virtual multicast trees paradigm. The optical multicast protocol considers free space inter-satellites links that provide high data rate transmission for real time services and it is based on the optical switching concept using optical codewords. The two proposed multicast modes are compared in terms of the traffic load and the number of used decoding functions. It is shown that the optical multicast mode based on the virtual multicast tree concept generates less traffic and uses less decoding functions when performing the optical multicast process on each intermediate satellite. However, it is demonstrated that each multicast mode is adapted to a type of payload (packet/burst) depending on the required processing delay for each multicast mode.**

Keywords–LEO constellation; all optical multicast; virtual multicast tree; structured codewords; tunable optical decoder.

I. INTRODUCTION

Satellite-based optical communication systems have become a most promising technology for their global coverage and high-speed inter-satellite and satellite-to-ground communication links. Compared to communication satellites in geostationary orbit, the communication links to Low Earth Orbit (LEO) satellites are characterized by lower propagation delay and lower link attenuation because of the shorter distance, resulting in the need for reduced transmission power. Consequently, LEO satellites are becoming increasingly important for real-time applications such as voice, teleconferencing and video traffic, which require a mechanism to deliver information to multiple recipients. However, due to the rapidly and periodically changing of the network topology caused by the high mobility of satellites, the routing in satellite networks face great challenges.

In [1], we propose the LEOCast protocol, which is an Optical Multicast Protocol for LEO Satellites based on Optical Codewords. This protocol can function in two modes: the first mode is based on the shortest paths and the second mode is based on the concept of virtual multicast trees. For the two modes, we identify each satellite in the LEO constellation network by an optical codeword. In the first mode, each intermediate satellite considers four codewords indicating its direct neighbors: two in the same orbit, and two in the neighboring orbits. A shortest path is defined in terms of the number of hops and is established by favoring the inter-orbit on the intra-orbit hop. In the second mode, the virtual multicast paradigm is used in order to underline that the multicast tree is not physically established but only built on the codewords structure used to switch traffic contrary to the first proposed mode that requires a route discovery process preceding the routing of a multicast traffic. In this approach, the virtual tree establishment consists on the management of codewords structure.

In the literature, several multicast routing protocols were proposed for satellite networks [2]–[11]. These approaches are mainly proposed in the context of IP (Internet Protocol) and mobile networks and then adapted to support multicast in LEO networks. However, the multicast process in satellite networks present more specific requirements compared to other types of networks such as the need of high data bit rate transmission and the scalability to support important multicast group and payload sizes. These requirements can be achieved by optical communication inter-satellite links.

The key contributions of our proposed LEOCast protocol with respect to the previously published research are:

1) The proposed multicast process is performed at the optical layer based on optical switching, which allows the multicast packets to be processed at very high bit rates (the order of gigabits/second) without conversion to the electronic domain. During the optical multicast process, the received packets are delayed in an optical buffer proposed in [12], in order to provide a tunable delay for real time traffic. We have chosen an optical multicast approach due to the fact that the transmission quality of the light beam is near perfect. Indeed, this latter will not be affected by the attenuation and the dispersion effects in free space.
2) The optical switching concept is based on the optical codewords, which are represented by a sequence of pulses. Indeed, a codeword is assigned to each satellite in the network and serves as an optical identifier of the satellite. Based on the received codeword and the structure we build in, the traffic will be multicasted to one or several directions allowing to reach the destination satellites.
3) The LEOCast protocol is scalable since its performance is not affected by the multicast group size and the member combination in the multicast group.

Therefore, the optical multicast module implemented in each intermediate satellite is at most composed of four tunable decoders.

4) The parameters of the proposed multicast module can be dimensioned depending on the traffic stream estimation. These parameters can be the number of tunable decoders, the number of loops that composes each decoder based on the length of codewords, the fiber length of the Virtual Optical Memory (VOM) [12] based on the size of packet, etc.

In this paper, we have extended the proposed work in [1], by adding the following contributions:

- At the presentation level: First, we have investigated recent multicast approaches proposed for LEO constellations, and we have presented the multicast requirements in such type of networks. Second, we have defined the optical codewords, their mathematical characteristics and the principle of their encoding/decoding process. Third, we have illustrated the optical codeword structures that are used by the two multicast modes, and we have discussed the type of payload (packet/burst) adapted to each multicast mode. Finally, we have presented the advantage of our multicast protocol compared to the other protocols by presenting the advantage of the optical link use. The new added points have allowed to complete and to extend our reference list.
- At the design level: we have improved the design of the optical multicast mode based on the virtual multicast tree concept. Indeed, the header attached to the payload is well structured in order to reflect the virtual multicast tree and to optically switch the multicast traffic without the need to establish the route from the source to destination satellites. We have also considered two multicast directions: front/backward instead of right/left. Infact, due to the mobility of satellites, we need to consider the direction of movement in the orbit and not the localization of satellites. Thus, each intermediate satellite must treat two types of traffic: an inter-orbit traffic appropriated to destination satellites in other orbits and an intra-orbit traffic appropriated to destination satellites in the same orbit.
- At the simulation level, we have extended the proposed simulation by assessing the optical signal quality when multicasted from a transmitter satellite to other satellite neighbors using Free Space Optical links (FSO) with specific characteristics (distance and wavelength). The objective is to evaluate the effect of these criteria on the output signal quality. The optical signal quality is assessed based on two main criteria: the wavelength and the distance between satellites. We can conclude, when performing the optical multicast process, that the optical signal quality increases when decreasing the FSO communication link wavelength and the distance between receivers.

The rest of the paper is organized as follows. Section II presents the proposed multicast approaches for the LEO constellation networks in the literature. The optical codeword concept, the code sequences characteristics and the codewords structures used by the two optical multicast modes are presented in Section III. The codeword based switching process, which is the main optical signal operation in each satellite, and the optical codeword association to each satellite in the constellation, are presented in Section IV. In Section V, the all optical multicast mode based on the shortest paths is described. In Section VI, the all optical multicast mode based on the virtual multicast trees is explained. The mobility management in the two all optical multicasting approaches is discussed in Section VII. Simulations and experimental results are given in Section VIII. Finally, Section IX concludes the paper.

II. Multicast in LEO constellation networks

With recent needs of high speed communication systems, Free Space Optical links (FSO) become a most promising technology for high-speed inter-satellite and satellite-to-ground communication links [13]–[15]. A free space optical communication system includes optical transmitter and receiver satellites.

This type of links is preferred over Radio Frequency (RF) communication because of having narrower beam widths due to use of lasers, reducing the size of used antenna, which reduces the weight of the satellite, minimizing the power used for the communication system, and offering higher data rate.

The advantages of an optical communication link compared to a RF link in free space are characterized by: 1) high data rate, 2) less transmitter power consumption, 3) terminal design with reduced size and weight, and 4) transparency to RF interference.

Satellites can be directly linked by Inter-Satellite Links (ISLs) to other satellites in the constellation as described in Figure 1. ISLs provide direct communication paths between satellites. The ISLs between a satellite and its neighbors in the same orbital plane are called intra-satellite links, and its links with its neighbors in neighboring orbits are called inter-satellite links.

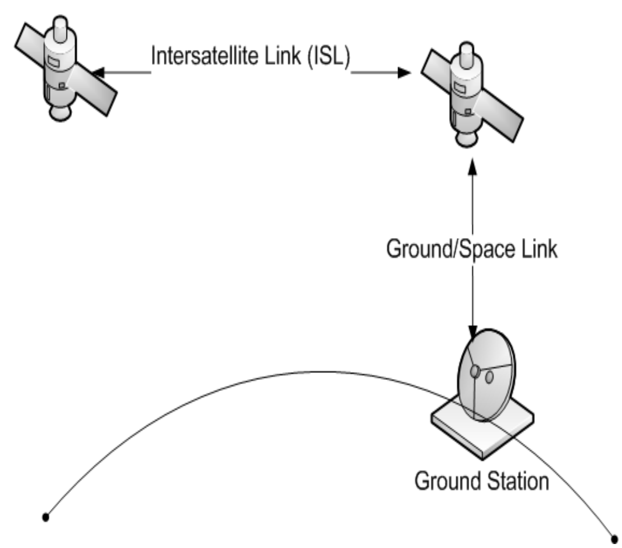


Figure 1. Inter-Satellite Links (ISL).

There are several orbits available for satellites to reside. The orbits are low Earth orbit (LEO), medium Earth orbit (MEO), highly elliptical orbit (HEO) and geosynchronous orbit (GEO). LEO satellites are orbiting at low earth orbits with an altitude generally between 500 km and 2000 km. Compared to communication satellites in geostationary orbit, the communication links to LEO satellites are characterized by lower propagation delay and lower link attenuation because of the shorter distance, resulting in the need for reduced transmission power. Since LEO satellites, presented in Figure 2, provide low propagation delay and low power requirements, they are becoming increasingly important for real-time services, which require a mechanism to forward data to several receivers. However, LEO satellite systems have mobile network topologies and this dynamic topology makes data multicast difficult.

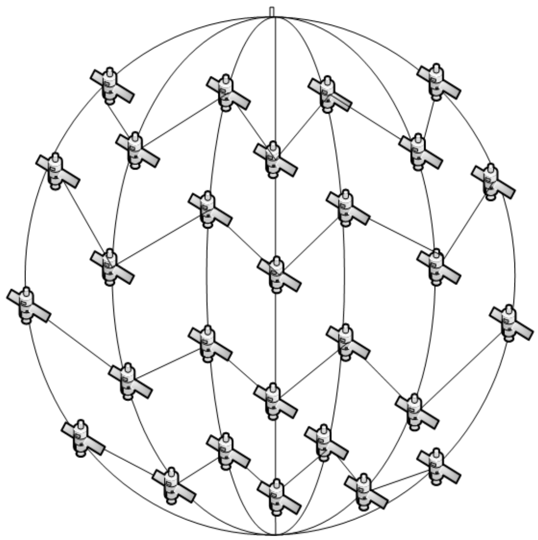


Figure 2. The topology of LEO satellite networks.

In the literature, some multicast protocols including Internet Group Management Protocol (IGMP) [2], Reverse-Path Multicast (RPM) [3], Distance Vector Multicast Routing Protocol (DVMRP) [4], and the Multicast Extensions to Open Shortest Path First (MOSPF) [5] consider periodic message exchanges for the establishment and the management of multicast trees, which can overload the network mainly in the case of long communication periods between source and destination users during a multicast session. At present, only a few multicast routing schemes in the literature, have been developed for satellite networks. In [6], a new core-based shared tree algorithm, viz Core-cluster Combination-based Shared Tree (CCST) algorithm and the weighted version (i.e., w-CCST algorithm) are proposed in order to resolve the channel resources waste problem in typical source-based multicast routing algorithms in LEO satellite IP networks. In [7], the authors associate to every multicasting routing problem a Steiner tree problem. In this paper, a dynamic QoS (Quality of Service) routing mechanism is proposed to support multiple Steiner trees. In [8], a fast iterative distributed multicast routing algorithm was developed based on the inherent characteristics of satellite networks, which using distributed computing model and significantly reducing the algorithm computational complexity. In [9], a QoS-Guaranteed Secure Multicast Routing Protocol (QGSMRP) is proposed for satellite IP networks using the logical location concept to isolate the mobility of LEO and HEO satellites. a novel triple-layered satellite network architecture including GEO, HEO, and LEO satellite layers is introduced. In [10], combining LEO satellites' advantages on transmitting real-time information with GEO's ability of big computing, a GEO/LEO Double-Layer Multicast Routing Algorithm (DLMRA) is given. The proposed algorithm aims to minimize cost of multicast trees, under the condition of the available bandwidth bound and multicast tree delay bound. In [11], the authors proposed a multicast routing for LEO satellite constellation networks with high performance. The algorithm uses the group members' geographic information to route multicast packets, with less memory, computer power and signaling overhead.

The proposed approaches are mainly proposed in the context of IP and mobile networks and then adapted to support multicast in LEO networks. However, a multicast protocol used for LEO satellites must address more specific requirements. First, a multicast protocol must provide low latency and low overhead when transferring multicast packets in the network since multicast applications are high data rate real-time application (such as voice and video traffic) and the number of multicast packets can be very important. Second, a multicast protocol must provide scalability when considering different traffic sizes, groups and connection delays. Furthermore, a multicast protocol must insure the security of connections when establishing multicast trees, constructing and transferring packets in the network. Thus, optical communications between satellites, which are not considered by the proposed approaches, can fulfill the cited requirements for a multicast protocol in LEO satellite constellations. In the literature, several approaches studied the optical multicast aspect in Wavelength Division Multiplexing (WDM) networks [16]–[19]. In this context, we propose the LEOCast protocol, which is a multicast protocol based on free space optical communication links between satellites.

III. CODEWORD BASED SIGNAL STRUCTURES

In this section, we introduce the optical codeword concept and its characteristics. Then, we present the codewords' structures used by the two modes of the LEOCast protocol: 1) the mode based on the shortest paths, and 2) the mode based on the virtual multicast trees.

A. Codeword concept

Optical encoding is based on optical codeword sequences. An optical codeword is a set of ("0","1") sequences of length N that satisfies certain auto-correlation and cross-correlation constraints. Each "0" or "1" of a sequence is called a chip.

Optical encoding has a wide range of novel and promising applications, such as label switching and Optical Code Division Multiple Access (OCDMA) multiplexing technology. In OCDMA, each transmitted data bit is optically encoded by a specific pulse sequence. The optical encoding operation consists in representing the data bit by a code sequence either in the time domain, the wavelength domain, or a combination of both (2D-coding) [20]. The decoding operation is performed by the receiver to recover the original data. We define optical

Codeword_1	Codeword_2	Codeword_n	Payload

Figure 3. Shortest path structure.

Codeword_Orbit_current Codeword_sat_dest	Codeword_Orbit_current_direction (front/backward) Codeword_sat_x	Codeword_Orbit_n Codeword_sat_dest	Codeword_Orbit_n_direction (front/backward) Codeword_sat_y	Payload

Figure 4. Virtual tree structure.

coding/decoding as the process by which a code is inscribed into, and extracted from, an optical signal.

B. Code sequences characteristics

An optical orthogonal code [21]–[23] is a family of (”0”, ”1”) sequences with maximum auto-correlation and minimum cross-correlation in order to optimize the differentiation between correct signal and interference. Code family that have the best orthogonal propriety is optical orthogonal codes (OOC). An OOC is defined as a collection C of codewords $C = \{c^1, ..., c^m\}$, characterized by a 4-uplet $(N, w, \lambda_a, \lambda_c)$, where N is the length of a codeword, and w is the Hamming distance of a codeword, and for each c^i, c^j:

- $\sum_{t=0}^{N-1} c_t^i c_{t+k}^i \leq \lambda_a$ for every $k \neq 0$
- $\sum_{t=0}^{N-1} c_t^i c_{t+k}^j \leq \lambda_c$ for every k

Hence, λ_a and λ_c define the constraints on the auto-correlation and the cross-correlation functions, respectively. The above conditions indicate:

- The number of ones in the zero-shift discrete auto-correlation function should be maximized.
- The number of coincidences of non-zero shift discrete auto-correlation function should be minimized.
- The number of coincidences of the discrete cross-correlation function should be minimized.

In the case where $\lambda_a = \lambda_c = \lambda$, the OOC is represented by the 3-tuple (N, w, λ) and called optimal OOC. $\mid C \mid= m$ is the cardinality of the code (i.e., the number of codewords contained in the code family). For a code characterized by (N, w, λ) the maximum number of codewords that belong to this code family, denoted by $\Phi(N, w, \lambda)$ should satisfy [24]:

$$\Phi(N, w, \lambda) \leq \frac{(N-1)(N-2)....(N-\lambda)}{w(w-1)(w-2)...(w-\lambda)} \qquad (1)$$

For $\lambda = 1$, the number of codewords is upper-bounded by:

$$\mid C \mid \leq \left\lfloor \frac{(N-1)}{w(w-1)} \right\rfloor \qquad (2)$$

where $\lfloor C \rfloor$ denotes the integer portion of the real number C.

A (N, w, λ) OOC can be considered as a family of sets (of cardinality w) composed by integers modulo N. Each among these sets specifies the positions of the nonzero bits in the codeword. For instance, the codeword 1101000 in the OOC characterized by $(7, 3, 1)$ can be represented by $\{0, 1, 3\} mod(7)$ because the positions of the bits set to one are respectively 0, 1, and 3.

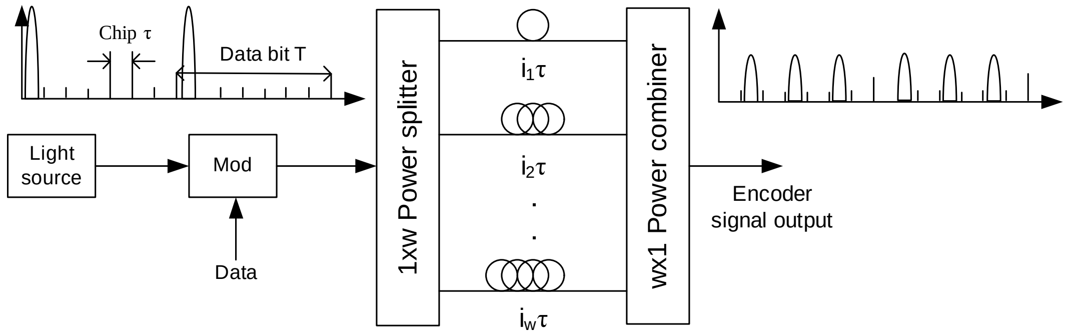
C. Path and tree based codeword for signal transmission

An optical codeword is assigned to each satellite in the LEO constellation network. Therefore, each satellite is uniquely identified by its assigned codeword, which is similar to an address in our case. The total number of associated codewords is equal to $N * M$ with N is the number of orbits (planes) in the constellation, and M is the number of satellites in each orbit. We notice that the number of codewords used for a LEO constellation network is very reduced compared to current terrestrial networks. We suppose that the codewords are generated, assigned and managed by a central entity implemented in a terrestrial station, which is the ground station in our case.

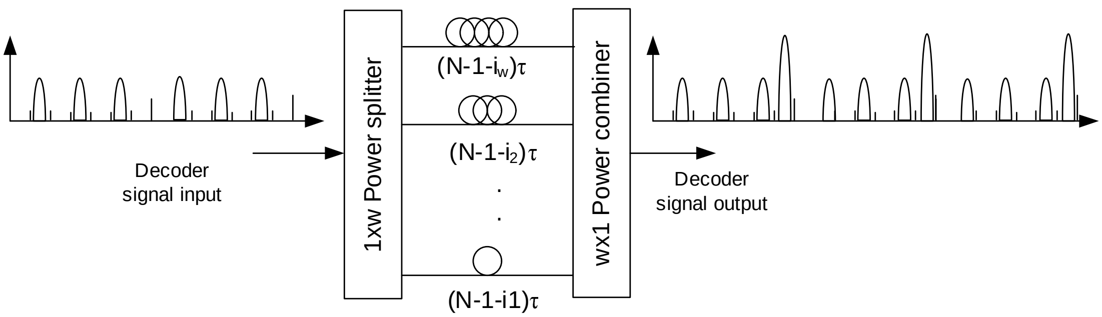
For the two optical multicast modes based the shortest paths and the virtual multicast trees, we consider the following header structures composed of a set of codewords. Figure 3 describes the header structure considered by the first proposed multicast mode based on the shortest paths and that will be presented in the following section. In this approach, we associate to the payload a set of optical codewords that indicate satellites in a shortest path. Figure 4 presents the header structure considered by the second proposed multicast mode, which is based on the paradigm of the virtual multicast trees that will be described in the next sections. As we can notice, the structure of codewords associated to the payload is more complicated compared to the first mode. Indeed, each codeword is structured in two parts: the first part identifies the orbit in the LEO constellation and the second part identifies a satellite in the constellation.

IV. CODEWORD BASED MULTICAST

In this section, we present the codeword based switching process, which is the main optical signal operation performed by each satellite. Furthermore, we describe the optical codeword association to each satellite in the constellation.



(a) All-optical encoder



(b) All-optical decoder

Figure 5. All-Optical Encoding/Decoding

A. *Codeword based switching*

In this work, we propose an optical multicast process based on codewords. Also, we provide a function so that each satellite can optically switch a multicast traffic based on codewords properly structured. Thus, each satellite implements an optical switching module based mainly on two optical operations, which are the optical codeword matching and the deviation of the traffic to the adequate direction. The received codeword is split to a set of decoders; if a codeword is matched by a decoder, an optical switching gate will be activated by a pulse to forward the multicast traffic to the direction that allows reaching the destination satellite. Indeed, each decoder is configured with a codeword associated to a satellite neighbor. Consequently, each multicast module is composed of only 4 tunable decoders associated to the 4 satellite neighbors.

One of the key issues to consider OOC codewords in the optical signal processing operations is how to encode and decode the received OOC. Thus, we need to design and develop encoding and decoding techniques that can generate and reliably recognize appropriate code sequences. Optical encoders and decoders are major components to achieve optical processing signal operations. The optical encoder encodes only the information bit ”1” and does not produce any optical pulse when the information bit ”0” is transmitted.

In a direct sequence optical code, each codeword can be presented by its corresponding codeword block, where each element in every block indicates the position of ”1” in the codeword. Therefore, the structures and operational principles of any temporal incoherent optical encoders/decoders are similar with each other. Each codeword in an OOC $(N, w, \lambda_a, \lambda_c)$ corresponds to a codeword block $\{i_1, i_2, ..., i_w\}$, where i_j represents the position of the j^{th} ”1” of the codeword, $0 \leq i_j \leq N-1$. A fixed and tunable one-dimensional incoherent optical encoder and decoder can be composed of an optical power splitter, a number of fixed or tunable optical delay lines and an optical power combiner.

As it is shown in Figure 5a [25], an optical encoder for a 1-D $(N, w, \lambda_a, \lambda_c)$ OOC consists of a $1 \times w$ optical power splitter, w fiber-optic delay lines and a $w \times 1$ optical power combiner. The delay of the j^{th} fiber-optic delay line is $i_j\tau$, $0 \leq i_j \leq N-1$, where N is the code length of the optical orthogonal code, w is the code weight, and τ is the width of a chip (i.e., the time-width of an optical pulse). At the beginning of a data bit cycle of a user, the light source sends an optical pulse with time-width τ into the optical modulator. The optical modulator outputs an optical pulse when the data bit is “1” and the optical modulator outputs nothing when the data bit is “0”. Then, the optical pulse corresponding to the data bit “1” is encoded by an optical encoder whose output is an optical pulse-signal waveform matching an optical orthogonal codeword. Because there is no optical signal to be input into the optical encoder for a data bit “0”, nothing is output from the optical encoder.

As depicted in Figure 5b [25], a fixed optical decoder whose structure is the same as its corresponding encoder except that the delay of its j^{th} fiber-optic delay line is changed into $(N-1-i_j)\tau$, $0 \leq i_j \leq N-1$. When the input of the decoder is the output signal from its corresponding encoder, its output is an auto-correlation function of its corresponding OOC codeword.

Finally, the data bit will be restored after the optical-to-electrical conversion and threshold decision. If the decoder

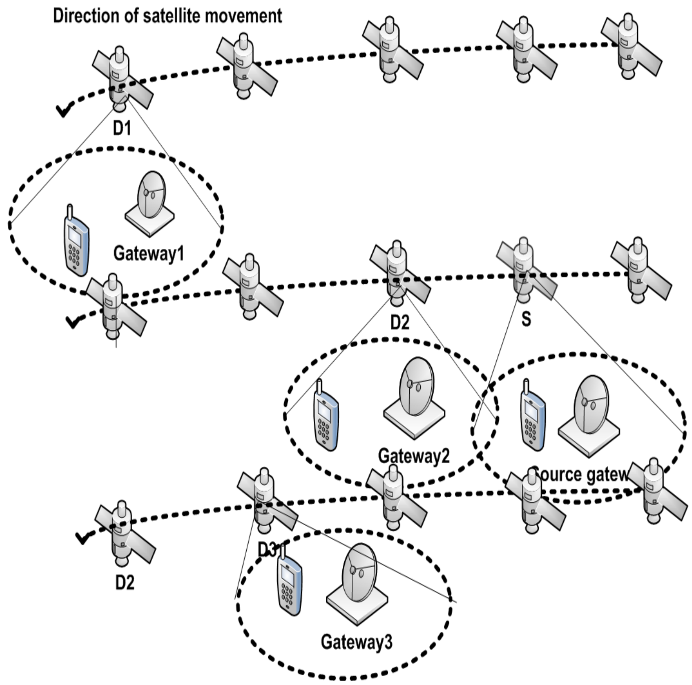


Figure 6. Multicasting in LEO satellite constellation (Figure 1, [1]).

input is an encoded waveform from other OOC codeword, its output is a cross-correlation function. Since an auto-correlation peak does not occur there is no data output.

B. Multicast

Since LEO satellites provide short round-trip delays, they are becoming increasingly important for real-time applications, such as voice and video traffic. Many applications require a mechanism to deliver information to multiple recipients, as illustrated in Figure 6.

Since each group of destination users has a geographic location, each group is covered by a different distribution satellite that we call also a destination satellite. At t_0, a source user is attached to a source satellite and a group of destination users is served by a destination satellite via a wireless link. On the contrary, the links between satellites, which are free space optical communication links.

During the communication period, the group of destination satellites are changed due to the mobility of the source and destination satellites. Thus, the management of the multicast process between the source and the destination satellites must be done during communication periods that may last for hours.

V. LEOCast protocol: All optical multicast mode based on the shortest paths

In this section, we present an optical multicast mode based on the shortest paths. In this approach, each intermediate satellite considers four codewords indicating its direct neighbors: two in the same orbit, and two in the neighboring orbits. A shortest path is defined in term of number of hops and is established by favoring the inter-orbit on the intra-orbit hop. In order to establish the shortest paths to a list of destination satellites, the route discovery process is initiated by the source satellite. The source satellite duplicates a Route REQuest message (RREQ) in order to send it to d destination satellites. The considered RREQ message format is composed of: the message identifier, the codewords associated to destination satellites, the satellite source address, the satellite destination address, and the communication time between the source user and the destination users.

At the reception of a RREQ message, the intermediate satellite adds its codeword to the Multicast list address and sends it to the nearest neighbor based on the Destination address in the RREQ message. A destination satellite that receives the RREQ message, sends a Route RESponse (RRES) message, which contains the shortest path to the source satellite. A path is formed by a list of codewords that denote the intermediate nodes in the shortest path.

After the route discovery process, a source satellite sends the multicast traffic to the d destination satellites. Each multicast packet is duplicated on the shortest paths established to destination satellites. The paths are composed of a list of codewords that characterize the list of intermediate nodes on the shortest path. Thus, a header that contains the path to a destination satellite is associated to each packet.

The design of the multicast module implemented in each satellite is illustrated in Figure 7. Therefore, an intermediate node that receives a multicast packet examines the received header optically by considering the following steps:

- the first codeword in the header, which indicates the current satellite, is extracted from the received list of codewords that compose the header and dropped;
- the packet and the new header are delayed in a VOM

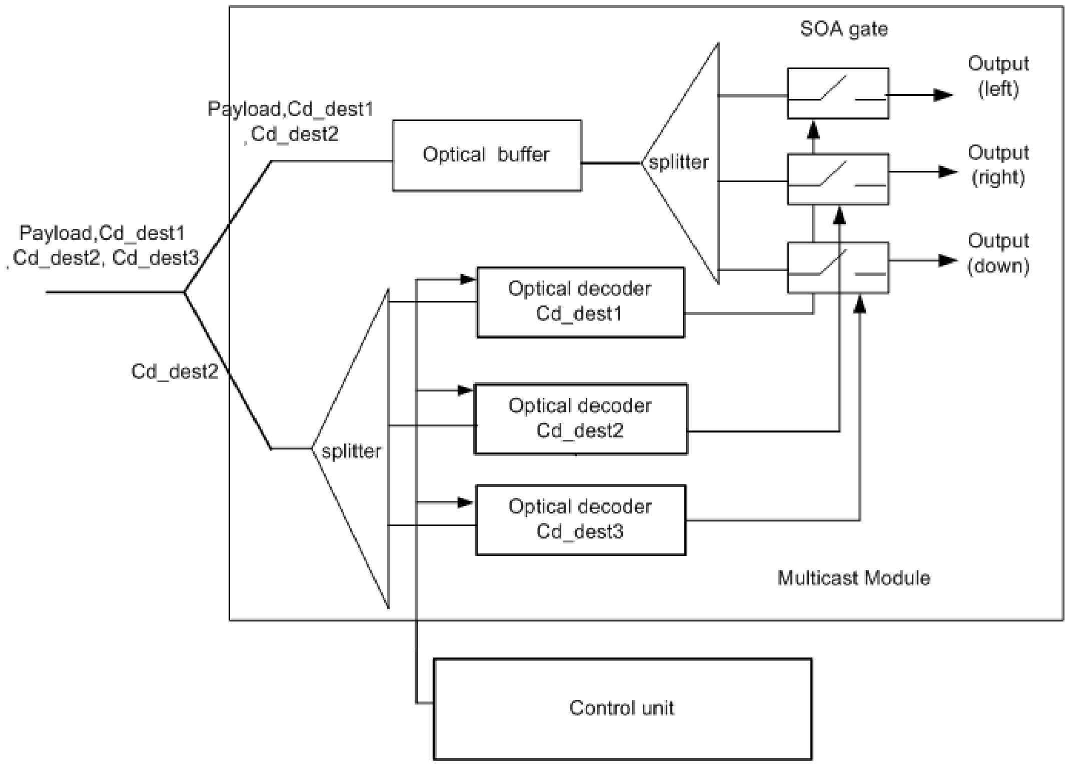


Figure 7. All optical multicast based on the shortest paths (Figure 3, [1]).

based on optical delay lines that is developed in [12];

- during the buffering delay, the second codeword in the received header is extracted from the received list of codewords and split to foure tunable decoders. Each decoder allows to match a codeword that characterize one neighbor of the current satellite; and
- if the second codeword matches a configured codeword, then the delayed packet and its corresponding header will be sent to the adequate next neighbor that allows to reach the destination satellite.

The multicast approach based on the shortest paths is more adapted to a payload with small size, which corresponds to a packet. Thus, the proposed multicast process requires a buffering delay of the payload in the optical buffer in order to optically switch the received payload to the next satellite in the shortest path in order to reach the destination satellite. Consequently, the buffering delay must not be more important than the size of the payload in order to have an optimized utilization of the bandwidth. In the proposed approach, the required buffering delay is simply the delay required to perform the matching process of one the four codewords corresponding to the direct neighbors of the current satellite.

VI. LEOCast protocol: All optical multicast mode based on multicast trees

In this section, we present an optical multicast approach based on the concept of virtual multicast trees. The virtual multicast paradigm is used in order to underline that the multicast tree is not physically established but only built on the codewords structure used to switch traffic contrary to the first proposed approach that requires a route discovery process preceding the routing of a multicast traffic. In this approach, the virtual tree establishment consists on the management of codewords structure. Thus, an optical codeword is structured in two parts as follows: the first part identifies the orbit and the direction (front or backward) conforming to the direction of the satellite movement in this orbit, and the second part identifies uniquely a satellite in the LEO constellation network. Therefore, the destination satellite can be either on the front or on the backward of an intermediate satellite or it can be localized in another orbit. The source satellite sends a multicast packet composed of the payload and a list of codewords that corresponds to the list of destination satellites. At the reception of a multicast traffic, an intermediate satellite can forward the traffic to the front or backward if the destination satellite is in its orbit or it switches the traffic to the following orbit.

In the following example, we consider the virtual tree example illustrated in Figure 8. Therefore, from a source satellite, the header to be sent, which is composed of a set of codewords, has the following structure: $CdOrbit_f^2$ CdD_1 CdD_3 — $CdOrbit_b^2$ CdD_2— $CdOrbit_b^3$ CdD_4 — $CdOrbit_f^3$ CdD_5, where $CdOrbit_f^i$ is the codeword that identifies the front direction in the $Orbit^i$ conforming to the satellite movement, $CdOrbit_b^i$ is the codeword that identifies the backward direction in the $Orbit^i$ conforming to the satellite movement, and CdD_j is the codeword that identifies the destination satellite D_j in the LEO constellation network. Thus, each optical codeword identifying an orbit is succeeded by a set of optical codewords identifying destination satellites in this

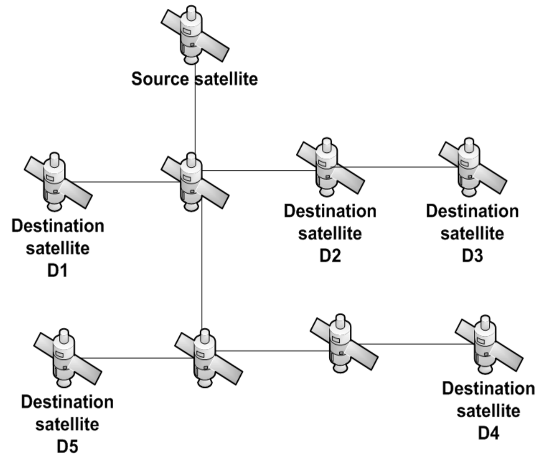


Figure 8. Virtual tree example (Figure 4, [1]).

orbit in a specific direction, front or backward (depending on the direction of the satellite movement in the orbit).

Each intermediate satellite manage a multicast traffic arriving in the same orbit (inter-orbit traffic) and from different orbits (intra-orbit traffic). This is achieved by considering a multicast module, which design is described in Figure 9. The multicast module is mainly composed of a set of optical buffers, tunable decoders and a control unit. The optical buffers delay the payload and the header during the matching process. The tunable decoders are configured by the control unit at the establishment of the multicast connection with the adequate codewords in order to send the multicast payloads to the destination satellites, which can be in the same orbit as the intermediate satellite or in the next orbits. At the reception of a payload and its corresponding header, which presents the multicast tree, the optical multicast process will be performed as follows:

- Different headers are extracted of the received header in order to obtain a separate header for each multicast direction (front, backward, next orbit conforming to the satellite movement in the orbit) and delayed with the received payload in different optical buffers. The considered optical buffer is called a VOM [12].
- The header, which corresponds to a list of structured codewords, is split in order to sequentially treat the codewords by the multicast module. The latter is composed of a set of tunable decoders configured with the following codewords by the control unit:
  - $CdOrbit^{current}CdD_{sat_dest}$, which indicates that the current satellite is a destination satellite.
  - $CdOrbit_f^{current}CdD_{sat_id}$, which indicates that the sat_id situated in front of the current satellite is a destination satellite.
  - $CdOrbit_b^{current}CdD_{sat_id}$, which indicates that the sat_id situated in backward of the current satellite is a destination satellite.
- A matching process is performed by the set of configured decoders on the received codewords that compose the header.
- Based on the result of the matching operation, the received payload will be multicast to one or several directions: front of the current satellite, backward of the current satellite and down to reach next orbits. Indeed, if $CdOrbit^{current}CdD_{sat_dest}$ is matched then the delayed payload will be treated. If $CdOrbit_f^{current}$ is matched then an optical switching gate is activated in order to send the delayed payload and its corresponding header to the neighbor satellite in front of the current satellite in order to reach destination satellites in this direction. The new header associated to the payload is composed only of codewords relative to destination satellites in front of the current satellite. If $CdOrbit_b^{current}CdD_{sat_id}$ is matched then the delayed payload and its corresponding header will be sent to the neighbor satellite in backward of the current satellite. If none of the configured codewords is matched, then a threshold detector commands an optical switching gate in order to send the delayed payload and its corresponding header to the neighbor satellite situated in the next orbit in order to reach destination satellites in other orbits.
  In our case, we need a high speed optical switching gate with a switching time window in order to get out the delayed signal from the VOM, which can be achieved for example by a Semiconductor Optical Amplifier gate (SOA). One of the most desirable properties of the considered SOA gate is the fast switching speed. Depending of the type of the SOA gate and the key temporal parameters of the SOA transit time, we obtain different switching window widths [26]–[28] as illustrated in Figure 10. For high-speed processing, short switching window is used. In our case, we have a synchronization issue that must be considered. Indeed, the switching window width T_{soa} must be sufficient to extract the header relative to a direction T_e and to get a copy of the delayed optical signal in the VOM T_s. Thus, in order to satisfy the synchronization constraint, we define the following relation: $T_s + T_e < T_{soa}$.

The multicast approach based on the virtual multicast tree

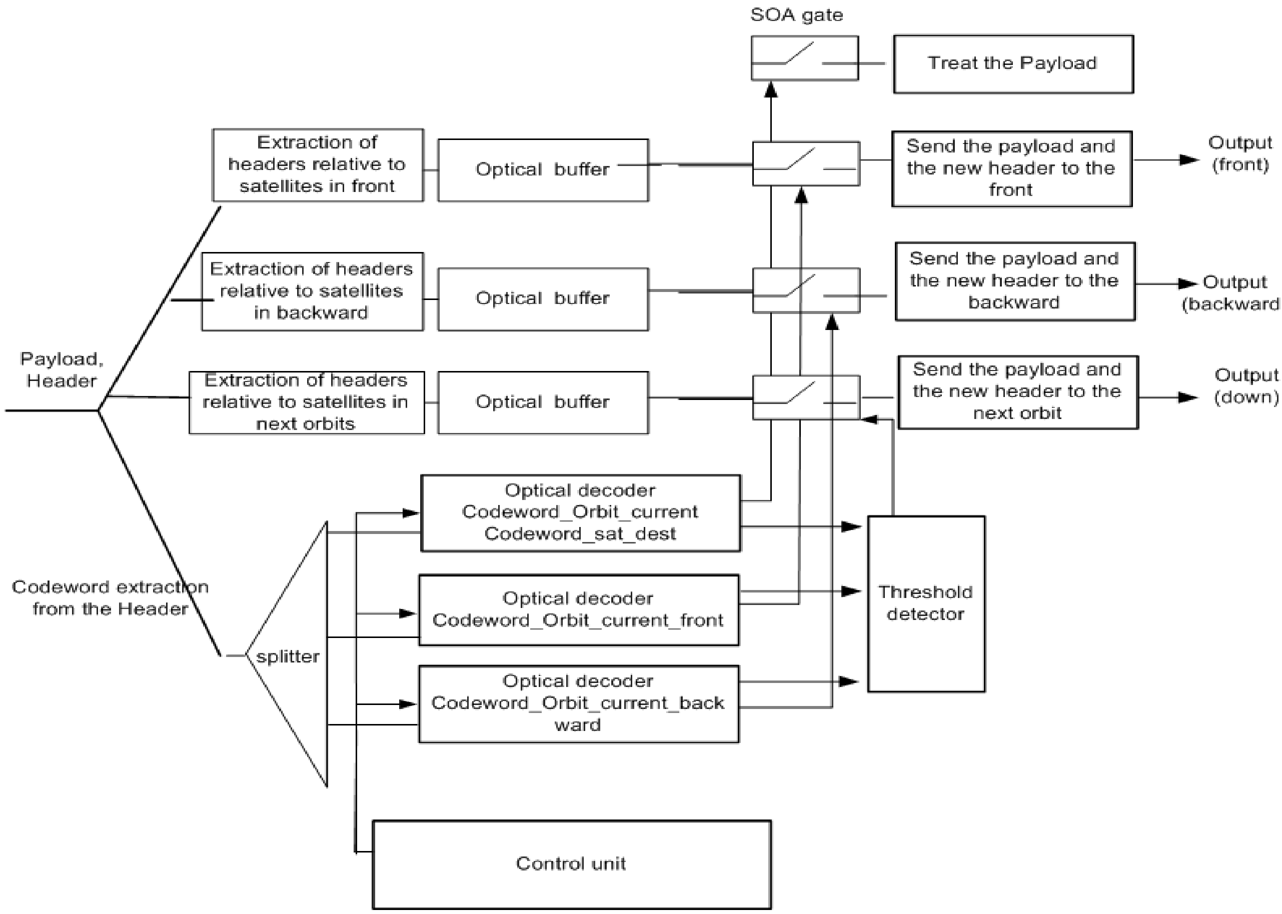


Figure 9. All optical multicast based on the Virtual multicast tree.

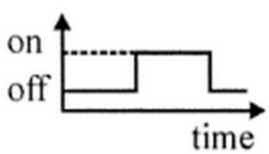


Figure 10. Switching window.

paradigm is adapted to a payload with important size, which corresponds to a burst (association of packets). Thus, the proposed multicast process requires a buffering delay of the payload in the optical buffer. This delay corresponds to the treatment of the set of codewords in the header sequentially and the generation of the new header to each burst to be switched to the adequate direction in order to reach destination satellites.

VII. Multicast coping with long period services

In this section, we study the mobility management in the two all optical multicast modes of the proposed LEOCast protocol.

A. Mobility management

Satellite movement results in challenging mobility management problems in LEO satellite networks. Due to the movements of the satellites and according to the movements of their coverage area and footprints, a group of destination users are served by several groups of satellites during a communication period that may lasts for hours, when transmitting voice/data/video to users in the earth. In this section, we present the mobility management for the two proposed optical multicasting approaches. For the two approaches, three types of handovers can occur during a multicast session period T. The first type of handover occurs when the source satellite moves and will not serve the users source of the traffic anymore. The second type of handover occurs when a destination satellite will not serve any destination user therefore it will be deleted from the group of the destination satellites. The third type of handover occurs when a destination satellite is added to the group of the destination users in order to serve one or several destination users.

1) Shortest paths management: The optical multicasting based on the Shortest paths requires the initiation of a novel Route request discovery process to the destination satellites when a source satellite is not in the coverage of source users anymore. The ground station must send the current list of codewords that indicate the current list of destination satellites, to the new source satellite. In the case where a destination satellite will not serve any destination user, it will be deleted from the group of destination satellites. This type of mobility does not require any treatment in the intermediate satellites. In fact, the multicast traffic to be switched by the intermediate satellites will no longer be switched in the direction of the deleted satellite since it is not in the coverage of the destination users. An updated list of codewords that corresponds to the new list of destination satellites, will be sent to the source satellite from the ground station. In the case where a destination satellite is added to the group of destination users in order to serve one or several destination users, the shortest path from the source to the new destination satellite must be established. An updated list of codewords will be sent to the source satellite from the ground station.

2) Virtual tree management: When a source satellite is not in the coverage of source users anymore, the optical multicasting based on the virtual tree concept does not require any route reestablishment process initiation to discover the routes to the destination satellites as it is the case in the optical multicasting based on the Shortest paths. Indeed, the multicast tree will be implicitly established by sending a multicast traffic that requires at each intermediate satellite to be optically switched. The ground station must send the current list of codewords that indicate the current list of destination satellites, to the new source satellite. In the case where a destination satellite will not serve any destination user, therefore, it will be deleted from the group of destination satellites. The ground station must send the current list of codewords to the new source satellite. And, the source replaces the codeword corresponding to the deleted destination satellite by the codeword associated to the new destination satellite and adds to the codeword, a special codeword that identifies the destination orbit and the direction (front or backward) in the orbit. The new codeword will be added to the header of the multicast traffic. In the case where a destination satellite is added to the group of the destination users in order to serve one or several destination users, its corresponding codeword will be added to the header of the multicast traffic.

3) Comparison: First, the two proposed optical multicasting approaches are compared in term of the number of used segments. The segments used by the approach based on virtual trees are those in the tree axis and in the front and backward of the axis. And the segments used by the approach based on the shortest paths are the total segments that compose the shortest paths to the destination satellites. Thus, the number of segments used in the first approach avoids the segments redundancy and consequently, it is smaller than the number of segments used in the second approach.

Second, the virtual tree construction favors the inter-orbit over the intra-orbit hop, which allows to have only one possible path to the destination satellite. This method is also considered in order to establish the shortest paths in the second approach. Indeed, the association of all established shortest paths gives the constructed virtual tree in the first approach.

Third, in the approach based on the shortest paths, when a satellite leaf handover occurs in the direction of satellites movement, a segment is removed of the path and the codeword of the removed destination satellite is removed of the list of destination satellites. In the contrary case, a segment will be added to the path and the codeword of the new destination satellite is added to the list of destination satellites. Thus, the handover in this approach consists on the increase or the narrowing of paths. When we consider the approach based on virtual trees, the destination satellites can be either a leaf or an intermediate satellite, which minimizes the increase or the narrowing of paths due to handovers compared to the first approach.

VIII. Simulations and Experimental Results

In this section, we assess the performance of the two proposed optical multicasting approaches in terms of traffic load (in erlang) and number of decoding functions. As simulation environment, we have considered Matlab tool. For simulation purpose, we consider the GE Starsys constellation topology [29], which forms a LEO network (6 orbits, each orbit has 4 satellites), a traffic load matrix. The traffic load matrix considers a traffic load only between each satellite and its four neighbors (up and down, front and backward). For each simulation, a source satellite number and destination satellite numbers are randomly generated.

In order to compare the efficiency of the two multicasting approaches in terms of the traffic load, we calculate the traffic load for each approach by considering the maximal path length/multicast tree depth and the mean on all shortest paths/segments of the tree. As illustrated in Figure 11, we can notice the similarity between the traffic load curves computed on the maximal path length and the depth of the multicast tree and considered for several multicast group sizes (4,8,12,16,20,24).

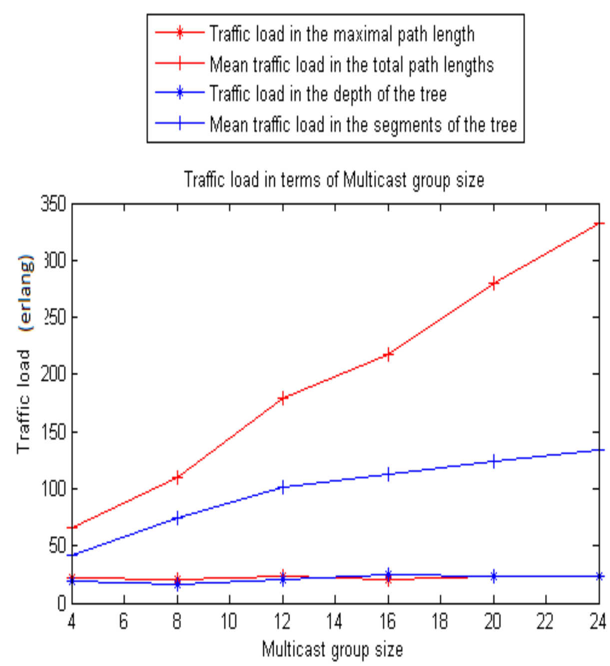


Figure 11. Traffic load in terms of multicast group size (Figure 6, [1]).

This similarity can be justified by the fact that the maximum path length corresponds to the depth of the multicast tree. We notice also that the mean traffic load on the total segments of the multicast tree is the half for a multicast group size equals to 24 satellites, which is optimized compared to the mean traffic load on the total path lengths. Therefore, we can deduce that the all optical multicasting approach based on the virtual multicast tree has the advantage to eliminate the redundant segments in the shortest paths established in the all optical multicasting approach based on the shortest paths.

The efficiency of the two multicasting approaches in terms of the number of used decoding functions is also assessed. We have considered a communication period of one hour between a source and destination users. During this period and due to the mobility of satellites, four source satellites and four multicast groups are considered. We calculate the number of used decoding functions for: 1) different multicast group sizes (4,5,6,7,8,9,10), 2) three path lengths less or equal to 1,2,3 hops. As illustrated in Figure 12, we notice that the number of used decoding functions for optical multicasting approach based on the shortest paths is five times greater than the number of used decoding functions for optical multicasting approach based on the virtual multicast tree for a multicast group size equals to 10 satellites. This can be the fact that the virtual multicast tree eliminates segments redundancy observed in the shortest paths approach.

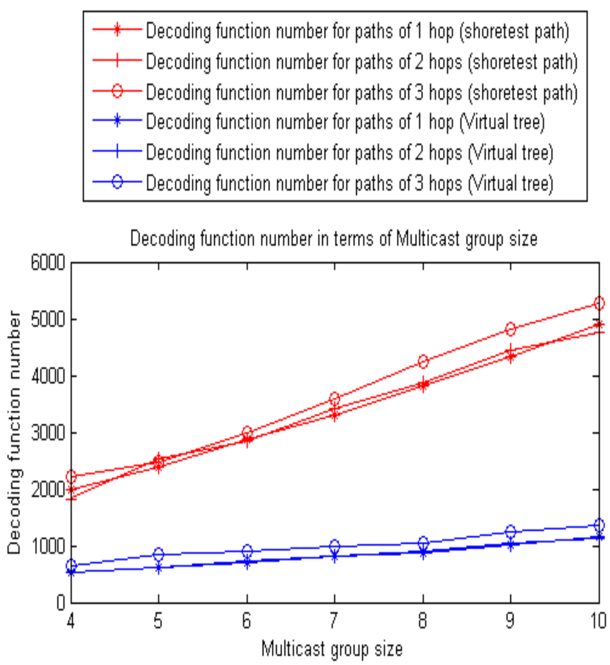


Figure 12. Decoding function number in terms of multicast group size (Figure 7, [1]).

Consequently, the optical multicasting approach based on the virtual tree decreases the traffic load observed in the network segments and uses less decoding functions for the optical multicasting. Thus, this approach is more efficient than the approach based on the shortest paths.

In Figure 13, we assess the optical signal quality after the multicast process to three receivers, which correspond to the three direct neighbors of the transmitter. To this objective, we consider Optiwave Optisystem as a simulation platform. We

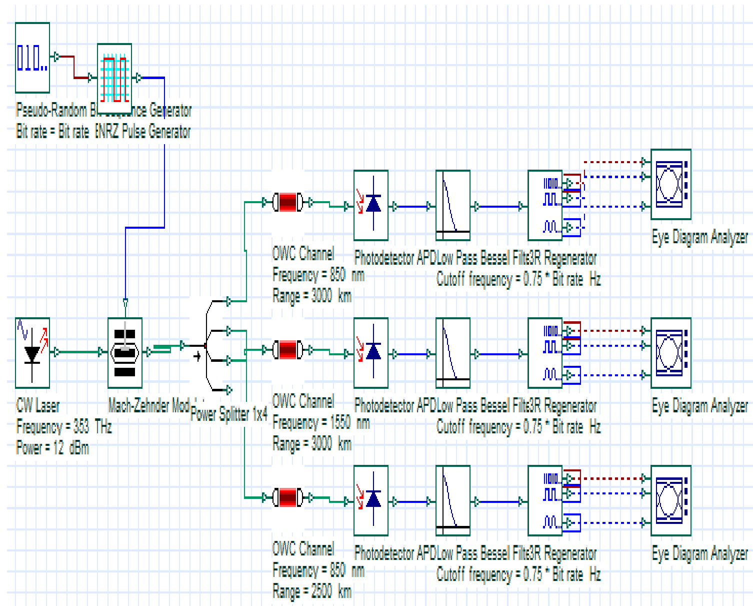


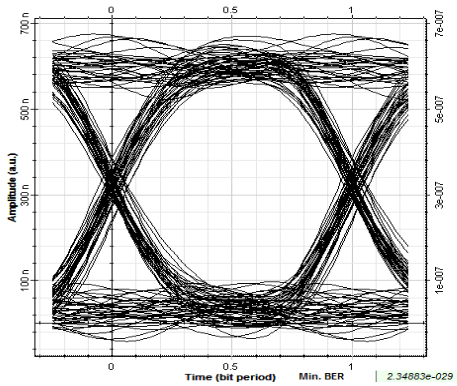
Figure 13. FSO link application in LEO networks.

consider a transmitter with a frequency of 353 THz and a power of 12 dBm, and we choose an Avalanche Photo Diode (APD) detector receiver, which is commonly used as FSO receiver [30], [31]. The optical signal quality is assessed based on two main criteria: the wavelength and the distance between satellites.

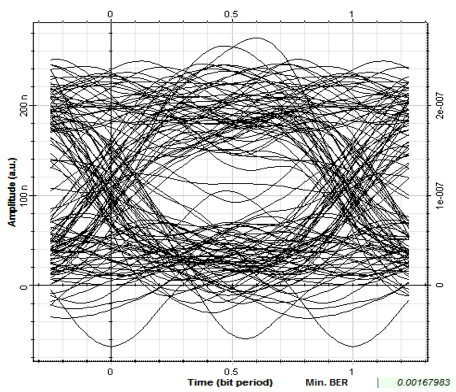
The eye diagram and the bit error rate are used as performance estimators in this simulation. Figure 14 describes the eye diagrams obtained when observing the received signal at the output of each receiver. The first eye diagram corresponds an ISL as an optical link with 850 nm and 3000 km. The second eye diagram corresponds to an ISL optical link with 1550 nm and 3000 km. The third eye diagram corresponds to an ISL optical link with 850 nm and 2500 km. Distortions shown on the eye diagram demonstrate that the best signal quality is obtained for the third receiver and the worst signal quality is obtained for the second receiver. Furthermore, the bit error rate obtained for the third receiver is lower than 10^{-9}, which is considered as an acceptable Bit Error Rate (BER) threshold, while the BER for the second receiver is greater than 10^{-9}. We can conclude that the optical signal quality increases when decreasing the FSO communication link wavelength and the distance between receivers when performing the multicast process optically. Furthermore, 1550nm lasers transmit more power than 850nm lasers for eye safety reasons (i.e. more power can be transmitted to overcome attenuation by aerosols). However, detectors in the 1550nm are typically less sensitive and have a smaller receive surface area when compared to silicon APD detectors that operate in the 850 nm wavelength.

IX. Conclusion

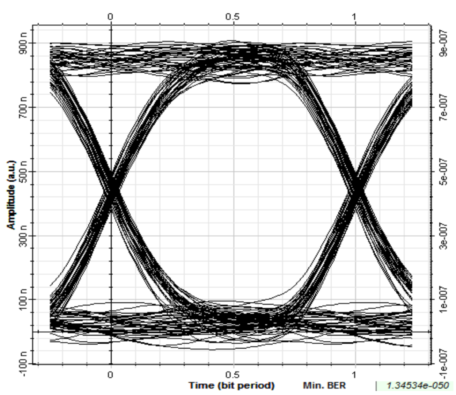
In this work, we propose an all optical multicast protocol LEOCast based on an optical switching technique that allows to perform the multicast of received traffic streams based on their optical codewords. Furthermore, this optical switching



(a) Wavelength:850nm, Distance:3000km



(b) Wavelength:1550nm, Distance:3000km



(c) Wavelength:850nm, Distance:2500km

Figure 14. Eye diagrams related to three receivers.

technique allows to perform several other functions with an optimized resource utilization. However, it is possible to use FSO links and consider an other switching technique with assuming the optical-electronic-optical conversion on intermediate satellites.

The proposed LEOCast protocol offers two multicast modes: the shortest path multicast mode and the virtual multicast tree mode. Each multicast mode is adapted depends on the buffer The multicast approach based on the shortest paths is more adapted to a payload with small size, which corresponds to a packet. Thus, the proposed multicast process requires a buffering delay of the payload in the optical buffer in order to optically switch the received payload to the next satellite. In this mode, the required buffering delay is simply the delay required to perform the matching process of one of the four codewords corresponding to the direct neighbors of the current satellite. The multicast approach based on the virtual multicast tree paradigm is adapted to a payload with important size, which corresponds to a burst (association of packets). This is due to the fact that this mode requires a more important buffering delay compared to the first mode. This latter corresponds to the treatment of the set of codewords in the header sequentially and the generation of the new header to each burst to be switched to the adequate direction in order to reach destination satellites.

REFERENCES

[1] M. Sliti, W. Abdallah, and N. Boudriga, "Optical multicast protocol for leo satellite networks," in *Proceedings of the 6th International Conference on Advances in Satellite and Space Communications (SPACOMM 2014)*, (Nice, France), pp. 29–35, 23 February 2014.

[2] S. Deering, "RFC 1112: Host extensions for IP multicasting," tech. rep., IETF, 1989.

[3] S. Deering and D. Cheriton, "Multicast routing in datagram internetworks and extended LANs," *ACM Transactions on Computer Systems (TOCS)*, vol. 8, pp. 85–110, May 1990.

[4] D. Waitzman, C. Partridge, and S. E. Deering, "RFC 1075: Distance vector multicast routing protocol," tech. rep., IETF, 1988.

[5] J. Moy, "Multicast routing extensions to OSPF," tech. rep., IETF, 1994.

[6] L. Cheng, J. Zhang, and K. Liu, "Core-based shared tree multicast routing algorithms for LEO satellite IP networks," *Chinese Journal of Aeronautics*, vol. 20, pp. 353–361, August 2007.

[7] Z. Yang, F. Long, F. Sun, and D. Wang, "A dynamic QoS routing mechanism based on steiner trees for LEO satellite networks," in *Proceedings of the International Conference on Networking, Sensing and Control (ICNSC)*, (Chicago), pp. 153–158, 10-12 April 2010.

[8] T. Mao, "A multicast routing algorithm for LEO satellite networks," in *Proceedings of International Conference on Future Computer and Communication*, (Wuhan), pp. 94–96, 6-7 June 2009.

[9] Z. Yin, L. Zhang, X. Zhou, P. Xu, and Y. Deng, "QoS-guaranteed secure multicast routing protocol for satellite IP networks using hierarchical architecture," *International Journal of Communications, Network and System Sciences*, vol. 3, no. 4, pp. 355–363, 2010.

[10] C. Yuan and X. Wang, "A multicast routing algorithm for GEO/LEO satellite IP networks," in *Proceedings of the International Conference on Dependable, Autonomic and Secure Computing (DASC)*, (Chengdu), pp. 595 – 599, 21-22 December 2013.

[11] Y. Ma, J. Su, C. Wu, X. Wang, W. Yu, B. Zhao, and X. Hu, "A source-based share-tree like multicast routing in satellite constellation networks," in *Proceedings of the International Conference on Mobile, Ubiquitous, and Intelligent Computing (MUSIC)*, (Vancouver), pp. 240 – 245, 26 - 28 June 2012.

[12] S. Batti, M. Zghal, and N. Boudriga, "New all-optical switching node including virtual memory and synchronizer," *Journal of Networks*, vol. 5, pp. 165–179, February 2010.

[13] X. Wu, P. Liu, and M. Matsumoto, "A study on atmospheric turbulence effects in full-optical free-space communication systems," in *Proceedings of the International Conference on Wireless Communications Networking and Mobile Computing (WiCOM)*, (Chengdu), pp. 1–5, 23-25 September 2010.

[14] Y. Koishi, Y. Suzuki, T. Takahashi, I. Mase, M. Jibiki, Y. Hashimoto, S. Murata, T. Yamashita, and K. Shiratama, "Research and development of 40gbps optical free space communication from satellite/airplane," in *Proceedings of the International Conference on Space Optical Systems and Applications (ICSOS)*, (Santa Monica), pp. 88 – 92, 11-13 May 2011.

[15] M. Rahman, S. Iqbal, and M. Islam, "Modeling and performance analysis of free space optical communication system," in *Proceedings of the International Conference on Informatics, Electronics & Vision (ICIEV)*, (Dhaka), pp. 211 – 218, 18-19 May 2012.

[16] W. Kmiecik and K. Walkowiak, "Survivable overlay multicasting in WDM optical networks with dual homing architecture," in *Proceedings of the International Conference on Optical Network Design and Modeling*, (Stockholm), pp. 19 – 24, 19-22 May 2014.

[17] D. D. Le, M. Molnr, and J. Palaysi, "Multicast routing in WDM networks without splitters," *IEEE Communications Magazine*, vol. 52, pp. 158–167, July 2014.

[18] R. Abbas, A. Masrub, S. Amin, and H. Al-Raweshidy, "High QoS and low cost wavelength allocation protocol in optical network using multicast technology," in *Proceedings of the International Conference on Innovative Computing Technology (INTECH)*, (London), pp. 356 – 359, 29-31 August 2013.

[19] H. Liu, T. Hu, Y. Chen, Q. Chen, and Y. Xie, "Optimisation of layer rate and wavelength allocation based on network coding for multirate optical multicast," *IET Communications*, vol. 8, pp. 1570 – 1577, June 2014.

[20] H. Ghafouri-Shiraz and M. M. Karbassian, *Optical CDMA Networks: Principles, Analysis and Applications*. Wiley-IEEE Press, 2012.

[21] Y. Li, J. Gao, Q. Li, and J. Liu, "A construction method of optical orthogonal code based on difference set," in *Proceedings of the International Conference on Network Infrastructure and Digital Content*, (Beijing), pp. 486 – 490, 6-8 November 2009.

[22] R. Chauhan, R. Kumar, M. Shukla, A. Sigh, and G. Bagaria, "Proposal for one dimensional optical orthogonal codes: Design, analysis, & algorithm," in *Proceedings of the International Conference on Communication Systems and Network Technologies (CSNT)*, (Katra, Jammu), pp. 514 – 519, 3-5 June 2011.

[23] D. Wu, H. Zhao, P. Fan, and S. Shinohara, "Optimal variable-weight optical orthogonal codes via difference packings," *IEEE Transactions on Information Theory*, vol. 56, pp. 4053 – 4060, August 2010.

[24] J.Salehi, H.Chung, and V.K.Wei, "Optical orthogonal codes: Design, analysis, and applications," *IEEE Transactions on Information theory*, vol. vol 35, pp. 595–605, May 1989.

[25] H. Yin and D. J. Richardson, *Optical Code Division Multiple Access Communication Networks Theory and Applications*. Springer Publications, 2007.

[26] Y. Guan and R. Wang, "Demonstration of an optical switch based on SOA-MZI operation at 10 gbit/s," in *Proceedings of the International Conference on Artificial Intelligence and Software Engineering*, (China), pp. 239–241, July 2013.

[27] R. Rostam, M. Wahid, S. Rais, and M. Faridus, "The effect of signal repetition rate, injected current and switching window on cross-phase modulation in SOA-NOLM," in *Proceedings of the IEEE Regional Symposium on Micro and Nanoelectronics*, (Malaysia), pp. 373 – 376, September 2011.

[28] H. N. Tan, M. Matsuura, and N. Kishi, "Parallel WDM regenerative waveform conversion for mixed NRZ and RZ transmission networks using a SOA-based multiple switching-window optical gate," in *Proceedings of the Optical Fiber Communication Conference and Exposition (OFC/NFOEC)*, (Los Angeles, USA), pp. 1–3, March 2011.

[29] M. Sturza, "LEOs - the communications satellites of the 21st century," in *Proceedings of Northcon/96*, (Seattle, Washington), pp. 114–118, 4-6 November 1996.

[30] K. Kiasaleh, "Performance of APD-based, PPM free-space optical communication systems in atmospheric turbulence," *IEEETransactions on Communications*, vol. 53, pp. 1455–1461, September 2005.

[31] P. Sun, M. Hayat, and A. Das, "Bit error rates for ultrafast APD based optical receivers: exact and large deviation based asymptotic approaches," *IEEE Transactions on Communications*, vol. 57, pp. 2763 – 2770, September 2009.

Fully Distributed Ubiquitous Information Sharing on a Global Scale for the Internet-of-Things

Victor Kardeby, Stefan Forsström, Patrik Österberg, and Ulf Jennehag

Department of Information and Communication Systems
Mid Sweden University
Sundsvall SE-85170, Sweden
Email: victor.kardeby@miun.se, stefan.forsstrom@miun.se, patrik.osterberg@miun.se, ulf.jennehag@miun.se

Abstract—**The Internet-of-Things will require ubiquitous information sharing between connected things on a global scale, which existing systems do not offer. Most current efforts focus on solutions for information dissemination, which induce single points of failure and introduce unnecessary communication delays. To this end we propose the SensibleThings platform, which is a fully distributed open source architecture for Internet-of-Things based applications. This article describes the major problems that Internet-of-Things platforms must address, our technical solution to these problems, and an evaluation thereof. We also present the current progress and a series of demonstrators, which show the wide range of applications enabled by the platform. Finally, we present how the platform will be used in our future research and potential spin off companies.**

Keywords–overlay;sensors;actuators;internet-of-things.

I. Introduction

This journal article is an extension of [1] and [2], where [1] is a best paper awarded publication at the 2014 International Conference on Digital Telecommunications (ICDT).

Today we can observe a large interest in applications that can utilize information from sensors attached to different things in order to provide more personalized, automatized, or even intelligent behavior. These are commonly referred to as Machine-to-Machine (M2M) applications [3] or Internet-of-Things (IoT) applications [4]. The IoT can be seen as a natural evolution of computer networking and communicating devices, from simple direct communication between computers, via globally connected computers, to small devices such as smartphones that are ubiquitously connected to the Internet. Together these form a worldwide network of interconnected everyday objects. Through the IoT, applications will display context-aware behavior [5] and even be able to have social interactions between themselves [6]. These applications may address a variety of areas, such as environmental monitoring (pollution, earth quake, flooding, forest fire), energy conservation (optimization), security (traffic, fire, surveillance), safety (health care, elderly care), and enhancement of social interactions. Furthermore, the IoT is surprisingly close to Mark Weiser's predictions made in 1991 on the computers of the 21st century [7].

There is also an interesting relationship between the IoT and big data [8], since all of the connected things will produce and consume large amounts of data. Current estimations are in the order of 50 billion connected devices year 2020 [9]. Thus, IoT applications will probably have a big impact on how we interact with people, things, and the entire world in the future. But in order to enable a widespread proliferation of IoT services there should be a common platform for dissemination of sensor and actuator information on a global scale. This is however a very difficult goal to achieve, because there is a large number of practical difficulties that must be solved. Therefore, the purpose of this article is to explore how to enable ubiquitous information sharing on a global scale for the IoT, using highly scalable fully distributed solutions. We present a realized solution called the SensibleThings platform, which is verified through a series of demonstrator applications and performance measurements.

The remainder of this article is outlined in the following way: Section II presents a list of requirements to evaluate potential platforms. Section III surveys current IoT architectures. Section IV presents our approach to address the requirements. Section V describes the SensibleThings platform which is our implementation of the approach. Section VI presents the verification of the platform and measurement results. Section VII discusses future IoT services. Finally, Section VIII presents the conclusion and future research.

II. Platform Requirements

To aid in evaluating IoT platforms, a list of application requirements has been constructed. This list is derived from previous and related work, for example, [2], [4], [10].

IoT applications spread very diverse areas, but common among many of them is the focus on many small devices on a global scale, low response times, reliable operation, and interoperability to support different hardware and features depending on the scenario. Therefore, we state that the majority of applications on the IoT will require the following from an underlying platform:

1) Capability of signaling between end points with low latency, without any unnecessary relaying of information.
2) Reliably handle transient nodes joining and leaving with high churn rates. Also avoiding choke points with significantly higher utilization than the rest of the system.

3) Ability to run on devices with limited computational and data-storage capacity.
4) Be extensive and adaptive to conform with a wide range of applications, devices, and future scenarios with currently unforeseen demands.
5) Easy to adopt and free to use in commercial products without restrictions in terms of software licenses and fees.

Some requirements have been left out of scope in this article because we focus on the open sharing of information where problems regarding security and privacy can be addressed at a later stage. In order to evaluate if a platform achieves our requirements we have also determined the following concrete and measurable metrics.

A. Evaluating Requirement 1

The first requirement on low response times cannot be measured using only the raw response times between the source and sink because they highly depend on the infrastructure in between. Since all IoT platforms are based on the Internet, all communication is inherently made using a best effort system. This makes it impossible to compare the raw response times from two different platforms. Hence, we evaluate the response time based on the ratio between the retrieval time compared to an ideal communication case. The ratio is calculated in (1), where $R_{latency}$ is the average ratio between the retrieval time $t^{retrieve}$ and the ideal round trip time t^{rtt} for each device i of the total devices N. Where $t^{retrieve}$ is the measured time to sends a sensor value from source to sink and t^{rtt} is the measured ideal round trip time, in this case ping.

$$R_{latency} = \frac{1}{N}\sum_{i=1}^{N}\frac{t_i^{retrieve}}{t_i^{rtt}} \tag{1}$$

B. Evaluating Requirement 2

The second requirement on reliability and scalability can be measured by the communication load on all ingoing devices. To calculate this we need to find the function $f_{avg}(N)$ in (2), of how the number of messages m handled per device per minute i changes as the number of devices N increase in the system. However, for most platforms the average number of messages per node and minute will logarithmically increase and converge to a constant.

$$f_{avg}(N) = \frac{1}{N}\sum_{i=1}^{N} m_i \tag{2}$$

The standard deviation of the amount of messages per device is also interesting for the scalability metric, since it can indicate central points of failure. The standard deviation as a function of the total number of devices N can be seen in (3). A constant or decreasing standard deviation function indicates an even system, where the new devices handle an even share of the workload, whereas an increasing standard deviation means that there is an uneven workload among the devices.

$$f_{stdev}(N) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(m_i - f_{avg}(N))^2} \tag{3}$$

C. Evaluating Requirement 3

The third requirement on resource efficiency can be measured in different aspects, such as energy consumption, processing time requirements, storage space requirements, and network load. The resource efficiency is difficult to calculate for a whole system, because it highly depends on the hardware present on the devices. Therefore, we define the resource efficiency as the maximal value U in the set of all resource utilizations u for different available resources j when the platform is running. See (4) and (5). The resource utilization for each resource is calculated by the amount of a resource used by the platform u_{used} divided by the available amount of the resource u_{avail}. If any resource utilization on any device in the system is close to 100%, the resources have been saturated and the system is no more resource efficient.

$$u = \frac{u_{used}}{u_{avail}} \tag{4}$$

$$U = max(\{u_1, u_2, ..., u_j\}) \tag{5}$$

D. Evaluating Requirement 4 and 5

The fourth and fifth requirement cannot be measured, only qualitatively evaluated. To be future proof the platform must have the ability to add new features without costly updating all ingoing devices individually. To study the cost, one has investigated the terms of usage for the platform to evaluate the type of costs that are related to the platform. These costs can be either traditional fees or resources such as bandwidth and computational power.

III. Related Work

There currently exist a vast amount of platforms which claim to enable an IoT, far more than can be listed in this paper. There are also standardization efforts being made by standardization bodies such as the Internet Engineering Task Force (IETF), the European Telecommunications Standards Institute (ETSI), and the Institute of Electrical and Electronics Engineers (IEEE). In general, most of these are focused on problems such as enabling Internet Protocol (IP) over radio communication and connecting different hardware to the Internet. Therefore, the standardization efforts are much less extensive in the layers closer to the application, which is what this article focuses on. As an example, the IETF are discussing standardization from the physical and Media Access Control (MAC) layer, covering the Internet Protocol Version 6 (IPv6) adaptation layer, the routing layer, and finally an application protocol [11]. The application protocol that they propose is, however, the Constrained Application Protocol (CoAP) which only provides simple transfer of sensor and actuator information directly between two known IP addresses. Meanwhile, the IEEE IoT standard P2413 [12] is still in its early stages, since their first working group meeting was in July 2014. Hence, no real results have so far been presented other than that the plan is to utilize existing technologies for the lower layer communication and that they are investigating which application level protocols to use in their standard.

In addition to the standardization efforts, there are multiple forums and alliances made of industry and academia participants to both promote the idea of an IoT and to establish

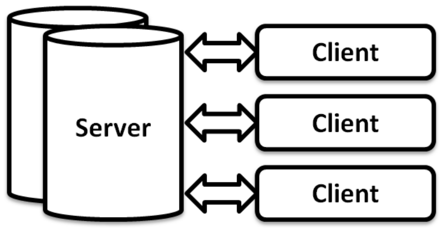


Figure 1. Overview of typical centralized architectures

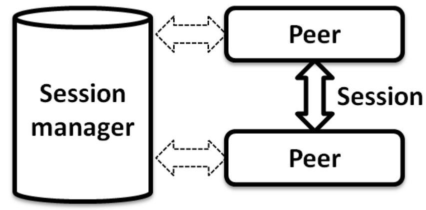


Figure 2. Overview of typical semi-distributed architectures

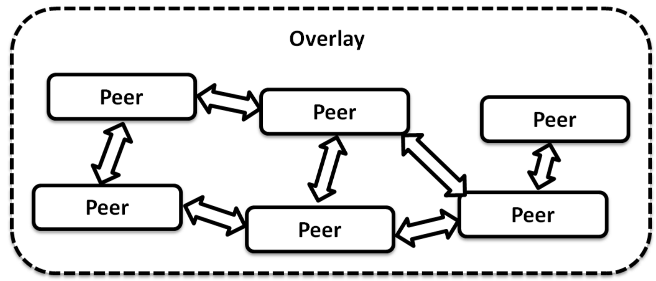


Figure 3. Overview of typical fully distributed architectures

some form of praxis and consensus. These include for example the AllSeen Alliance [13], European Research Cluster on the Internet of Things [14], HyperCat [15], Internet of Things – Architecture [16], the Internet of Things Council [17], the Internet of Things Europe [18], The Internet of Things Initiative [19], the IoT Forum [20], IP for Smart Objects Alliance [21], oneM2M [22], and Open interconnect [23].

Additionally, there are research [24] that investigates the possibilities and challenges for using software defined networking (SDN) [25] as an IoT platform, where [24] concludes that SDN will provide a unified method of managing network resources globally. In [26] the authors implements a vertical SDN controller that is used in conjunction with advanced network calculus- and genetic algorithm-techniques in order to improve data throughput, end-to-end delay and end-to-end jitter.

The many platforms found in related work can generally be categorized into three groups, centralized (or cloud distributed), semi-distributed, and fully distributed systems.

A. Centralized Systems

Most of the systems being released today focus on distributing the data on some form of cloud-based IoT architecture. The cloud is a concept that rise in popularity, but it is in many cases simply a new word for traditional web services. Very few commercial cloud-based systems explain how the distribution and synchronization is actually done inside their architecture, which type of virtualization they use, how many servers they have, etc. Either way, cloud-based systems can be considered as centralized systems since they always relay the sensor and actuator information through a centralized point, in this case a cloud (be it one or many connected servers). See Figure 1 for an overview of a centralized system. The main problems with these solutions are that they have difficulties achieving requirement 1 on direct communication between end devices, requirement 2 on no central points of failure, and requirement 5 on an open and free to use system, because they are based on large scale servers. Typical examples of these centralized or cloud-based architectures include: SicsthSense [27], ThingSpeak [28], Sen.Se [29], Nimbits [30], ThingSquare [31], EVRYTHNG [32], Paraimpu [33], Xively [34], XOBXOB [35], Thingworx [36], One Platform [37], Carriots [38], OpenIOT [39], SAP Internet-of-Things [40], and many more.

B. Semi-Distributed Systems

The semi-distributed systems are often based on session initiation protocols, whereas they afterward use direct communication between the connected devices. See Figure 2 for an overview of a semi-distributed system. Because of this, they usually contain a centralized point for coordinating the communication. Thus, semi-distributed systems are faster and to some extent easier to scale than centralized solutions, but they still have difficulties coping with requirement 2 and 5. Typical examples of these semi-distributed architectures include: ETSI M2M [41], SENSEI [42], ADAMANTIUM [43], and other platforms based on 3GPP IMS [44].

C. Fully Distributed Systems

Fully distributed systems operate in a peer-to-peer manner, where clients both store and administer the information locally on each entity without centralized components, see Figure 3. To achieve this, they often utilize hash tables to enable logarithmic scaling when the number of entities increases in magnitude. These systems do not contain any single point of failure and are thus more resilient, though the distribution itself often requires additional overhead in order to maintain an overlay. The main problem associated with fully distributed systems is, however, that they place a larger responsibility on the end devices, and thus have difficult to achieve requirement 3. Examples of such systems are the Global Sensor Networks (GSN) [45], the RELOAD architecture [46], and MediaSense [2].

IV. TECHNICAL PLATFORM

Because the existing systems are unable to address the five requirements to enable a large scale IoT stated in Section II to enable a large scale IoT, a new solution is needed. Therefore, we have developed a platform based on how well the different categories achieve the requirements. An overview of this platform can be seen in Figure 4. In short, the platform is based on connecting sensor and actuators to form an IoT using current IP networking, fully distributed systems, peer-to-peer communication, and distributed hash tables (DHT). DHTs are distributed systems, which enable the storing of key-value pairs that can be utilized as a distributed storage service. A DHT

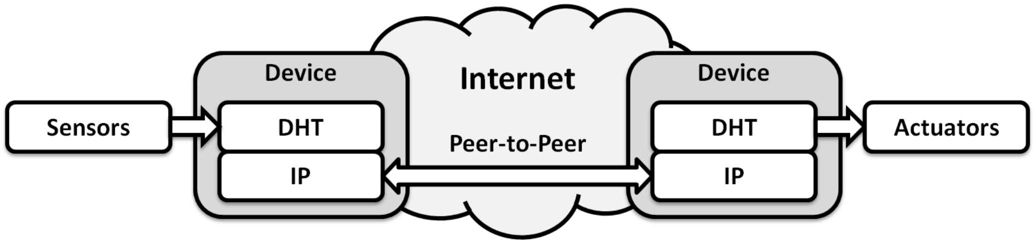


Figure 4. Overview of the proposed approach.

can therefore be seen as a form of DNS system but without the centralized servers. The DHT can be used to organize all connected entities in a logical structure which collaboratively stores data over distributed entities, where anyone can retrieve the data.

In detail, the platform should be IPv4 and IPv6 compatible which should make it future proof of many years to come, since the proliferation of the IPv6 standard is still not widespread. The platform should be scalable and avoid central points of failure by employing a fully distributed architecture, with no centralized points and a communication load that scales at worst logarithmically with the number of connected devices. The platform should employ peer-to-peer communication between the connected devices. As this type of communication is made directly between the data sources and the sinks without any unnecessary proxying of the data. Thus, only the normal Internet routing is added as delay, which is as optimal as one can achieve on the Internet. The platform should employ DHTs as an overlay network in order to become stable. Seamlessness in fully distributed systems is difficult to achieve, but there exists quite prominent NAT penetration techniques, which can be employed to solve most cases. Furthermore, constructing the ingoing protocols for lightweightness is a whole optimization problem in its own. But the idea is that the chosen DHT protocol should be lightweight by itself. Lastly, extensibility can be achieved by, for example, implementation specific solutions such as smart redistribution protocols, automatic updates, and distributed dynamic loading of components in runtime.

V. The SensibleThings Platform

In order to address the stated requirements we have created the SensibleThings platform, which is implemented as a layered architecture. A layered structure is chosen in order to more easily exchange modules as needed in both research and commercial applications. The layers and their default modules are explained in detail in the original article, but they are summarized here as well. The actual SensibleThings code is based on a fork of the MediaSense platform from 2013, which is a realization and implementation of the MediaSense architecture explained in [2].

A. Conceptual solution

The SensibleThings platform can be applied to a wide range of scenarios because of its versatile features. For example, how to perform low latency communication and finding information within the platform. This section presents these features and the corresponding conceptual solutions inside the platform.

1) Communication: Peer-to-peer communication is employed as the solution for the communication in the SensibleThings platform. The reasons behind this are many, but most important are the performance benefits. Peer-to-peer communication has an inherently low source to sink delay, because all communication is made with as few intermediate steps as possible. A fully distributed peer-to-peer system has no central points of failure and should thus be resilient to infrastructure failures. In a peer-to-peer system all participants provide their own network capacity, which is important because there is a significant investment involved in maintaining an IoT architecture. However, this is also a drawback since it causes vulnerability to denial of service attacks and an overlay must therefore be maintained to achieve good functionality and reliability. The are also many different transport layer protocols that could be utilized to send the peer-to-peer communication. Where one might strive for certain capabilities of the protocols, such as the flow control in TCP and the encryption of SSL. However, the conceptual idea in the SensibleThings platform is to enable the communication with as low overhead as possible, but still maintaining some type of reliability of the communication.

2) Finding information: The conceptual solution for finding information in the SensibleThings platform has been to employ a DHT. The SensibleThings platform associates the IP addresses of sensors with a Universal Context Identifier (UCI)[1] in the DHT. The UCIs can be seen as the unique identity of a sensor and looks like a combination between a URL and an E-mail address. To be specific, a UCI follows the structure of a Universal Resource Identifier according to RFC 3986[47] but we omit the scheme here to save space.

```
user@domain/path[?options]
```

Thus, a typical example a Celsius temperature sensors UCI owned by the user named Victor Kardeby is:

```
victor.kardeby@miun.se/temperature?unit=celsius
```

However, most DHT's only support the finding and resolving of an UCI to an IP address, they do not support searching for information. For example, the DHT can resolve a previously known UCI to an IP address, but it can not find the UCI's of sensors in a particular city. Hence, the conceptual solution for this is to employ distributed searching algorithms in the key space to enable intelligent searching of information and meta information in the system.

3) Publish and Subscribe: Most retrieval of information on the IoT is based on the publish and subscribe paradigm, hence the SensibleThings platform should support this. The conceptual solution for subscriptions in the Sensiblethings platform is to create a distributed subscription system where each device handle its own subscribers. Thus, notifying them whenever the sensor value is updated.

4) Security: The information on the IoT often originates from sensors, which have sensed their surroundings. Therefore, the information can possibly contain private information, which could be utilized by malicious users. Hence, there is a need for security in the SensibleThings platform, both in the form of encryption of the data and authentication of those who may access it. Security is out of scope, but our conceptual solution is to encrypt the peer-to-peer communication, which can be done in a number of different ways. But it is important to remember that the encryption should be done without the use of a centralized authority and as lightweight as possible, because of the devices with limited hardware resources. After the communication has been encrypted, the authentication problem can be addressed. However, authentication in the platform should also be constructed in a fully distributed manner with low overhead. Thus, some type of distributed authentication and trust system should be employed.

5) Mobility: Another big problem is the mobility and connectivity of devices. In the course of seconds a device might change its IP address multiple times as it switches between a home WiFi and 3G/4G connectivity when a person leaves a house. This spontaneous change of IP address makes traditional mobility services ineffective[48] due to frequent updates to third party support. The lookup system could become unreliable as it might have obsolete information in the DHT. The conceptual solution to this is to employ different fully distributed solutions to reduce the convergence time of the DHT's IP address records and to estimate a persons behavior in order to predict the change and update the DHT beforehand.

6) Persistence: To store large amounts of information in a fully distributed peer-to-peer system is difficult, since each node might have limited capabilities. Furthermore, because each node is responsible for its own sensors and is the only source of its information, they are responsible for the persistence of their own information. The conceptual solution to create persistence in the SensibleThings platform is to introduce cloud based persistence for storing of important information and as an offloading system when a node exceeds its ability to persist its data. The cloud based persistence should only be used as a backup for extremely vital information, which should never be lost, such as history of medical sensors.

7) Reasoning: Intelligent reasoning is one of the final purposes of connecting sensors to together, to make applications alter their behavior depending on the context of the users. In order to create intelligent reasoning, both large statistical data and machine learning techniques can be used. These require computationally heavy calculations, data mining, storage space, and other resources. Generally, intelligence and reasoning is solved on the application layer at the respective endpoint and thus is not a part of the platform itself. Therefore, an application only has access to data that the user would be authenticated to access. However, data on the platform should of course be opened up to support more ubiquitous types of data mining. The SensibleThings platform has an extensive add-in system, which can add functionality, such as intelligent reasoning, as needed in runtime. Different types of intelligent reasoning engines have been applied on the data from the platform. For example, creating intelligence from continually changing user profiles called context schemas, which are based on relevancy of the information and derived from the algorithms in [49]. Furthermore, different types of machine learning techniques have been applied to estimate the context of a person, based on information from the IoT [50].

8) Interoperability: There exists a wide range of different sensor devices, which have different capabilities and support different protocols, both open and proprietary. This makes it difficult to build a platform that can utilize all the different hardware and to create interoperability between them. However, in the latest years there seem to have arrived a consensus between the different Wireless Sensor Network operating systems, since both major operating systems TinyOS [51] and Contiki [52] support the Constrained Application Protocol (CoAP) [53]. Therefore, the SensibleThings platform also support sensors and actuators, which expose CoAP interfaces. Furthermore, as the SensibleThings platform is far from the only platform for enabling IoT applications, there is a need to create interoperability between platforms as well. The conceptual solution to this in the SensibleThings platform is to create add-ins that can bridge between different platform technologies. For example, create an add in for bridging sensors connected to different REST-based [54] cloud services.

B. Architecture Layers

The SensibleThings platform is divided into five different layers, which can be seen in Figure 5. The interface layer exposes the platform's Application Programming Interface (API) to the applications, the add-in layer makes it possible to extend the platform with additional functionality, the dissemination layer addresses finding and retrieving data from other entities, the networking layer handles the IP connectivity, and lastly the sensor and actuator layer that connects sensor and actuators. Thus, each layer focuses on specific problems and the remainder of this section will describe the layers and their implementation.

1) Interface Layer: The interface layer is the public interface that applications use to interact with the SensibleThings platform. It includes a single component, the SensibleThings application interface, which is a generic API for developers to build their own applications on top of. The main problem that the interface layer addresses is related to requirement 5 on easy usage, namely how to make the platform easy to understand and easy to implement applications with. Different approaches were explored, but since almost all communication on the platform is done asynchronously, the listener Java pattern is typically used in the application interface. Hence, almost all interface access with the platform is done through normal function calls, whereas the values are returned in event listeners. The sensor location address scheme is abstracted such that they are represented by objects acquired by resolving UCIs, as defined in Section V-A2. In short, the SensibleThings platform has the following basic interface:

- **SensibleThingsPlatform(Listener)** a constructor that joins the distributed system and sets the listeners to use

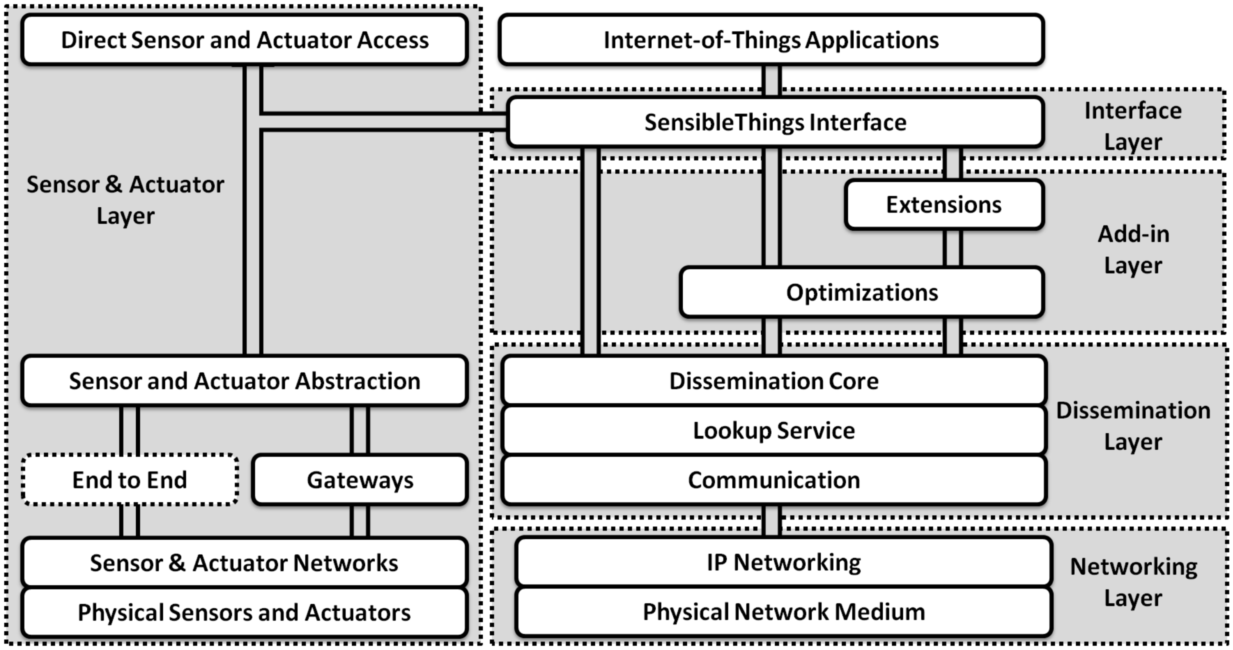


Figure 5. Overview of the SensibleThings platform's architecture

- **Register(UCI)** registers the specified UCI in the system
- **Resolve(UCI)** resolves a UCI to an node address
- **Get(UCI, Node)** retrieves the value of sensor, given its UCI and resolved node.
- **Set(UCI, Node, Value)** sets an actuator, given its UCI, resolved node, and the value to set
- **Shutdown()** performs a graceful leave from the distributed system

2) Add-in Layer: The add-in layer enables developers to add optional functionality and optimization algorithms to the platform beyond the basic primitives offered by the interface layer. For example, add-ins can help the platform meet specific application requirements, such as handling the available capacity in regards to computational power and bandwidth. The add-in layer deals with requirement 4 on being extensive. It manages different extensible and pluggable add-ins, which can be loaded and unloaded in runtime when needed. These add-ins are divided into optimization and extension components, but the platform can include any number of them at the same time. The add-in layer therefore handles how the add-ins should be managed, loaded, and the API's chain of command. In the current platform, there are still some limitations as these issues have not been prioritized. For example, some add-ins hijack functionality of the platform when enabled. Thus, some add-ins become mutually exclusive and will not function properly together. Examples of currently implemented add-ins are: caching, buffering, publish/subscribe, streaming, and password authentication.

3) Dissemination Layer: The dissemination layer enables sharing of information between all entities that participate in the system and are connected to the platform. A variant of the Distributed Context eXchange Protocol (DCXP) [55] is used, which offers communication among entities that have joined the distributed system, enabling exchange of context or sensor information with low response times. The operation of the DCXP includes first resolving of UCIs and subsequently transferring information directly between the peers. Therefore, the dissemination layer includes three components, a dissemination core, a lookup service, and a communication system. The dissemination core exposes the primitive functions provided by DCXP, the lookup service stores and resolves UCIs within the system, and the communication component abstracts transport layer communication. In short, the dissemination layer enables registration of sensors in the platform, resolving the location of a sensor in order to find it, and the communication to retrieve the actual sensor values.

The main design choices faced when developing the dissemination layer was regarding the choice of lookup service that supports requirements 1, 2, and 3, namely quick dissemination, being reliable with good scalability, and lightweight operation. There exists a number of DHTs that the platform could use, where we have chosen to focus on three prominent ones, namely Chord [56], Kelips [57], and P-Grid [58]. All three choices have their separate advantages and disadvantages. Chord uses a ring structure, which is difficult to maintain and has a logarithmic lookup time $O(log(N))$. Kelips uses affinity groups with a much simpler synchronization scheme and has fixed lookup time of $O(1)$, but it does not scale as well and has a larger overhead. P-Grid has a trie based structure with a logarithmic lookup time of $O(log(N))$ with load balancing and extra features such as range queries, but is also quite complex and difficult to maintain. In the current platform, both Chord and Kelips are completely reimplemented to operate within the platform and with the same license as the rest of the code. We have also experimented with the currently available P-Grid code, but since that is using multiple source code licenses (including propagating open source licenses), it cannot be a part of the SensibleThings code

TABLE I
SERIALIZATION METHOD COMPARISON.

Message Type	Binary	Java	Java compressed
Get	75 bytes	157 bytes	137 bytes
Kelips sync	856 bytes	2991 bytes	1252 bytes

at this stage. Currently, the platform defaults to the Kelips DHT implementation, simply because that code is more stable than the Chord implementation when nodes join and leave rapidly.

Another important design choice faced in the dissemination layer was the choice of communication protocol. Requirement 1 states that the communication should be fast with low overhead. Therefore, the aim was to have the useful payload data already in the first packet. Because of this, a variant of a Reliable User Datagram Protocol (RUDP) [59] is utilized as the default protocol. The problem with RUDP is however that the packets are sent in clear text, but to support industry applications the platform must provide the possibility of encryption. There exists several approaches for enabling this, such as different key exchange schemes with varying degrees of security and overhead. In the end, the decision was to support standard Secure Sockets Layer (SSL) encryption through the Java Secure Socket Extension to make it possible to encrypt the data if needed. The encryption is however only useful to prevent eavesdropping, not man in the middle attacks, because all certificates will be self signed by the end devices. There is also a significant overhead related to SSL, and since there is an initial handshake the useful data will no longer arrive in the first packet.

We have also made design choices in relation to the serialization of messages, namely how the messages are coded when sent over the Internet (before any encryption). Furthermore, minimizing the message size is tightly coupled to requirement 1 on fast dissemination and requirement 3 lightweight operation. For example, a binary serialization format is most suitable from a performance perspective, but a text based format is most usable from a human-readable perspective and a code-specific format is the easiest to program. In the end, the choice was to support all different serialization formats but the default is set to Java's object serialization with added GZIP [60] compression, to make it easier to develop new extensions. Likely, the Java serializer will be replaced in the future, in order to make transitions to other platforms and programming languages feasible. Table I shows a comparison of the message sizes depending on some of the different options for serialization. Where the Get message is a typical small message and the Kelips sync message is a typical large message.

4) Networking Layer: The networking layer enables communication between different entities over current IIP based infrastructure, such as fiber optic networks or wireless and mobile networks. Hence, the networking layer is separated into two inner components, an IP network and the physical network medium. In short, the networking layer thus abstracts any underlying IP-based network architecture. The problem faced in the networking layer was related to requirement 2 on stability and seamless communication. The first versions of the platform did not take NAT and firewalls into consideration, it only worked if all devices was on the public Internet. However, today almost all consumer devices are connected to the Internet through either NAT or some type of firewall, either in their home or at their work, but also on the mobile phone networks. The NAT and firewall problem is however only a question of configuration if the user is allowed to enable features such as port forwarding on the NAT routers. But that this rarely the case.

Multiple approaches were considered, ranging from IPv6 solutions, to Universal Plug and Play (UPnP), different hole punching techniques, and finally simple proxy solutions. The chosen solution first tries the normal approaches, such as direct connections and UPnP. If this fails, it instead utilizes distributed proxy nodes in the system. However, the proxy solution stands in direct contradiction to requirement 1 on real-time communication without unnecessary relaying of information and requirement 2 on no central points of failure. Because of this, it is only used as a last resort when there are no other possible options. There also exists problems related to the capacity of the Internet connections and the network delay, but since the SensibleThings platform is built on top of the existing Internet architecture, it cannot affect these parameters. Therefore, as long as the useful payload data is sent in the first packet, as in the RUDP implementation, it is considered to be transmitted as fast as possible by the underlying network infrastructure.

5) Sensor and Actuator Layer: The sensor and actuator layer enables different sensors and actuators to connect into the platform. The sensors and actuators can vary greatly and the platform therefore offers two options to connect them. Firstly, they can be connected directly if they are accessible from the application code, such as in the case of smartphone sensors. Secondly, the sensors and actuators can connect through the sensor and actuator abstraction. The abstraction enables connectivity either directly to wireless sensor networks or via more powerful gateways. Hence, the sensor and actuator layer is separated into five components: the directly accessible sensors and actuators, an abstraction component, different sensor and actuator networks, sensor and actuator gateways, and the physical sensors and actuators.

In the sensor and actuator layer, there were problems with the actual sensor hardware platforms that is available today, especially in regards to requirement 3 on being lightweight. Different vendors of sensors have different platforms that the sensors run on, especially when it comes to connecting large Wireless Sensor Networks (WSN). Typically, cheap analog sensors can be connected directly to a more powerful device, such as a smartphone or a Raspberry Pi [61]. But to connect traditional WSN architectures such as TinyOS [51] or Contiki [52], the platform must communicate via CoAP [53] or other lightweight protocols that they can handle. Therefore, in most of the examples we have utilized either smartphones with sensors already built in, or Raspberry Pi devices with attached sensors. However, any device that can run the Java code for the platform can be a part of the system, and any low end device that can communicate via CoAP can easily be connected via a more capable device.

C. Source Code License

One purpose of the platform is to make it available for industry partners to develop their own applications and then commercialize the products, see requirement 5. This requirement made it impossible to use a strict and propagating open source license such as GNU General Public License (GPL). The amount of external code should also be kept to a minimum in order to enable an open source community centered around the platform with the power and possibility to easily incorporate changes into the code without navigating multiple licenses. In the end, the decision was to use the GNU Lesser General Public License (LGPL) that allows companies to make commercial products on top of the platform, without forcing their products to be open source as well, while forcing changes to the core to propagate back to the platform and allow the creation of an open source community centered on the platform.

VI. RESULTS

The current results include launching our new development website for the SensibleThings platform (www.sensiblethings.se). This website will act as a portal for all developers who want to utilize the platform in their applications. For example, the website has developer packages to download, the complete source code, example applications, exercises, tutorials, and a wiki. All the code and resources is provided free and under the LGPL version 3 open source license. The remainder of this section will present measurements made on the platform, an evaluation of the platform in relation to the requirements, and some of the demonstrator applications, which have been built using the platform.

A. Measurements

Initial testing, demonstration, and evaluation of the platform has been conducted using a testbed with fixed and mobile access to the Internet. In terms of performance we have measured the platform to be on par with UDP traffic. This was made in two steps, first general expressions were created for the amount of messages sent in the platform, and secondly actual measurements of these expressions were performed in the testbed. The expressions were derived from the life cycle of a device within the system. This life cycle can be seen in Figure 6, which shows the external join rate λ_{join} and the amount of messages per join m_{join}. The rate of resolve operations in an application $\lambda_{resolve}$ and the amount of messages per resolve $m_{resolve}$. The rate of get operations in an application λ_{get} and the amount of get messages per get operation m_{get}. The rate a device has to perform maintenance λ_{maint} and the amount of messages per maintenance run m_{maint}. Finally, the leave rate of devices in the testbed λ_{leave} and the amount of messages for a single device to leave m_{leave}. General expressions for the total amount of messages sent within in the testbed can be seen in (6) to (9). In these, M_{join} represents the total amount of join messages being sent within the platform, which is expressed as the rate of new devices λ_{join} times the amount of messages for a single join operation m_{join}. In similarity, M_{leave} represents the total amount of leave messages being sent within the system, which is expressed as the rate of devices leaving λ_{leave} times the amount of messages for a single leave operation m_{leave}. The variable M_{maint} represents the total amount of maintenance messages being sent within the system. It is expressed as the rate of performing maintenance λ_{maint}, times

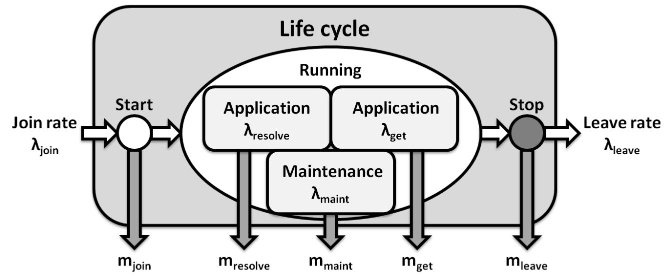


Figure 6. The life cycle of a device.

the amount of messages for a single maintenance operation m_{join} per device, times the total amount of devices in the testbed N. The variable M_{appl} represents the total amount of application messages being sent within the system. It is expressed as the rate of the resolve operations $\lambda_{resolve}$ times the amount of resolve messages for a single resolve operation $m_{resolve}$ and the rate of the get operations λ_{get} times the amount of get messages for a single get operation m_{get}. By summarizing these operations M_{join}, M_{leave}, M_{maint}, and M_{appl} we get the total amount of messages being sent within the whole system M_{total}

$$M_{join} = \lambda_{join} * m_{join} \tag{6}$$

$$M_{leave} = \lambda_{leave} * m_{leave} \tag{7}$$

$$M_{maint} = N * \lambda_{maint} * m_{maint} \tag{8}$$

$$M_{appl} = \lambda_{resolve} * m_{resolve} + \lambda_{get} * m_{get} \tag{9}$$

$$M_{total} = M_{join} + M_{leave} + M_{maint} + M_{appl} \tag{10}$$

We have also formalized expressions for the latencies in the testbed, which can be seen in (11) to (14). In these, L_{join} represents the latency for a device to join the testbed, which is expressed as the processing delay induced by the devices $d_{processing}$ and the amount of messages required for joining m_{join} times the network delay $d_{network}$. In similarity $L_{resolve}$ represents the latency for a device to perform the resolve operation in the testbed, which is expressed as the processing delay and the amount of messages for a single resolve operation $m_{resolve}$ times the network delay. Furthermore, L_{leave} represents the latency for a device to gracefully shutdown and leave the testbed, which is expressed as the processing delays and amount of messages for the graceful shutdown m_{leave} times the network delay. Lastly, L_{get} represents the latency for a device to get a sensor value from another device in the testbed, which is expressed as the amount of messages required for retrieving the data m_{get} times the network delay and any processing delays.

$$L_{join} = m_{join} * d_{network} + d_{processing} \tag{11}$$

$$L_{resolve} = m_{resolve} * d_{network} + d_{processing} \tag{12}$$

$$L_{leave} = m_{leave} * d_{network} + d_{processing} \tag{13}$$

$$L_{get} = m_{get} * d_{network} + d_{processing} \tag{14}$$

We can practically measure (8), (11), (12), (13), and (14) directly in the testbed environment. All other expressions are too highly dependent on the churn rates and the specific

TABLE II
LATENCY MEASUREMENTS FOR JOIN, RESOLVE, AND LEAVE OPERATIONS

Operation \ Device	Dell 780		Nexus 4	
	μ	σ	μ	σ
Join (L_{join})	464 ms	7.67 ms	2580 ms	1320 ms
Resolve ($L_{resolve}$)	36.8 ms	13.5 ms	423 ms	175 ms
Leave (L_{leave})	1.87 ms	0.129 ms	21.9 ms	12.6 ms

TABLE III
LATENCY MEASUREMENTS FOR THE GET OPERATION (L_{get})

Source \ Sink	Dell 780		Nexus 4	
	μ	σ	μ	σ
Fujitsu TX100	4.77 ms	0.145 ms	154 ms	43.5 ms
Padphone infinity	280 ms	136 ms	437 ms	92.8 ms
Pi_1 (Telia)	41.4 ms	2.08 ms	248 ms	40.6 ms
Pi_2 (Servanet)	47.1 ms	11.7 ms	209 ms	38.3 ms
Pi_3 (BBB)	76.3 ms	11.5 ms	277 ms	64.9 ms
Pi_4 (SUNET)	94.7 ms	61.2 ms	303 ms	68.9 ms
Pi_5 (Bahnhof)	106 ms	24.9 ms	289 ms	97,9 ms
Pi_6 (Acreo)	111 ms	36.7 ms	317 ms	123 ms
Pi_7 (Telia)	197 ms	24.6 ms	449 ms	59.7 ms
Pi_8 (Telia)	213 ms	35.1 ms	643 ms	98.2 ms
Pi_9 (Telenor)	370 ms	89.8 ms	1260 ms	937 ms

application scenarios, for any practical and useful measurements to be made. The measurements were performed using a total of 9 Raspberry Pi devices, 2 desktop computers, and 2 mobile device. All with heterogeneous connectivity and different network properties. The performance evaluation was performed on one of the mobile devices (a LG Nexus 4) and one of the desktop computers (a Dell OptiPlex 780), in order to evaluate the effects of the mobile Internet connection and its weaker hardware. Both the desktop computers was connected directly to the campus network with a public IP address and the mobile devices was connected via Telia's mobile 3G/4G network available at the campus in Sundsvall. The maintenance overhead of the whole testbed M_{maint} in (8) with all 13 devices connected, was measured to be on average 82.6 messages per minute with a standard deviation of 3.75 messages. The latencies for the different operations L_{join}, $L_{resolve}$, L_{leave} can be seen in Table II. Where, μ stands for the numerical average and σ for the standard deviation of the measurements. Finally, we measured the latency for retrieving the actual sensor values L_{get} from all the other connected devices, namely a desktop computer (a Fujitsu Primergy TX100), another mobile device (a Padphone infinity), and 9 Raspberry Pi devices. The results from these measurements can be seen in Table III. Where the Dell desktop computer and the Nexus 4 was used as measuring sinks and all the other devices acted as sensor sources.

All the measurements was setup with the two desktop computers and Pi_1, Pi_2, and Pi_6 connected directly with public IP to the platform, while the two mobile devices, Pi_3 to Pi_5, and Pi_7 to Pi_9 was connected via different NAT solutions. An overview of the measurement setup can be seen in Figure 7. To be specific, both desktop computers were connected via Mid Sweden University's SUNET 100/100 Mbps network with public IP addresses. The Nexus 4 was connected via Telia 3G 10/2 Mbps connection behind carrier grade NAT. The Padphone infinity was connected via Telia 4G 20/2 Mbps connection, also behind carrier grade NAT. Pi_1 was connected via Telia ADSL 30/12 Mbps connection with public IP. Pi_2 was connected via ServaNet 100/100 Mbps connection with

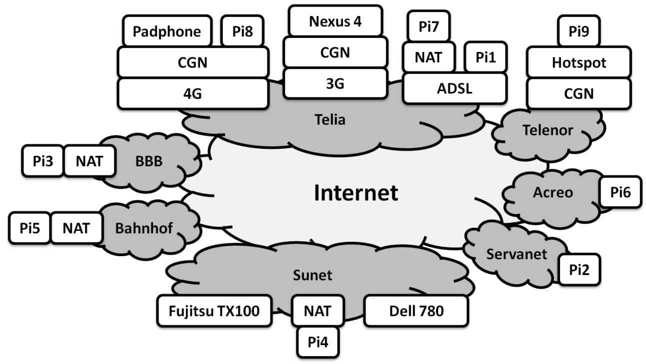


Figure 7. The measurement setup.

a public IP address. Pi_3 was connected via Bredbandbolaget (BBB) 100/10 Mbps connection behind a home NAT router. Pi_4 was connected via Mid Sweden Universities SUNET 100/100 Mbps network behind an enterprise NAT router. Pi_5 was connected via Bahnhof 100/100 Mbps connection behind a home NAT router. Pi_6 was connected via Acreo 100/100 Mbps connection with public IP. Pi_7 was connected via Telia ADSL 8/1 Mbps connection behind a home NAT router. Pi_8 was connected via Telia 4G 40/2 Mbps connection behind a mobile Internet router. Pi_9 was connected via Telenor 3G 10/2 Mbps connection behind both carrier grade NAT and mobile hotspot NAT.

B. Evaluation

In order to ensure that the platform achieves the requirements from Section II and evaluation of them have been made in the currently available platform using the Kelips lookup and RUDP communication. The first requirement on being fast was evaluated using (1), which states how efficient the communication is, as a percentage of the ideal case i.e., ping. To evaluate this, a subset of the setup as the previous measurements were used. We found that the $R_{latency}$ ratio of our platform is calculated to be 5.09 or around 500% using equation (1). This mean that the response times of our platform is 5 times worse than the most optimal ideal case of the shortest response time made possible by the underlying networks. Hence, there is room for improvement in terms of low latency, which is understandable because some of the nodes in the measurements were behind NAT and thus had to be tunneled.

The second requirement was evaluated using two metrics. The first was how the function for the average amount of messages per node changes as the number of nodes increase. Basically the communication load on each node in the system. This was measured in an emulated environment, where a single powerful computer could run multiple instances the platform nodes. Thus, the amount of nodes was incremented between 10 to 100 nodes in steps of 10 for each run. In each of these runs, the total number of messages was counted during a 10 minute period, after which the total was averaged down to messages per minute and then averaged in according to (2) with the number of nodes. The results from these measurements can bee seen in Figure 8, which shows that the current Kelips lookup has a decreasing average communication load, which converges to a constant about 11 messages. This is

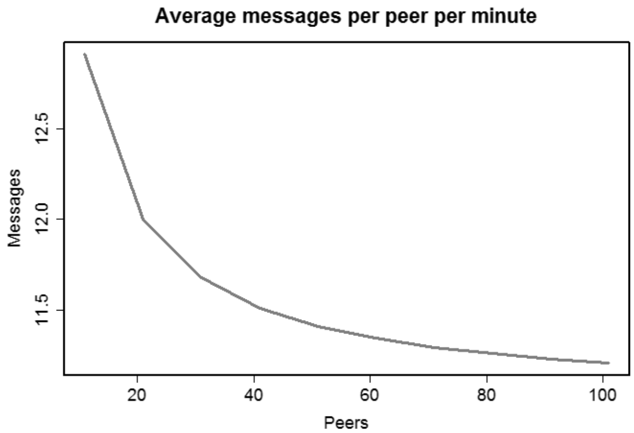


Figure 8. A graph of the average communication in the platform.

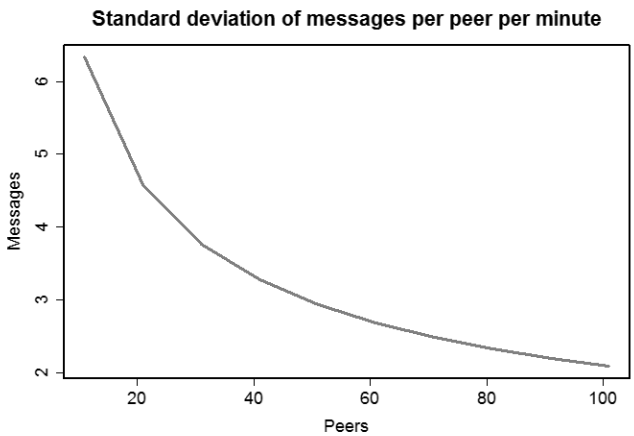


Figure 9. A graph of the standard deviation of the communication load in the platform.

because each additional node only adds a fixed amount of new messages into the system. This show that the Sensiblethings platform scales well, since each new device does not induce a heavier communication load for the ingoing devices. The reason for the higher average with few nodes and decreasing average is because the bootstrap device sends 21 extra messages compared to the other nodes.

In relation to requirement 2, the standard deviation of these values are interesting as well. As it indicated how even the system is and an increasing standard deviation indicates potential central points of failure. A graph of this can be seen in Figure 9, which shows that the standard deviation is decreasing toward a constant around 2 messages. Hence, the system becomes more evenly loaded as new nodes join the platform and there does not seem to exist any central points of failure. In similarity with the average, the only reason for the high standard deviation with few nodes is because the bootstrap device sends more messages compared to the other nodes.

In relation to requirement 3 on being lightweight, the resource utilization has to be evaluated. To evaluate this we tested three different devices running the platform, a Dell 780 workstation, a Raspberry Pi, and a Nexus 4 Android phone. In detail, we measured the memory usage, the processor utilization, and the network utilization of each of these devices when they run live platform. The memory usage and processor utilization measurements was taken directly from the device. But to determine the network utilization, first the maximum network capacity had to be determined. This was determined by using online broadband test tools [62], which show the practical maximum network capacity. This was done because the devices can not achieve their theoretical network speeds because of hardware limitations. For example, the Raspberry Pi has a 100/100 mbit network card but can practically only come up to around 70/30 mbit in actual throughput. Table IV show the results from all the utilization measurements, and thus the utilization metric was measured to be 26.7% in the platform, because that was the utilization on the most heavily loaded device property in our test. Namely, the processor load on the Nexus 4 device.

TABLE IV
UTILIZATION MEASUREMENTS.

Device	Dell 780	Raspberry Pi	Nexus 4
Processor	0.258%	11.0%	26.7%
Memory	1.029%	5.27%	2.10%
Network download	0.0194%	0.0552%	0.209%
Network upload	0.0200%	0.114%	2.10%

Requirement 4 on extensibility can be consider achieved because of the many add-ins developed for the platforms, as well as its component based nature. Where, for example, the lookup service can be exchanged in the future for a more suitable one. Thus, the platform is extensible and future proof because changes can be made to the platform as new standards and paradigms emerge. Lastly, requirement 5 on being easy to adopt and free to use, can be considered achieved since there exists simple tutorials and example code on the website. There is also no monetary cost directly involved in the platform because each node brings their own resources to the system and it has an open source code license that allows commercialization.

C. Demonstrator Applications

Proof-of-concept demonstrator applications have been built using many different devices, sensors, and actuators, in order to show the versatility of the SensibleThings platform. In most demonstrators we have utilized computers, Android smartphones, and Raspberry Pi devices to show the show that the platform works across different devices and network connections. In detail, this article will present functional demonstrators in three areas where the platform can applied to. The first type of application area is sensor reading applications, which simply retrieves sensor values over the platform to monitor things. The second application area is the intelligence home, which is a combination of sensors and actuators. Lastly, the third application, in which we show demonstrator applications of is surveillance applications, to monitor an area with images and video feeds. These are all typical IoT scenarios, but the platform itself is versatile enough to be applied to even more areas. We can foresee possible applications of the platform ranging from health-care, logistics, emergency response, tourism, and smart-grids, to more social-

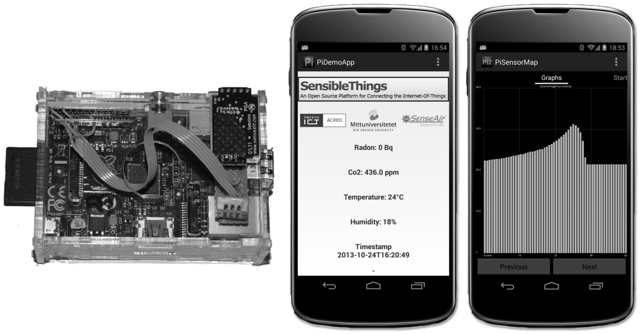


Figure 10. Sensor reading example applications.

Figure 11. Intelligent home example application.

oriented applications such as crowd sourcing, dating services, and intelligent collaborative reasoning.

1) Sensor Reading Applications: The first demonstrator applications are simple sensor reading applications. In these demonstrators we have attached different types of sensors to Raspberry Pi devices and connected the to different types of networks. Figure 10 shows an example Raspberry Pi sensor device and two example android applications, which retrieves the sensor values in a ubiquitous real-time manner trough the SensibleThings platform. The left application show radon measurements from a Raspberry Pi device with an attached radon sensor, which, for example, a landlord or house owner can monitor their properties with. The right application show a graph of a temperature sensor attached to another Raspberry Pi device, which can be used, for example, to monitor the health status of sensitive objects in logistics, such as food or medicine.

2) Intelligent Home Applications: The second demonstrator application is a typical intelligent home scenario, where one wants to control appliances in a home. To this we have utilized an intelligent wall socket, which has built in sensor and actuator functionality. For example, it can turn and off each socket and measure the power consumption of the connected appliance, as well measuring the environmental temperature and humidity. Figure 11 shows this intelligent wall socket and a simple intelligent home application, both connected to the platform. The application can retrieve the sensor values from the socket, control the power, and measure the consumption of the appliances that are plugged in.

3) Surveillance Applications: The third demonstrator application is a typical surveillance scenario, where one wants to monitor a specific area. For example, to detect trespassers or to avoid collisions with obstacles. To this demonstrator application we have utilized Raspberry Pi devices with attached camera sensors, to then transfer the images over the platform. Figure 12 shows this Raspberry Pi camera device and an applications, which utilize the data from the camera. The application retrieves the camera images and performs image analysis to detect obstacles and anomalies. In this case, to detect if a person some other obstacle is present on a railroad track.

Figure 12. Surveillance example application.

VII. Future Internet-of-Things Services

The IoT is often discussed [63], [64], [65], as one of the next steps in the evolution of everyday computing. However, the proliferation has so far only been marginal. Thus, the future of IoT services is still quite uncertain. Where the most prominent one is that there is still no real killer application for the IoT. The scenarios often discussed when advocating an IoT (logistics, health-care, intelligent home, surveillance, etc.) are simply not engaging enough to start the revolution. Furthermore, people are generally skeptical to sharing sensor information [66], which are tightly coupled to themselves, even though some people still share almost anything on social media such Facebook, Twitter, and Google+. Either way, we believe the open sharing of information is paramount to the proliferation of the IoT and without a killer application to really show the benefits to the common user, the proliferation will probably not happen.

The are also uncertainties in the business model [67], for example, the IoT will consume quite significant amount of network resources. This raise questions, such as who shall build this infrastructure and who shall stand for this cost. Basically, how can sensor manufacturer make money on IoT devices, how can operators make money on IoT traffic, and how can application developers make money from applications that utilize information from the IoT. There is also the question of ownership of the information, who owns the information on the IoT and thus determines the cost of using it. Furthermore, as we presented in Section III, there is still no single standard or commonly accepted solution for the IoT, there are still questions regarding the protocols and architectures to use. We believe that because of the vision of an open flow of information on the IoT, an open protocol, and free open source implementations has the most advantages, which is why we built the SensibleThings platform. But the standardization committees might have a different points of view on the openness of the information and protocols and the patent situation also plays a large role in the standardization efforts.

Because of all the uncertainties we believe that the true killer application of the IoT is caught in a form of catch 22. We believe that this killer application is the creation of truly intelligent behavior in all different applications, which can understand the situations and context of their users and react accordingly. The closest service available today is Google Now [68], which tries to achieve this behavior with the information Google has available. However, to truly create this type of functionality it will require large amounts sensor information, which is not yet available, since people do not have the hardware or are not sharing it yet. Furthermore, truly context aware applications also require large amounts of processor power to run different machine learning algorithms, which some one must pay the cost for. And since there is real business model yet, no one can take this cost responsibility for it. Hopefully, there will some type of consensus in the IoT research soon. So that the common user can start seeing the benefits what a global and ubiquitous IoT with open sharing of information can offer.

VIII. Conclusion and Future Work

In this paper, we presented the challenges we have encountered when researching and developing the SensibleThings platform. The first contribution is the identification of the requirements for a functional and fully distributed IoT platform in Section I. The second contribution is a short survey of existing IoT platforms, presented in Section III. The main contribution is however the SensibleThings platform itself, which is shown to fulfill all the stated requirements. The platform can disseminate information to end devices quickly with low overhead (requirement 1). It is reliable and operates without any central points of failure (requirement 2). The platform is lightweight and can run on mobile devices or Raspberry Pi devices (requirement 3). It is extensible (requirement 4) with all its add-ins and functionality. Finally, the platform is licensed under a well known and widely accepted open source license making it free to use and at the same time encouraging development of commercial a products (requirement 5). Therefore, our conclusion is that when compared to related work, the SensibleThings platform can be classified as a fully distributed solution, where the overhead and communication is kept as lightweight as possible. This is the only type of solution that will scale for billions of connected devices and still be able to disseminate sensor information with real-time demands.

Our future work are directed toward improving and optimizing the existing code, as well as investigating other possible choices of the DHT and communication protocol. We will also develop more extensions to satisfy specific applications demands, such as seamless integration with other IoT platforms and cloud infrastructures. Finally, we will investigate the possibility to create a formal open communication standard of the platform.

Acknowledgment

This research has been supported by grant 00163383 of the EU European Regional Development Fund, Mellersta Norrland, Sweden. The authors also want to thank our partners from industry and academia, in particular Acreo AB and Stockholm University.

References

[1] S. Forsström, V. Kardeby, P. Österberg, and U. Jennehag, “Challenges when realizing a fully distributed internet-of-things-how we created the sensiblethings platform,” in ICDT 2014, The Ninth International Conference on Digital Telecommunications, 2014, pp. 13–18.

[2] T. Kanter, S. Forsström, V. Kardeby, J. Walters, U. Jennehag, and P. Österberg, “Mediasense–an internet of things platform for scalable and decentralized context sharing and control,” in ICDT 2012, The Seventh International Conference on Digital Telecommunications, 2012, pp. 27–32.

[3] G. Wu, S. Talwar, K. Johnsson, N. Himayat, and K. Johnson, “M2M: From mobile to embedded Internet,” IEEE Communications Magazine, vol. 49, 2011, pp. 36–43.

[4] L. Atzori, A. Iera, and G. Morabito, “The internet of things: A survey,” Computer Networks, vol. 54, 2010, pp. 2787–2805.

[5] J. Hong, E. Suh, and S. Kim, “Context-Aware Systems: A Literature Review and Classification,” Expert Systems with Applications, vol. 36, 2009, pp. 8509–8522.

[6] L. Atzori, A. Iera, and G. Morabito, “From" smart objects" to" social objects": The next evolutionary step of the internet of things,” IEEE Communications Magazine, vol. 52, no. 1, 2014, pp. 97–105.

[7] M. Weiser, “The computer for the 21st century,” Scientific american, vol. 265, no. 3, 1991, pp. 94–104.

[8] D. E. O’Leary, “Big data, the internet of things and the internet of signs,” Intelligent Systems in Accounting, Finance and Management, vol. 20, 2013, pp. 53–65.

[9] L. Ericsson, “More than 50 billion connected devices,” 2011.

[10] H. Sundmaeker, P. Guillemin, P. Friess, and S. Woelfflé, Vision and challenges for realising the Internet of Things. EUR-OP, 2010.

[11] Z. Sheng, S. Yang, Y. Yu, A. V. Vasilakos, J. A. McCann, and K. K. Leung, “A survey on the ietf protocol suite for the internet of things: Standards, challenges, and opportunities,” IEEE Wireless Communications, vol. 20, 2013, pp. 91–98.

[12] P2413 Standard for an Architectural Framework for the Internet of Things (IoT). IEEE. [Online]. Available: http://iot.ieee.org/ [retrieved: December, 2014]

[13] AllSeen Alliance. [Online]. Available: https://allseenalliance.org/ [retrieved: December, 2014]

[14] European Research Cluster on the Internet of Things. [Online]. Available: http://www.internet-of-things-research.eu/ [retrieved: December, 2014]

[15] Hyper/Cat. [Online]. Available: http://www.hypercat.io/ [retrieved: December, 2014]

[16] Internet of Things Architecture (IoT-A). [Online]. Available: http://www.iot-a.eu/public [retrieved: December, 2014]

[17] Internet of Things Counsil. [Online]. Available: http://www.theinternetofthings.eu/ [retrieved: December, 2014]

[18] Internet of Things Europe. [Online]. Available: http://www.internet-of-things.eu/ [retrieved: December, 2014]

[19] Internet of Things Initiative (IoT-I). [Online]. Available: http://www.iot-i.eu/public [retrieved: December, 2014]

[20] IoT Forum. [Online]. Available: http://iotforum.org/ [retrieved: December, 2014]

[21] IP for Smart Objects (IPSO) Alliance. [Online]. Available: http://www.ipso-alliance.org/ [retrieved: December, 2014]

[22] oneM2M. [Online]. Available: http://www.onem2m.org/ [retrieved: December, 2014]

[23] Open interconnect. [Online]. Available: http://openinterconnect.org/ [retrieved: December, 2014]

[24] A. L. Valdivieso Caraguay, A. Benito Peral, L. I. Barona López, and L. J. García Villalba, “SDN: Evolution and Opportunities in the Development IoT Applications,” International Journal of Distributed Sensor Networks, vol. 2014, 2014, pp. 1–10.

[25] F. Hu, Q. Hao, and K. Bao, “A survey on software-defined network and openflow: From concept to implementation,” Communications Surveys Tutorials, IEEE, vol. 16, no. 4, Fourthquarter 2014, pp. 2181–2206.

[26] Z. Qin, G. Denker, C. Giannelli, P. Bellavista, and N. Venkatasubramanian, "A Software Defined Networking architecture for the Internet-of-Things," in 2014 IEEE Network Operations and Management Symposium (NOMS). IEEE, May 2014, pp. 1–9.

[27] SicsthSense. [Online]. Available: http://sense.sics.se [retrieved: December, 2014]

[28] ThingSpeak. [Online]. Available: https://www.thingspeak.com [retrieved: December, 2014]

[29] Sen.se. [Online]. Available: http://open.sen.se [retrieved: December, 2014]

[30] Nimbits. [Online]. Available: http://www.nimbits.com [retrieved: December, 2014]

[31] Thingsquare. [Online]. Available: http://thingsquare.com [retrieved: December, 2014]

[32] EVRYTHNG. [Online]. Available: http://www.evrythng.com [retrieved: December, 2014]

[33] Paraimpu. [Online]. Available: http://paraimpu.crs4.it [retrieved: December, 2014]

[34] Xively. [Online]. Available: https://xively.com [retrieved: December, 2014]

[35] XOBXOB. [Online]. Available: http://www.xobxob.com [retrieved: December, 2014]

[36] ThingWorx. [Online]. Available: http://www.thingworx.com [retrieved: December, 2014]

[37] One Platform. [Online]. Available: http://exosite.com/products/onep [retrieved: December, 2014]

[38] Carriots. [Online]. Available: https://www.carriots.com [retrieved: December, 2014]

[39] Openiot: Open source solution for the internet of things into the cloud. [Online]. Available: http://openiot.eu [retrieved: July, 2014]

[40] Sap internet of things. [Online]. Available: http://www.sap.com/pc/tech/internet-of-things.html [retrieved: December, 2014]

[41] ETSI, "Machine-to-machine communications (m2m); functional architecture," (TS 102 690 V1.1.1), Tech. Rep., 2011.

[42] M. Presser, P. Barnaghi, M. Eurich, and C. Villalonga, "The SENSEI project: integrating the physical world with the digital world of the network of the future," IEEE Communications Magazine, vol. 47, no. 4, 2009, pp. 1–4.

[43] H. Koumaras, D. Negrou, F. Liberal, J. Arauz, and A. Kourtis, "ADAMANTIUM project: Enhancing IMS with a PQoS-aware multimedia content management system," International Conference on Automation, Quality and Testing, Robotics, vol. 1, 2008, pp. 358–363.

[44] G. Camarillo and M.-A. Garcia-Martin, The 3G IP multimedia subsystem (IMS): merging the Internet and the cellular worlds. Wiley, 2007.

[45] K. Aberer, M. Hauswirth, and A. Salehi, "A middleware for fast and flexible sensor network deployment," in Proceedings of the 32nd International Conference on Very Large Data bases, 2006, pp. 1199–1202.

[46] C. Jennings, B. Lowekamp, E. Rescorla, S. Baset, and H. Schulzrinne, "REsource LOcation And Discovery (RELOAD) Base Protocol," RFC 6940 (Proposed Standard), Internet Engineering Task Force, Jan. 2014. [Online]. Available: http://www.ietf.org/rfc/rfc6940.txt

[47] T. Berners-Lee, R. Fielding, and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax," RFC 3986 (INTERNET STANDARD), Internet Engineering Task Force, Jan. 2005, updated by RFCs 6874, 7320. [Online]. Available: http://www.ietf.org/rfc/rfc3986.txt

[48] G. Chellani and A. Kalla, "A review: Study of handover performance in mobile IP," CoRR, vol. abs/1312.2108, 2013, pp. 137–151.

[49] S. Forsström and T. Kanter, "Continuously changing information on a global scale and its impact for the internet-of-things," Mobile Networks and Applications, vol. 19, no. 1, 2014, pp. 33–44.

[50] S. Forsström and V. Kardeby, "Estimating contextual situations using indicators from smartphone sensor values," in 2014 IEEE International Conference on Internet of Things (iThings), 2014.

[51] P. Levis, S. Madden, J. Polastre, R. Szewczyk, K. Whitehouse, A. Woo, D. Gay, J. Hill, M. Welsh, E. Brewer, and D. Culler, "TinyOS: An operating system for sensor networks," in Ambient intelligence. Springer, 2005, pp. 115–148.

[52] A. Dunkels, B. Gronvall, and T. Voigt, "Contiki-a lightweight and flexible operating system for tiny networked sensors," in Local Computer Networks, 2004. 29th Annual IEEE International Conference on. IEEE, 2004, pp. 455–462.

[53] Z. Shelby, K. Hartke, and C. Bormann, "The Constrained Application Protocol (CoAP)," RFC 7252 (Proposed Standard), Internet Engineering Task Force, Jun. 2014. [Online]. Available: http://www.ietf.org/rfc/rfc7252.txt

[54] R. T. Fielding and R. N. Taylor, "Principled design of the modern web architecture," ACM Transactions on Internet Technology (TOIT), vol. 2, no. 2, 2002, pp. 115–150.

[55] T. Kanter, S. Pettersson, S. Forsstrom, V. Kardeby, R. Norling, J. Walters, and P. Osterberg, "Distributed context support for ubiquitous mobile awareness services," in Communications and Networking in China, 2009. ChinaCOM 2009. Fourth International Conference on, Aug. 2009, pp. 1–5.

[56] I. Stoica, R. Morris, D. Karger, F. Kaashoek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup service for internet applications," in in Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, vol. 31. New Yourk, NY, USA: ACM Press, 2001, pp. 149–160.

[57] I. Gupta, K. Birman, P. Linga, A. Demers, and R. Van Renesse, "Kelips: Building an efficient and stable p2p dht through increased memory and background overhead," in Peer-to-Peer Systems II. Springer, 2003, pp. 160–169.

[58] K. Aberer, "P-grid: A self-organizing access structure for p2p information systems," in Cooperative Information Systems. Springer, 2001, pp. 179–194.

[59] T. Bova and T. Krivoruchka, "Reliable udp protocol," IETF, Tech. Rep., 1999.

[60] P. Deutsch, "GZIP file format specification version 4.3," RFC 1952 (Informational), Internet Engineering Task Force, May 1996. [Online]. Available: http://www.ietf.org/rfc/rfc1952.txt

[61] Raspberry pi. [Online]. Available: http://www.raspberrypi.org [retrieved: December, 2014]

[62] L.-P. G. Hansen, I. C. Østhus, and P. Heegaard, "Online broadband test tools," in Proceedings of Norsk informatikkonferanse 2009, 2009.

[63] P. Balamuralidhara, P. Misra, and A. Pal, "Software platforms for internet of things and m2m," Journal of the Indian Institute of Science, vol. 93, no. 3, 2013, pp. 487–498.

[64] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," Future Generation Computer Systems, vol. 29, no. 7, Sep. 2013, pp. 1645–1660.

[65] S. Hachem, A. Pathak, and V. Issarny, "Service-Oriented Middleware for the Mobile Internet of Things: A Scalable Solution," in IEEE GLOBECOM: Global Communications Conference. Austin, United States: IEEE, Dec. 2014. [Online]. Available: https://hal.inria.fr/hal-01057530

[66] R. Roman, J. Zhou, and J. Lopez, "On the features and challenges of security and privacy in distributed internet of things," Computer Networks, vol. 57, no. 10, Jul. 2013, pp. 2266–2279.

[67] S. Turber, J. vom Brocke, O. Gassmann, and E. Fleisch, "Designing business models in the era of internet of things," in Advancing the Impact of Design Science: Moving from Theory to Practice. Springer, 2014, pp. 17–31.

[68] Google Now. [Online]. Available: http://www.google.com/landing/now/ [retrieved: December, 2014]

Towards Optimized Performance in Military Operations

Major Tapio Saarelainen, PhD; IARIA Fellow
Army Academy
Research and Development Division
Finland
tapio.saarelainen@mil.fi

***Abstract*—Succeeding in military operations requires that the available resources be timely assessed and optimized. Sustaining combat performance asks for suitable tools for organizing services and resources that allow drawing from Situational Awareness, Common Operational Picture and utilizing Graphic User Interfaces. When aiming at maximum impact with the resources available, Service Oriented Architecture (SOA) can be used as a tool in the process of organizing the services. This paper presents an idea-phase introduction of how to facilitate military decision-making process with the assistance of the systems described here. The objective involves increasing the performance capability to allow time-efficient execution of military operations. The paper examines the optimization of services requested by means of SOA when viewing a military maneuver as a Business Process (BP) and sketches a preliminary Grid for Interface Assisted Decision Making (GIADM) to facilitate timely and accurate data processing. GIADM exploits the data offered by means of SOA, the Resource Manager (RM), and BP, and aims at facilitating the decision making process of a military commander at the planning stage of a given operation including evaluating Courses of Action (COA). The paper sets out to identify how to utilize the collected data from the battlespace to empower the process of decision-making and how to benefit from SOA in the creation process of Grid for Interface Assisted Decision Making.**

Keywords-Grid for Interface Assisted Decision Making, Business Process, military operation, Resource Manager, Service Oriented Architecture.

I. INTRODUCTION

Requiring the optimization of the resources available, any military maneuver, including the dismounted company attack, serves as an example of a military operation in which the optimization of resources is the key to success [1]. This paper uses dismounted company attack as an example of a process that requires the constant ability to transmit relevant data and allocate requested services. In military operations, the destruction of mechanical and electrical adversaries continues to be necessary. Material destruction focuses on the enemy's manpower and machinery, when electronically performed destruction can be understood as deleting the adversary's capability to use electronic devices.

The asymmetric nature of war requires improved capabilities in allocating the resources available. This sets increasing demands for Situational Awareness (SA), the data necessary for commanders executing operations in the battle space: precise data concerning location information, updated data involving performance capabilities of own troops and their current operation status.

This paper examines the optimization of services requested by means of SOA when viewing a military maneuver as a Business Process and discusses a possible method for organizing the distribution of requested services in light of the set of assisting tools available real-time. Although these types of systems tend to be based on classified data, this paper exclusively relies on public domain sources.

This paper argues that even though the requested services can be processed with the described tools (Resource Manager and Scheduler) and outlined methods of organizing services (Business Process –like actions assisted with SOA), the system still remains incomplete because it lacks a key tool, a field-tested grid that would take into account all the relevant and constantly changing variables, such as the planned duration of an operation, the type of terrain where the operation is taking place, the expected troop performance and the capabilities of a specific military unit. None of the expected performance or capability values have been determined, evaluated or inserted as part of any Battle Management System. This paper sketches the idea of a Grid for Interface Assisted Decision Making (GIADM) to facilitate timely and accurate data processing.

Military commanders depend on automatic data collecting processes and tools to simplify complicated military maneuvers. To optimize performance, a military operation, an attack, can be simplified in the form of a Business Process (BP) by means of Service Oriented Architecture (SOA). The element named Resource Manager is a tool to be used in managing and allocating the existing resources. A BP -like orchestration of systems and services may improve the overall performance of military operations executed. This may result in improved overall performance capabilities while executing missions in the battlespace benefitting from SOA, the RM and the Scheduler [1].

Comparing modeling war as a process assisted with Service Oriented Architecture with other types of approaches would be difficult, as these types of models tend to be labeled as classified data. Hence, comparing different models remains outside the scope of this paper. Consequently, the details concerning network topology and energy supply are excluded from this paper as well.

The military objectives concerning digitizing dismounted soldiers, sensing their environment and sharing information, will likely require as much as twice the power required by soldiers today [2]. Issues related to power usage remain outside the scope of this study for the reasons stated above.

This paper is organized as follows: Section II introduces the related work, Section III discusses the Military Operation as a process, Section IV deals with Service Oriented Architecture and the allocation of requested services, and Section V examines a maneuver, a dismounted company attack, as a Business Process. Section VI combines the Business Process and the Resource Manager. Section VII explains the significance of communication and networks in the process of war, and Section VIII discusses the use of Unmanned Vehicles (UVs) and Section IX introduces a Grid for Interface Assisted Decision Making (GIADM). Section X concentrates on the Graphic User Interface (GUI). Section XI discusses on security and the Scheduler and Section XII concentrates on benefits and drawbacks of the introduced system. Section XIII concludes the paper with Section XIV suggesting further work.

II. Related Work

Since the existence and utilization of military troops, the necessity to improve the performance of soldiers and overall military performance has loomed large. This translates into the survival of the fittest. As we speak, in the era of Information Technology and beyond, the need to use advanced electrical devices to gain increased military performance has become a military research and development stable. When different types of Battle Management Systems have been applied to assist in using weapon systems, the performance of these Battle Management Systems have been linked to command and control systems to facilitate automated decision making.

In the field of command and control, the need for Artificial Intelligence has been recognized from the perspective of modeling command and control [3]. As noted in [4], command and control increasingly necessitates the use of automated decision assistants and automated decision makers to aid in executing the complex and dynamic operations performed in the battlespace. The presented blueprint in [4] also provides design guidance for resource management.

As explained in [5], a new method of firepower allocation is presented, which uses the structure of Multi-Agent system to boost the firepower application as a process. Also a Multi-Agent System (MAS) in Computer Generated Forces (CGF) has been discussed in [6] in an environment of Tank combat unit. In [6] the hierarchical CGF agent model is designed to support the planning and execution of different tasks. Furthermore, a modeling method of agent-based Computer Generated Forces is discussed in [7] and a flowcart of simulation process has been created complemented by architecture on the combat simulation model.

As discussed in [8], it is necessary to support commanders' continuous access to information and to automate the Military Decision-Making Process (MDMP). As [8] describes, with the use of the computational model R-CAST, cognitive agents can effectively assist the commander and thus bring about improved cognitive performance in military decision making tasks.

Service Oriented Architecture (SOA) can be identified as a central factor in military environment, especially in creating capabilities. As identified in [9], SOA can be useful in providing required capabilities for the military in the needs concerning command, control, communications, computers, intelligence, information, surveillance and reconnaissance (C^4I^2SR). SOA can assist in gaining increased capabilities in Situational Awareness, improved Qulity of Service (QoS) and in decision making process. Soldier's overall performance remains critical. As described in [10], soldiers' performance has been studied by means of creating a map featuring a Bayesian Soldier Performance Model. The importance of the performance of a single soldier can never be over emphasized from the perspective of a military operation.

The papers referred to above leave outside their scope the orchestration and allocation of services in the battlespace, such as processing fire support orders, requesting evacuation, and gaining information from designated areas. Moreover, these papers leave outside their scope the military performance capability offered for both a soldier and a military unit with the assistance of SOA. This requires that SOA be properly connected with the executing processes of a military operation. In addition, thre is room in the existing discussion for examining the use of Unmanned Vehicles as robotic platforms for communication purposes.

This paper views a military operation as a Business Process. The use of SOA has been introduced together with possible services to be orchestrated and allocated as listed in the previous paragraph. Also, a Resource Manager and its functions have been introduced and combined with Business Processes and SOA. This paper emphasizes the significance of ubiquitous communication services and a reliable, operationally secure network system in executing military operations. Soldiers' performance has been evaluated with the assistance of Psycho Physical Factor. The formula used is linked to commanding troops (identified as nodes) in a military operation. A new tool to assist commander's decision making process has been introduced. The tool is named Grid for Interface Assisted Decision Making (GIADM) The paper offers a perspective for comprehending the complexity of executing a military process.

III. Military Operation as a Process

Operating in military settings relies on the existing resources available real-time. These existing military resources can be understood as comprising troops or units (personnel), logistics (resupply materiel, medical supplies) and heavy machinery (tanks, armored vehicles), ammunition (cartridges, shells) and services (transportation, resupply, evacuation). These resources have to be available and accessible when executing a military operation. War can be modeled as a Business Process and Service Oriented Architecture (SOA) can used as an assistant in military operations, as described in [11]. Military operations are composed of parts, such as Information Management, which is linked with Network operations connected to Network

Centric Warfare (NCW). NCW adopts the characteristics of modern age into the military battlespace, especially when applying SOA in issues concerning battlespace Situational Awareness [12]. As communication systems remain vital in executing military processes of any kind, a functional and reliable communication network system ensures that issues related to Situation Awareness (SA) and Common Operational Picture (COP) can be tackled. The following Figure 1 displays communication systems related to Information Management and Network Operations.

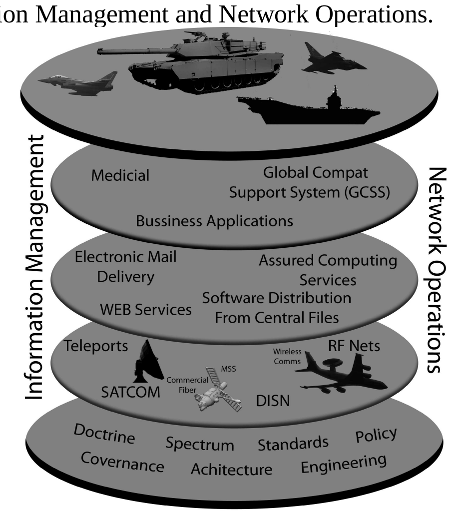


Figure 1. Relation of Information Management and Network Operations.

When applying the Business Process approach to Service Oriented Architecture, the services available are viewed as corresponding with existing real resources. These real resources can be more easily understood as services produced by the military troops themselves or by neighboring troops or services being served from the higher echelon. These services can be seen, for instance, as maintenance of machinery, close-air support (CAS), evacuation and resupply.

The typical military operation carried out by a dismounted company is an attack maneuver. A dismounted company attack is composed of the actions performed by the company itself, as well as services offered by a higher echelon. A dismounted company attack is performed by using personnel (soldiers) equipped with personal weapons and heavy machinery (trucks, armored personnel carriers, artillery). A dismounted company attack is illustrated in Figure 2.

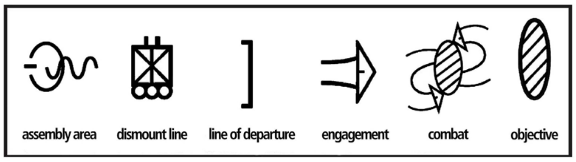


Figure 2. Dismounted Company attack as a process.

A planned dismounted attack usually starts from the assembly area, moves on to the dismount line, via a line of departure, advances to engagement, results in close combat and ends when the set objective is reached. The SOA BP approach can increase the probability of success of an attack by empowering the human-based decision-making process with computers. This can enable an optimal use of resources, and thereby improve overall performance in operations.

A successful military operation requires coordination and timing of existing resources. The availability of real resources in a given place and time is limited and needs to be meticulously scheduled. Usually, SOA services are assumed to be independent of each other but this assumption is no longer valid if SOA services represent real existing resources. The Resource Manager is a necessary element in SOA architecture as discussed in [1]. In the case of a dismounted company attack, successful performance requires that the requested services, for instance, processing a fire support order, are allocated timely and accurately. This sets demands for enhanced Situational Awareness. In the utilization of SOA in military operations, challenges of real-time SOA must be solved [1]. To successfully execute BPs, the Business Process Execution Language (BPEL) is required, as argued in [13].

Organizing of troops and making decisions concerning weapon selection equals a process, in which automation is needed. Situations alter rapidly in the battlespace. Human capabilities are very limited as regards simultaneously monitoring different sensors, a number of screens and communicating with all the decision makers while executing commands. All the matters concerning Computers, Command, Control, Communication, Intelligence, Information, Surveillance and Reconnaissance (C4I2SR) and involving targeting and weapon systems by default value tend to change rapidly and, obviously, any change in one system can result in an immediate change in another system. An example of this would be a weapon system with the effective range of 5000 meters. When a targeting system recognizes the potential target at 7000 meters, locks into the potential target at this distance and transmits the data to the decision maker, the target has already travelled possibly several kilometers before the command for executing the firing mission arrives. By the time the firing order is executed, the target can be out of range of the weapon system (which was the set 5000 said), depending on the target's pace. Therefore, automation is needed, but only to an extent; otherwise, own or friendly aerial targets may get annihilated. This decision making process when applying targeting and weapon systems is also affected by matters concerning Situational Awareness and Common Operational Picture as described in Figure 3.

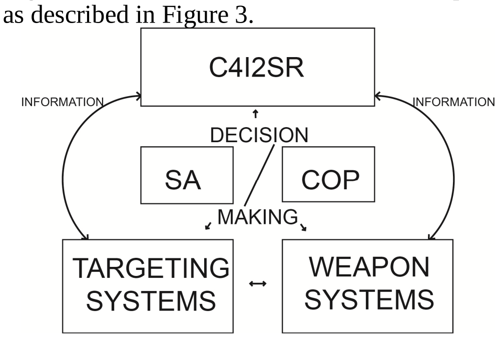


Figure 3. Decision making in applying targeting and weapon systems.

IV. Service Oriented Architecture and Services

To utilize Service Oriented Architecture in a military operation, its capabilities and nature need to be comprehended and to be aware of the services available to organize the system to support a military operation, such as a dismounted company attack.

When needing to affect the sequencing and pace of events in the name of increasing efficiency, SOA is needed as an accelerator. In a successful military operation assisted with SOA, the end result can be seen as an operation with a minimum amount of collateral damage and fratricide. Moreover, a successful operation can be seen as efficient use of resources (ammunition, troops, vehicles, medical support and time). Thereby SOA is viewed as an aid for a military commander in the decision making process, for instance, when choosing Courses of Actions (COAs). Figure 4 depicts an example of two different types of COAs in the battlespace. In the first COA, the objective is to stop an armored enemy, whereas in the second COA, the objective is to destroy a Command Post.

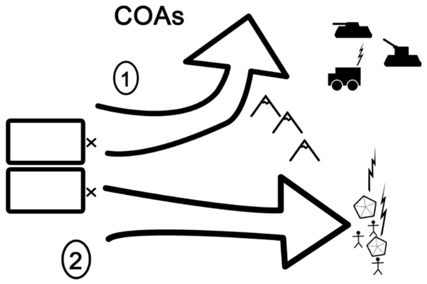


Figure 4. Different types of Courses of Actions.

SOA enables organizations and entities to operate complicated systems and enhance interoperability, collaboration, see [1], and foster the reusing of components and interfaces. SOA can be used in service collaboration. With the correct framework, SOA allows publishing services in a service registry and exchanging data through the Simple Object Access Protocol (SOAP) [13]. SOA offers an adjustable solution for systems integration, applications, protocols, data sources and processes to form a cohesive system that supports the execution of critical BPs [2]. SOA can be used as a collaboration tool in crises management and military environments if the challenges of real-time SOA [14] are solved.

In order to successfully execute BPs, the Business Process Execution Language (BPEL) is required, as argued in [11]. In military systems, the adoption of SOA principles can beneficially result in the overall improvement of system flexibility and maintenance. SOA provides the user with richer information sets via the ability of Web Services to reach out through the networks, see [15]. In the process of achieving greater interoperability, SOA can be used by utilizing service oriented migration and reuse technique, described in [16].

In Network Centric Warfare contexts, SOA has been recognized to act as an enabler of services. SOA is an architecture style that encourages loose coupling between services to facilitate interoperability and the reuse of existing resources as described in [17]. SOA is seen as a tool in enabling agility to handle the changing dynamic evolution needed in network enabled capability, see [18]. The concept of NCW can be viewed as an integration of assets to meet a mission objective, as discussed in [18]. NCW fosters SOA to achieve flexible forces, which are constantly ready and deployable, capable of dynamic changes and evolution to achieve realizable effects. To benefit from SOA in an optimal way, organizations require a comprehensive and applicable SOA governance framework to implement the management and control mechanisms in the system, as argued in [19].

It has been pointed out [20] that Shared SA is in central role for network-enabled capabilities, as described in [20]. In NEC, SOA is most commonly realized through Web Services' GUIs, as discussed in [21], using Extensible Markup Language (XML) formatted documents, see [20]. As evident, XML WS have been recently used to implement SOA enabling the building of BPs by dynamically calling services from the World Wide Web.

SOA is an open concept and supports plug-and-play capabilities of heterogeneous software and hardware components, with the implementation of Web Services, which is probably so far the most popular implementation of SOA, as discussed in [22]. For this reason, SOA has been selected as the architectural solution for the C4I2SR systems for the Finnish Defence Forces [23]. SOA is seen as an enabler in crises management organizations for delivering data and services across political, organizational and cultural boundaries as well as addressing the issues of information sharing regardless of where required data is stored, as concluded in [24]. The global information grid is an essential vehicle in the execution of SOA and for the transformation of data.

In tactical operations, the significance of the real-time location data plays an important role. The tools available include different types of Tactical Battle Management Systems (BMS) for dismounted combat to produce the location information of own troops and precise target designation. The tools for target designation and air-land coordination are necessary requirements for success in operations as described in [25]. Furthermore, air-to-ground communications are described in [26].

Requirements related to improved Situational Awareness, communications and networks are described in [27], grid computing in the battlespace plays an important role as described in [28], and enhancing squad communications with the assistance of smart phones is described in [29]. Lastly, multiplication of various technologies is introduced in [30]. Their overall purpose is to increase the performance of a Future Force Warrior. The inputs of all these tools and networks can be calculated with the assistance of tools used in Business Processes and with the assistance of Service Oriented Architecture.

There has to be a carefully defined military operation before SOA can be adopted in a specific process. The planning sequence requires modeling the military operation into small fragments and the relations and functions between different fragments have to be planned. Moreover, parts and functions of Business Processes have to be opened and studied before embedding SOA into BP by examining the necessary services and their features.

The variety of services used in BPs may be in the operational use of a single unit or several units at the same time. This requires an efficient orchestration of services to maintain service control. SOA can be seen as an enabler in the process of executing military operations as BPs.

The offered services during an advancing dismounted attack are listed in Table I. Most of these services can be pre-programmed to concern the wanted product-line Future Force Warrior (FFW) level. The company commander utilizes various services (fire support orders, location services, medical care, resupply, evacuation, geographical information system -map-service, Blue Force Tracking) while executing the commanded attack from the assembly area to the objective. Table I illustrates possible services available for a dismounted company attack.

TABLE I. LIST OF PRE-PROGRAMMED AND ADDITIONAL SERVICES IN A DISMOUNTED COMPANY ATTACK.

Area of dismounted attack		*Basic services for the Warriors*	*Advanced services (platoon leaders and above)*
1	Assembly area	Location data, terminal guidance to the dismount line	Blue Force data, evacuation
2	Dismount line	Blue Force data, evacuation, resupply	Blue Force data, evacuation
3	Line of departure	Blue Force data, evacuation, resupply	Precision location data, fire support
4	Engagement	Precision location data, fire support	Air-strike
5	Combat	Reinforcement, evacuation, resupply	Air-strike, preparing instructions to the following mission
6	Objective	Evacuation, resupply, reinforcement, precise location data	Air-strike, next mission objective and its time-frame

Fulfilling a requested service asks for the requested service to be available and within range. When dealing with Fire Support Orders (FSO), the range limitations of artillery units are critical. An artillery unit has to be located within appropriate range, and it has to be ready to intake Fire Support Orders and execute them in the required time frame, precisely as ordered.

In a military operation, it is important both to optimize the resources and also minimize collateral damage. SOA can act as a functioning part of the weapon selection process, especially in offering the weapons available for the use of a company commander. Figure 5 indicates the caption from a Battle Management System from a target rich environment. The mission is to destroy the enemy and simultaneously minimize fratricide and collateral damage.

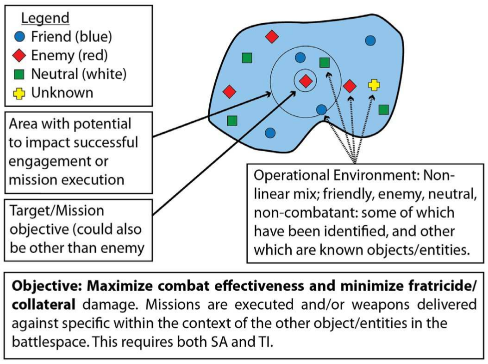


Figure 5. Challenges in weapon selection process.

Battle Management Systems are strongly linked to Command and Control systems. Typically Command and Control systems provide support for Battle Management Systems in various tasks, such as operational planning, obtaining Situational Awareness, creating Common Operational Picture and sustaining the decision making process. The goal of Command and Control systems is to increase the operational capabilities of the operating units from the battalion level to a single platform and soldier level [31].

When we observe the possibilities to fuse the information, Battle Management Systems are usable to an extent. Battle Management Systems use Battle Management Language (BML), which also enables automatic information fusion, as discussed in [32]. As BML represents an unambiguous language for military communication, it is a usable tool in a digitized battlespace environment. As described in [32], the vocabulary of BML is based on a Joint Consultation, Command and Control Information Exchange Data Model (JC3IEDM). Several versions of BML are available to meet the requirements of a consumer for different purposes.

Military commanders need to process a vast amount of data for decision making purposes. The formula to calculate the cumulative uncertainty can be found in [32]. This formula produces a result, which still has to be evaluated by a human being. The significance of a human-being in a decision loop continues to be indispensable. Digitized tools need to be utilized especially in the planning stage of an operation.

V. BUSINESS PROCESS

The orchestration of Business Processes requires a tool for allocating resources, the Resource Manager (RM). The tool has been described in [33]. The RM sorts out and lines up the requested services. These requested services can be understood as service actions performed or produced by humans or machines, involving transportation, issuing Fire Support Orders or evacuation. Furthermore, a single service can be understood as a data exchange process. As militaries implement the framework of Network Centric Warfare with a continuing need to automate the command and control (C2) tools utilized in military, the tempo of operations must be taken into consideration. The collected data need to be processed, analyzed, verified, transmitted, and finally stored. SOA can be identified as a technology that can satisfy these needs of network centric operations. The starting point in the BP approach to SOA is that the main business operations of the organization are described by SOA BPs. The Business Processes are chains of logic that request SOA services. In the case of a military setting, the Business Processes represent military operations as depicted in Figure 6.

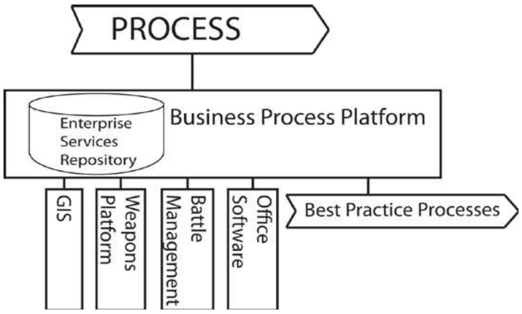


Figure 6. Busines Process Platform as a service enabler.

SOA can be utilized in reorganizing the military organization when casualties affect the command chains of the military organization. This is described in [34]. In Network Centric Warfare one key aspect is to offer a valid and accurate Common Operational Picture (COP) for the operating troops in the battlespace. A basic requirement for a combat leader is to have a possibility to lead the troops and gain good COP. An important aspect in distributing information in the battlespace is the amount and quality of information shared at different levels. A squad leader needs only basic information of the task while the battalion leader needs the information in a much wider scope and rearranged. The amount of information allocated needs to be set to a level where the person responsible for making decisions is able to perform timely and make accurate decisions.

SOA offers a way to organize this system in the form of versatile and varying types of processes and services. One terminal can use a number of services over a network to gain the information needed. All the terminals also act as services themselves and this creates a network of various services. The strong point of this system is that it is not dependent on any node of the system and in ideal case every terminal can change the current level. Each Warrior type has to be able to act as a leader at the next hierarchy level. For example, a Basic Warrior has the capability to act as a squad leader if needed. The terminals are also able to discover the services dynamically, which is an important aspect in a versatile and changing system in the form of databases and UDDI - registries to each node.

A constructive idea in SOA is its process ideology. In a military system the composition of the unit together with its performance are in a key role in executing the operations. Military units suffer from casualties and their performance value tends to change in an unpredictable manner. A military organization can be described at different abstraction levels seen in Figure 7.

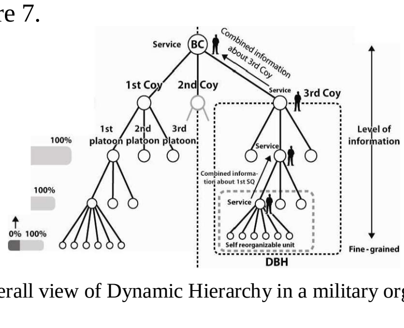


Figure 7. Overall view of Dynamic Hierarchy in a military organization.

Psycho Physical Factor can be calculated by using varying and measured variables of individuals' physical state. These values can be identified as, for example, heart rate, sweep of sweat and hormonal activity. Creating this factor is described in [26] and the formula can be useful in calculating the performance of a military unit. Varying variables can be measured via a wrist embedded computer. Obviously, to gain reliable values of each soldier and his or her current state of health, a monitoring period of at least six months is required in advance before starting field tests. Figure 8 features the formula created in [34], which describe calculating Psycho Physical Factor and accounts for command and control relations in a military organization.

As demonstrated in [9], human variables are important in the quest for answers to improve the soldier performance, including the measuring of stress, sleep deprivation, fatique and mental and physical trauma together with information overload, mission duration and overall energy balance. In [9] the Bayesian network model of Soldier performance was utilized. In addition, the formula for Psycho Physical Factor is relevant when calculating the performance values for military units and organizations varying in size (team, platoon, company, battalion). Here the Dynamic Battlefield Hierarchy plays an important role.

Figure 9 represents the increasing abstraction level in Dynamic Battlefield Hierarchy. Dynamic Battlefield Hierarchy is utilized in organizing and re-organizing military units. The performance of soldiers can be calculated, and as a result, the performance level of a whole military unit can be determined. When the performance is calculated, this value can be inserted as a mathematical value into different types of decision making processes. The following Figure 9 introduces a simplified view of the performance of a single military unit. In this figure the stress level of each soldier has been taken into consideration as well as the number of battles they have encountered and participated. Their recovery times have been calculated together with the stress level and the level of missing soldiers from the organization. As a result, each military unit (team, platoon, company, and battalion) can be identified with a performance value. The required power for each attack can be calculated beforehand once the enemy has been assessed first (the number of soldiers and amount of weaponry). The system described in Figure 8 was validated and laboratory tested as described in [34].

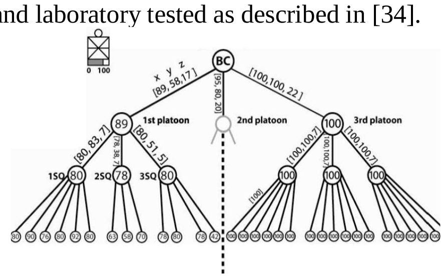


Figure 8. Calculating the performance of a military unit [27].

VI. Business Process and Resource Manager

The orchestration of Business Processes requires a tool for allocating resources, the Resource Manager (RM). The tool has been described in [33]. The RM sorts out and lines up the requested services. As militaries implement the framework of Network Centric Warfare with a continuing need to automate the Command and Control (C2) tools utilized in military, the ever-increasing tempo of operations must be taken into consideration. The collected data need to be processed, analyzed, verified, transmitted, and finally stored. SOA can be identified as a technology that can satisfy these needs of network centric operations. The starting point in the BP approach to SOA is that the main business operations of the organization are described by SOA BPs. The Business Processes are chains of logic that request SOA services. As regards military settings, the Business Processes represent military operations as explained earlier and depicted in Figure 6.

SOA-technology involves assisting, planning and executing events performed in military operations. Business processes are executed on a specific Business Process platform. Services and platforms, including geographical information-services, weapons platforms, and Battle Management Systems, are linked to the Business Process Platform to obtain optimal results. When a Future Force Warrior can benefit from the possibilities offered by a successful adoption of BP and SOA, the result can be improved overall performance in military operations.

Figure 9 describes how the Business Process approach can improve the performance of an FFW. Several battlefield sensors gather data from the battlespace. The collected data are then automatically transmitted to be analyzed in a command post. Various battlefield sensors transmit data to a context-aware reasoning layer. In this layer, data are converted to context and an inference engine transmits the data to a ubiquitous main layer for analyzing purposes. The data are verified, analyzed and transmitted as information for the execution of the operation.

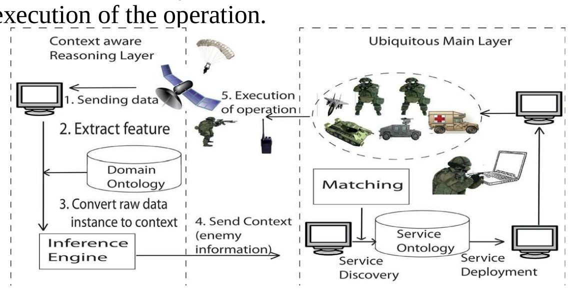


Figure 9. Increased FFW Performance can be gained via successful data utilization and analyzing processes.

Several of the needed services require real-time resources. These services can be identified, for example, as collecting SA data and issuing Fire Support Orders. Thus, the services and their use must be scheduled and sequenced to sustain the processes. The RM sorts out and lines up simultaneous requests concerning the requested service. The RM serves as the element, which provides the needed services for User Groups (UGs). Services can be either pre-programmed on demand or be available on request basis. The RM as a tool is located at the battalion level. The user groups send a request for the demanded service. The UGs are then authenticated, their privileges verified, after which the request is transmitted to the RM. The key functions of the RM are: 1) to receive the request of a required service, 2) to organize the line of user groups in the correct order depending on the UGs' privileges and battle-situation, 3) to check whether the service is available and within range, 4) to provide the User Groups with the answer, which is either the requested service or a rejection of the service.

The RM functions as a fully automated chain of functions in certain processes [33]. The key function of the person in the loop is to monitor the flow of events and to interfere with the chain of events if an unpredicted anomaly occurs in the process. As the RM is a critical resource, it must be physically protected against enemy actions.

The role of the RM is central in the allocating of resources in the BP. The RM communicates with four intermodules. The RM graphical user interface provides the core interface between all the presented modules and the Local Area Network (LAN), as shown in Figure 10. The LAN is utilized as a battlefield network or a community network as it can be used in a wide area of networks. However, the sharing of networking environment and its resources remains challenging in that searching for information and asking for resources turn out to be challenging when lacking proper search mechanisms. Each module has pre-defined and precise functions. First, the file and resources sharing module communicates with the RM GUI in conjunction with the sharing and the download module. The file and resource transfer and download module supports and enables the transfer or download of the searched file or resource from the other node connected to the network. The shared files and resources are listed on the RM GUI, where the listed and downloaded files can be examined. It is obvious that the same identified services are requested simultaneously. Therefore, the composition of the RM needs to be stable and reliable. Figure 10 illustrates the composition and function of the RM.

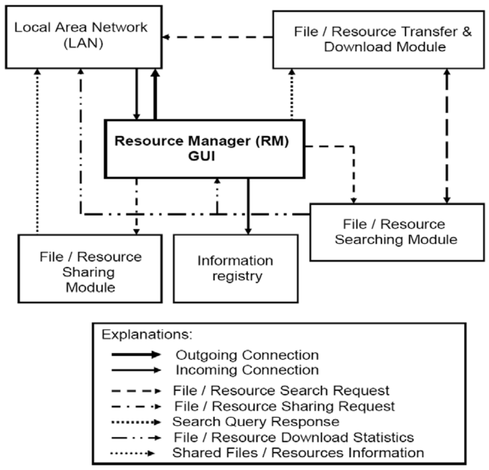


Figure 10. The composition and function of the RM.

The fire support system is essential in the battlespace. The fire support system requires an algorithm to support its optimized function as mentioned in [31]. The algorithm is used in the evaluation process of executing the firing mission. The critical features include the availability of the weaponry in question, its range, and its effect on a specific target and the status of the weaponry (i.e., in use/waiting for a fire support command). The example presented in Figure 11 depicts the processing of fire support order requests inside the RM as an informal Specification and Description Language (SDL) diagram. This action performed by the RM is essential to proceed in the process of offering requested service/s [33].

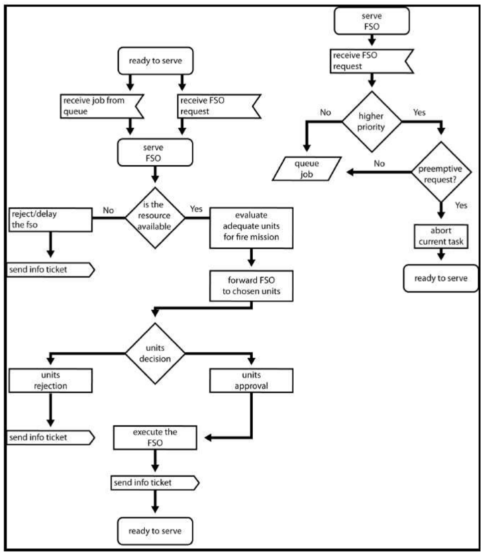


Figure 11. The processing of Fire Support Order requests in RM in an informal SDL diagram.

In a Fire Support Order the decision to execute the requested action, to fire, is made by a human being, an officer. In the future, when fire support decision algorithms are more fine-tuned, the decision to execute the mission can be forwarded to a Battle Management System. The BMS will then select the weapon and suitable ammunition for a specific fire mission. The BMS can also pre-evaluate the amount of ammunition type needed (how many shells it takes to destroy the target).

Each request comes with a time-stamp and own identification, contains route data and is traceable whenever tracking data are required. Each request is categorized according to an urgency class and its execution process is continuously monitored and evaluated. Once the request has been executed, it will be filed as a completed task in the common database. The tracking data of the completed request can be retrieved for analyzing purposes at any time by the system operator.

From the perspective of SOA and the Resource Manager, the described systems enable improved capability to align the listed and noticed services. When necessary services can be recognized and their state and phase have been identified, it may be possible to efficiently utilize own resources, for example, in a dismounted company attack. The process of a dismounted company attack was depicted earlier in Figure 2.

Operating in a military setting requires organizing matters related to command and control. It is mandatory to be able to sustain the composition of troops while soldiers get killed, wounded and missing in action. Dynamic hierarchy and performance value calculation were described earlier in Figure 7 and Figure 8. The orchestration of command chains and troops require reliable performance from SOA and the RM, but also a reliable means of communication and ubiquitous communication networks to support selected frequencies and used waveforms and bandwidths.

Although the requested services can be processed with the listed tools (Resource Manager and Scheduler) and outlined methods of organizing services (Business Process – like actions assisted with SOA), the system still remains incomplete because it lacks a key tool, a grid that takes into account all the relevant and constantly changing variables, such as the planned duration of an operation, the type of terrain where the operation is taking place, the expected troop performance and the capabilities of a specific military unit. None of the expected performance or capability values have been determined, evaluated or inserted as part of any Battle Management System.

VII. COMMUNICATION AND NETWORKS

The significance of functional communication systems and appropriately planned networks is substantial. To execute effective military operations as part of Network Centric Warfare, communication and network systems have to be carefully tailored in order to ensure that the services used can be located, identified and orchestrated together with SOA and with the assistance of the RM. Allocating and resourcing services becomes impossible without ubiquitous and secured network and communication systems.

The primary function of a communication system is to support Command and Control and support those actions. The main objectives of the systems involve providing Common Operational Picture of the battlespace in near-real time and sharing data among the battlespace systems. The systems utilized facilitate the fusion and display of intelligence information to commanders at all levels and handles the exchange of targeting data from sensor to weapon systems. This loop is called a loop from sensor-to-shooter-loop; the shorter the loop, the more effective the weapon system. Communication systems used have to support various waveforms, frequencies, transmission protocols, offer adequate bandwidth and Quality of Service (QoS). SOA can be seen as an organizing tool of systems utilized. Figure 12 explains the network system from sensor to shooter.

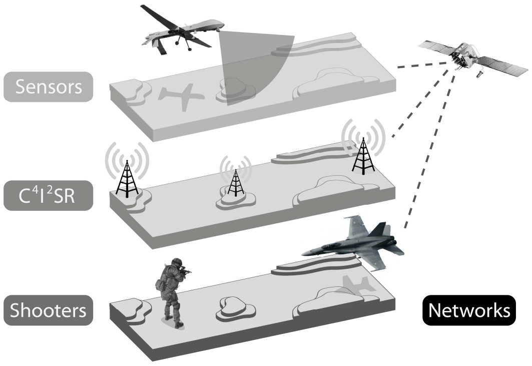


Figure 12. Network from sensor to shooter.

The speed and versatility of military operations keeps increasing unpredictably. Asymmetric and hybrid warfare set challenges for maintaining the ubiquitous networks in battlespace. If the Resource Manager does not identify the existing service, it can neither offer nor allocate it to any consumer (soldier).

The amount and versatility of collaboration tools also increases in the battlespace. Commanders' digital assistants of several types require constant and reliable connectivity to a functional network system with adequate bandwidth and the possibility to transmit data in different formats and also live footage from the battlespace. Different collaborating entities benefit from various tools and programs. Interfaces between military organizations can be challenging. A battalion commander and Company Commander use different tools for command and control. The used tools may rely on different frequencies and waveforms. The user challenges that surface are versatile when military units cooperate with civilian authorities or with Non-Governmental Organizations, when waveforms and tools may vary to a great extent. These challenges have been discussed in [35].

The concept of Network Centric Warfare has been created to support agile and versatile military operations in rapidly altering battlespace. To offer, allocate and trace services requires a comprehensive network system to ensure the ability to use these services. Without a ubiquitous and reliable communication network system, the execution of a successful military operation is endangered and likely to fail. In modern warfare communication services cannot be constantly produced by the performance of military satellites. Ubiquitous network systems require simple and local solutions in military operations executed in a low level (company and below).

Recent developments in modern military battlefield systems greatly influence today's military commanders and their performance. Modern devices such as smartphones are today supported by a complex infrastructure that enables high-speed communications and access to voice and data services anywhere, anytime. This is true only in theory. Constant communicating capability is hampered by latencies, disrupted connectivity in electromagnetic spectrum. Typical tactical operations benefit from the systems based on Local Area Communications (LAN). These systems are increasingly being employed to keep platoon and company level soldiers and assets connected with one another and their battle-group commander. Voice communications are being replaced with secure data connections as latest developments in the technology sphere indicate [2]. Moreover, communication systems based on Internet Protocol (IP) suite of products can be used to create a complete Wide Area System (WAS) with a number of Local Area Networks (LAN) in order to provide the required services for the consumer of a modern tactical communication network system. The maximum number of users in this type of network system is claimed to be 300 [2].

An example of typical soldier-worn personal radios are compact, cost effective and military graded transceivers that ensure secure and reliable voice and data communications for squad teams and higher echelons in many tactical scenarios, such as dismounted, vehicular, naval and amphibious operations. New radios are embedded with GPS/GLONASS receivers for automatic position reporting. Personal radios offer lightweight and low consumption transceivers operating in the frequency of 2.4 GHz Industrial, Scientific and Medical (ISM) band with a very low Radio Frequency (RF) signature modulation scheme based on spread spectrum technology (DSSS) and providing a robust, reliable and low probability of detection time division multiple access (TDMA) waveform. Radios typically feature advanced encryption standard (AES) encryption providing a very high security level on the transferred audio and data systems, while security keys can be downloadable. Personal tactical radios can provide several tactical communication services, such as full-duplex voice conferencing, GPS reporting, e-mail, chat, file transfer, as well as real-time video streaming [2]. On top of these systems, some of these can be seen as a wireless extension for an Intercom System and Combat Net Radios; they allow users to be connected to the vehicles and their main radio equipment via independent push-to-talk (PTT) selection. Software Defined Radios (SDRs) are on their way to tactical military environment. Typical SDRs' features include: Multiband 30 – 512 MHz, multimode, multi-mission, software programmable architecture, Low Probability of Detection (LPD) and Low Probability of Identification (LPI), simultaneous voice and data, near-real time data transfer for sensor to shooter applications and lastly integration to Strategic Communication Systems [2].

VIII. THE USE OF UNMANNED VEHICLES

The possibilities to benefit from Unmanned Vehicles (UVs) as relay-stations and sensor nodes deserve to be considered. Issues such as Low Probability of Detection and Low Probability of Identification have to be taken under close study and evaluation. When communication systems and services can produced with the assistance of UVs, they are easier to replace and the re-routing of the lost connection or connectivity can be arranged with a better probability compared to a satellite-based communication system of a higher echelon.

A typical UV can be a dispensable robot capable of overcoming the communications problems that soldiers encounter in built-up areas. Robots are usable in land and at sea, or air-borne, representing Unmanned Aerial Systems (UAS) or Unmanned Aerial Vehicles (UAVs) [36], [37]. Robots are utilized to reconfigure the prevailing communication infrastructure enabling the soldiers to stay connected with their lines of communication in versatile and altering urban communication battlespace. Throwable robots can act as an assisting tool in gathering data from the harsh and hostile battlespace [2]. It is crucial to understand the challenges posed by network connectivity constraints in the formation and reconfiguration of the network system in the environment, which UVs operate. The problem of formation control with network connectivity is discussed in [38].

SOA can serve as a tool in military operations in organizing and orchestrating networks and their topologies and in the process of re-routing the connection between users. Challenges related to Persistent Tactical Awareness (PTA) and Persistent Tactical Surveillance (PTS) [2] can be more easily solved if SOA and UVs can be utilized in gathering data and transmitting these data from the sensor to the shooter in an agile manner. Moreover, Command Posts create and update their Situational Awareness and Common Operational picture on the basis of the raw data produced by sensors embedded in the battlespace. Functional and reliable communication remains critical in Network Centric Warfare.

And lastly, communication networks have to be pre-planned to support sensor-throwable and sensor-shootable elements in different types of sensors. An example of the latter ones is described in [38]. Figure 13 depicts the functions of the Sensor Element Munition divided in different phases. When the data can be gathered beyond the horizon discreetly, the possibility of a success of own dismounted company attack is increasing.

The data collected with the assistance of Sensor Element Munitions can be used as real-time data to be inserted in the decision making process after the significance and reliability of these data have been evaluated and verified.

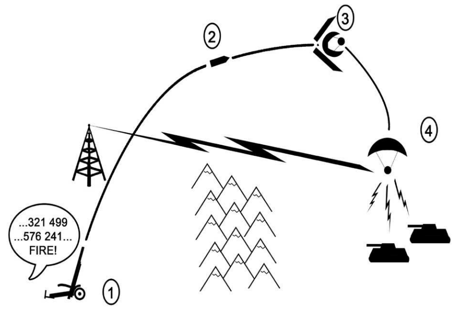


Figure 13. The usage of Sensor Element Munition.

IX. GRID FOR INTERFACE ASSISTED DECISION MAKING (GIADM)

As noted, neither the expected performance values nor the capability values of any military units have so far been calculated, assessed or inserted in any Battle Management System. The following discusses a preliminary idea concerning how to account for expected troop performance as a determined value together with time and terrain. The sketched Grid for Interface Assisted Decision Making (GIADM) exploits the data offered by means of SOA, the RM, and BP, and does so by using only one solution, which maximally exploits the results of all the systems. This type of an assisting tool needs to be algorithm-based because of the infinite number of affecting variables in a given military setting. For end-user purposes, the tool needs to be screen- or GUI-embedded, depending on the level of operation and that of a given military commander in an organization.

The proposed tool, a GIADM, aims at facilitating the decision making process of a military commander and can be benefitted in all types of military operations (defensive, offensive, special and Military Operations Other Than War), and is calculated on the basis of gathered data to support an existing military operation during its early planning stage. As an assisting tool, IADM serves the planning phase of an operation at all the army levels and draws on critical information based on the gathered knowledge on both own forces and the adversary, its actions and strength. Critical information includes time, distances between objects and troops, weapons available and the variety and amount of different types of supplies. In addition, features related to the prevailing weather, type of troops used, number and type of vehicles, type and nature of terrain and the capability of troops to execute a commanded mission are identified and quantified to determine the possibilities to execute a commanded mission. The levels of varying set objectives or actions can be marked on GIADM as depicted in Figure 14. To utilize GIADM, Graphic User Interfaces have to support its use. GUIs and their performance capabilities are discussed in Section X.

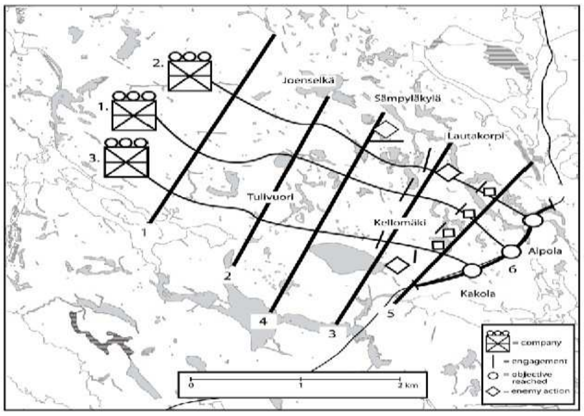


Figure 14. Set objectives marked on Grid for Interface Assisted Decision Making (GIADM).

When SOA is utilized to calculate the relative strength between military entities, the resources at hand may be optimized. This optimization means timing the usage of own resources, the amount of military force (troops, weapons, transportation, evacuation) available and its placement. When utilizing SOA and the Resource Manager in the planning process of organizing a military operation, the probability of a successful operation of a military maneuver may increase. This means saving own troops, fuel, ammunition, transportation vehicles and medical supplies.

These resources can be seen as services in SOA, the Resource Manager, and a resource or service in a Business Process. When all the parameters of the attack terrain, relative strengths of the entities (own, opposite force), amount of troops, different types of capabilities in performing operations are set in GIADM as variables, GIADM displays the results as suggestions on how to perform the commanded operation, as depicted in Figure 15, which visualizes only the pace of an operation from the perspective of a single soldier. Moreover, the use of friendly artillery fire can be presented in Graphic User Interface as a factor a single soldier has to take into account when performing a commanded task. This feature of the GIADM may aid in attempting to minimize instances of fratricide.

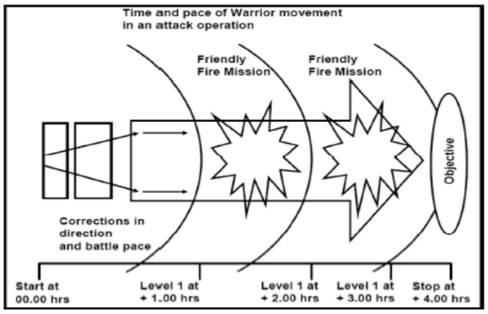


Figure 15. Suggestion of an operation, as visualized in GIADM.

X. GRAPHIC USER INTERFACE

Graphic User Interfaces are vital visualization platforms in the digitized military battlespace, in which military operations are executed speedily in constantly changing situations. Graphic User Interfaces can be identified as organic tools in command and control systems in contemporary warfare at all the levels of command and control systems; the higher the level in an organization, the more sophisticated the GUI.

The utilization of SOA in Network Centric Environments requires that the usability of Graphic User Interfaces be taken into consideration. GUIs have to be simple tools that assist the execution of a commanded military operation as visualization platforms of Battle Management Systems. When everything functions as planned, BMS can be seen as a system used to combine all the sensors, entities and devices to be used by higher echelon. BMS enables C2 processes executed in the battlespace. BML can be used in the interaction between Command and Control systems and as tools for course of action planning and analysis. When BPs can be embedded together with the assistance of an appropriate Scheduler, a Graphic User Interface of a Future Force Warrior may look as depicted in Figure 16.

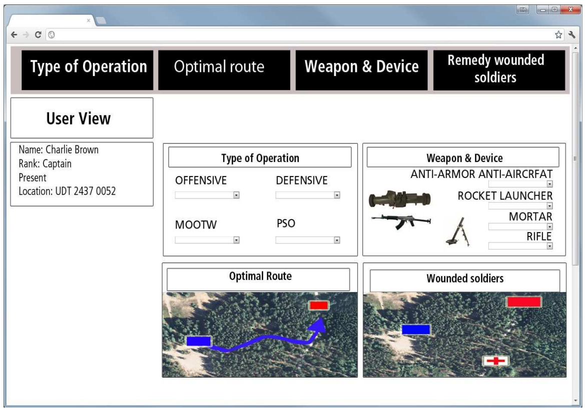


Figure 16. A view displayed by a Graphic User Interface.

GUI can serve both as a display and end-user device of a GIADM, especially at the lower command levels (company and below).

XI. SECURITY AND SCHEDULER

To account for operational security, there are protocols to identify the credentials of the requester entity by applying a security, authentication and agreement tool embedded in the RM. Before any tasks are ordered to be executed or resources allocated for use, the task or resource request is processed via the described system, as presented in Figure 17. An incoming task passes through a preliminary phase, in which it is checked and identified. Once the task has been verified and approved and transmitted from a trusted and secure cooperation entity, it will be processed via a series of approval and authorization policies.

Security issues remain critical also when dealing with unmanned aerial vehicles utilized in Network Centric Warfare at tactical level as described in [30]. The collected data have to be secured to remain intact and coherent when passing through different interfaces from the sensor to the shooter.

The described process ends with a phase in which the common language and tools are selected and then the given request moves forwards inside the RM. The overall description of the whole concept consists of three major parts and functions: 1) SA comprehending the existing solutions and tools, 2) command and control tools, and 3) information repository. These three together enable the command and control process and saving of logdata for further analyses. These functions presuppose the RM and the Scheduler to share and distribute the tasks and resources.

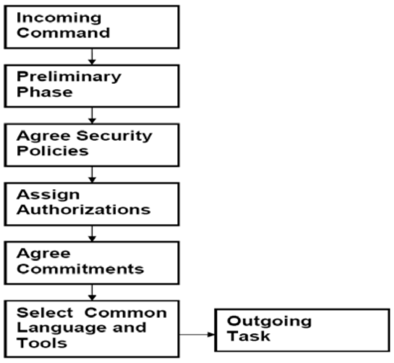


Figure 17. Security, authentiction and agreement system.

To provide the requested service, the RM requires one more component [34]. This critical component for the military use of SOA, which relies on the utilization of the RM is called the Scheduler. The role of the Scheduler is to coordinate processes to maximize the performance of resources and to reduce fratricide and collateral damage. The Scheduler enables militaries to execute various operations simultaneously but still under a strict command and control. The issue of simultaneously operations is solved by the element named Battlefield Secure Scheduler (BSS). This component uses two different methods of sharing calendar, Pre-Shared Scheduler (PSS) and Dynamic Schedule Update (DSU). The Scheduler functions together with the RM and utilizes SOA as a process. These elements can be recognized in Figure 18, which introduces the process from an incoming command/task to an outgoing command/task.

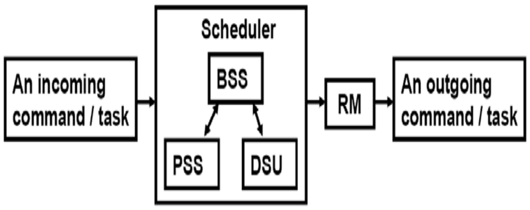


Figure 18. The elements inside the scheduler and the permeable command and control -process.

XII. BENEFITS AND DRAWBACKS

The delicate system introduced, the sketched Grid for Interface Assisted Decision Making exploits the data offered by means of SOA, the RM, and BP, but it does this by using only one solution, which exploits the results of all the systems. This means that the system can malfunction for various reasons. Challenges related to sustained energy supply have to be solved to enable the function of different processes. The orchestration of the system can also fail because of intentional enemy action (jamming, a virus, a worm). The system needs to be equipped with an analyzing program, which indicates when the system functions properly. If the system malfunctions and retrieving the capabilities becomes impossible, the system becomes useless for an FFW. This asks for an easily replaceable and fault-tolerant system with inbuilt check-in routines. Otherwise, traditional methods in orchestrating services need to be adopted.

By adopting SOA and embedding business processes into the existing command and control -system, the overall performance of military operations may be improved. With the assistance of the RM, limited military resources may be allocated more efficiently to the users requiring for services. When the Scheduler is implemented together with the RM into the BPs, it may be possible to increase the speed of decision making process together with the execution capability. The allocated resources available may then be used optimally. This means shorter execution times, and an increased amount of data for improved decision making. The overall system performance may be optimized with the assistance of these tools.

Offering a service of ubiquitous computing to battlespace commanders increases the possibility to utilize the resources available. This may foster a rapid decision making process especially when SOA can be embedded in the decision-making systems. As described in [22], SOA must deliver a solution that crosses existing boundaries as well as addresses the issues of information sharing regardless of where that information is stored.

By adopting these introduced elements into BMS together with SOA, it may be possible to gain improved capability to execute operations. This may also mean reduction in time needed for allocating resources and result in improved overall performance with minimized operations' execution times. Besides, with the improved level of SA, fratricide and collateral damage may possibly be reduced.

The system presented here is free to be adopted and tested. To create a functioning, automated business process applicable for future battlespace purposes requires development and testing, which accounts for matters related to adequate bandwidth, transmission power, connectivity, and operational security.

XIII. CONCLUSIONS

The introduced tool, GIADM, represents a follow-up idea-phase development stage in the quest for a comprehensive assisting decision making tool. GIADM combines gathered information from the existing resources and platforms connected to the Battle Management System. GIADM functions on top of BMS, SOA, RM and other possible tools for assisting in the decision making process and location of troops.

Synchronizing a sustainable performance of military troops, achieving the set objectives, optimizing the use of unmanned vehicles and timing of operations remain necessary. These in turn affect planning the resources to be orchestrated and allocated to minimize casualties and ensure short recovery times.

A commander has to be able to use an optimum amount of force. Because of the inevitable casualties, troops must be reorganized periodically in an unpredicted manner. SOA can be benefitted in the action of reorganizing the composition of troops and their resources. Resource Manager is integrated into SOA and Business Processes in order to assist the commander to target the reinforcement of troops and services to the area of operation in which they are mostly needed. The functions of Resource Manager take the use of Unmanned Vehicles into consideration. The data collected by robotic vehicles in inserted in the command and control system via Battle Management Systems of various types.

The data collected are then inserted in the tool labeled Grid for Interface Assisted Decision Making. After the analysis, a commander receives a suggestion how to execute an operation in relation to time, troops, terrain and the type of an operation with the objective to minimize the needed amount of supplies and services and thus save resources for the follow-up action.

The timely reorganizing of troops requires that the data collected have to be presented in an appropriate manner depending on the level of a military troop commander (platoon leader, company commander, battalion commander). Simple to use, affordable Graphic User Interfaces are in essential role in low level operations.

Militaries adopt digital systems created for assisting the process of decision making and operational planning. These systems continue to need further development in order to be fully adopted in operational use. When militaries continue to aim at optimized performance with the assistance of Service Oriented Architecture, the systems embedded in SOA and Business Processes have to be properly organized and orchestrated. The amount of interference between different systems has to be minimized with careful planning sequences already when systems are being designed and built. This ensures that the relevant data remain intact and uncorrupted. Unfortunately, neither the technical readiness level of these systems nor the capability to link these systems together is currently at an implementation level. Some of the challenges have been identified in this paper to contribute to the process of embedding digitalized automated systems in the battlespace.

XIV. Further Work

When the process of war has been modeled to resemble a Business Process, the performance of Future Force Warriors may be optimized with the assistance of processes assisted with SOA. The result of this may be seen as an agile and modular military performer with improved capabilities and improved Situational Awareness, and the capability to utilize ever diminishing resources more optimally with decreased instances of fratricide.

Further work related to modeling a war as a Service Oriented Architecture assisted business process must concentrate on the orchestration of systems. Challenges posed by operational security cannot be neglected, as they continue to disrupt digitized battlespace. Issues related to sustaining an adequate level of constant energy flow and protection against violations caused by electronic warfare must be studied, tested and solved before adopting GIADM in operative use.

Both extensive funding and considerable human resources are required to run the extensive number of validity tests necessary, which would result in getting all the systems up (i.e., BP, RM, Scheduler) running synchronized with GIADM. Firstly, tens of thousands of simulation laps involving each operational scenario type, including defensive, offensive, and Special Operations, are required before implementing the introduced system/prototype into any real-time military exercise performed. Secondly, once any resources have been allocated for and invested in implementing a GIADM type of approach to complete a functional command and control tool that assists in decision making, the system design can no longer be accessed in any public domain data sources.

References

[1] T. Saarelainen, “Modeling war as a business process with the assistance of service oriented architecture,” ICDT 2014, The Ninth International Conference on Digital Telecommunications, February 23 – 27, 2014 Nice - France, pp. 1 – 7.

[2] M. Abdullah and D-P. Merklinghaus, “Mobile battlefield systems,” Vol XXXVIII, Issue 6/2014, Military Technology, ISSN 0722-3226, pp. 54 – 62.

[3] L. Jianying and L. Xiaolong, “Artificial intelligence in modeling command and control,” Second International Conference on Computer Modeling and Simulation, 22 – 24 January. 2010, Sanya, Hainan, pp. 383 – 386, Print ISBN: 978-1-4244-5642-0, doi: 10.1109/ICCMS.2010.449.

[4] J. Scholz, D. Lambert, D. Gossink, and G. Smith, “A Blueprint for command and control: automation and interface,” 15th International Conference on Information Fuison (FUSION), 9 – 12 July 2012, Singapore, pp 211 – 217, Print ISBN: 978-1-4673-0417-7.

[5] J. Du, Q. Liang, R-J. Xu, and Y-c. Xia, “Research of firepower allocationbase on battlefield situation perceive in aggregated CGF,” International Conference on Computer Science and Electronics Engineering (ICCSEE), 22 – 25 March 2012, Hangzhou, pp. 523 – 526, Print ISBN: 978-1-4673-0689-8, doi: 10.1109/ICCSEE.2012.337.

[6] S. Shaobin, L. Xiehua, P. Jun, and Z. Yingchun, “Decision-making and planning of computer generated forces,” 2012 IEEE International Conference on Computer Science and Automation Engineering (CSAE), 22 – 27 May 2012, Zhangjiajie, pp. 476 – 480, Print ISBN: 978-1-4673-0088-9, doi: 10.1109/CSAE.2012.6272641.

[7] S. Guojun, M. Xiaodong, and Z. Haijing, “Research on behavior modeling method of agent-based CGF,” Journal of Software, Vol 7, No 4 (2012), 913-918, April 2012, pp.913 – 918, doi:10.4304/jsw.7.4.913-918.

[8] J. From, P. Perrin, D. O´Neill, and J. Yen, “Supporting the commander’s information requirements: automated support for battle drill processes using R-CAST,” Military Communications Conference 2011 (MILCOM2011, 7 – 10 November 2011, Baltimore, MD, U.S.A., pp. 1523 – 1528, ISSN: 2155-7578, doi: 11.09/MILCOM.2011.6127523.

[9] C. Ortega, T. Brown, J. Ibbotson, and R. Hancock, “Improving WSN application QoS and network lifetime management using SOA strategies,” Military Communications Conference 2011 (MILCOM2011, 7 – 10 November 2011, Baltimore, MD, U.S.A., pp. 1580 – 1585, ISSN: 2155-7578, doi: 11.09/MILCOM.2011.6127534.

[10] K. A. McFarland, “A performance map framework for maximizing soldier performance,” Doctoral Dissertation, Dept. of Mechanical Engineering, United States Army, http://hdl.handle.net/2152/ETD-UT-2011-05-3201, accessed on November 5th 2014 at 8.32 am.

[11] W. T. Tsai, Q. Huang, J. Xu, Y. Chen, and R. Paul, “Ontologybased dynamic process collaboration in Service-Oriented Architecture,” in Proceedings of IEEE International

Conference on Service-Oriented Computing and Applications (SOCA'07), 19-20 June 2007, Newport Beach, CA, pp. 39 – 46, doi10.1109/SOCA.2007.35.

[12] P. Lalbakhsh, N. Sohrabi, and M. N. Fesharaki, "The role of Service Oriented Architecture in battlefield situational awareness," International conference on Computer Scientific and Information Technology, pp.476 – 479, ISBN 978-1-4244-4519-6.

[13] M. Panahi, W. Nie, and K-J. Lin, "The design and implementation of service reservations in Real-Time SOA," in Proceedings of IEEE International Conference on e-Business Engineering (ICEBE2009), 21-23 October 2009, Macau, pp. 129 – 136, doi 10.1109/ICEBE.2009.26.

[14] A. Gravel, X. Fu, and J. Su, "An analysis tool for execution of BPEL services," in Proceedings of the 9th IEEE International Conference on E-Commerce Technology and the 4th IEEE International Conference on Enterprise Computing, E-Commerce, and E-Services (CEC-EEE2007), Tokyo, pp. 429 – 432, doi 10.1109/CECEEE.2007.19.

[15] M. Jiang and A. Willey, "Service-Oriented Architecture for deploying and integrating enterprise applications," in Proceedings of 5th Working Conference on Software Architecture (WICSA 2005), Pittsburg, PA, pp. 272 - 273, doi 10.1109/WICSA.2005.60.

[16] G. Lewis, E. Morris, and D. Smith, "The Service-Oriented Migration and Reuse Technique (SMART)," Technical report CMU/SEI-2005-TN, Software Engineering Institute, in Proceedings of 13th International Workshop on Software Technology and Engineering Practice, September 2005, Budapest, pp. 222 – 229, doi 10.1109/STEP.2005.24.

[17] M. Medlow, "Extending Service-Oriented Architectures to the deployed land environment," Journal of Battlefield Technology, vol. 13, No. 1, March, 2010, pp. 27 – 33.

[18] L. Liu, D. Russell, N. Looker, D. Webster, and J. Xu "Delivering sustainable evolutionary Service-Oriented Architecture for network enabled capability," in Proceedings of International Workshop on Verification and Evaluation of Computer and Communication Systems (VECoS2008), pp. 1 – 11.

[19] F. Hojaji, M. Reza, and A. Shirazi, "Developing a more comprehensive and expressive SOA governance framework," in Proceedings of the 2nd IEEE International Conference on Information Management and Engineering (ICIME), 16-18 April 2010, Chengdu, pp. 563 – 567, doi 10.1109/ICIME.2010.5478046.

[20] K. Lund, A. Eggen, D. Hadzic, T. Hafsoe, and F.T. Johnsen, "Using web services to realize service oriented architecture in military communication networks," Communications Magazine, vol 45, Issue 10, pp. 47 – 53, doi 10.1109/MCOM.2007.4342822.

[21] J. He, I-L. Yen, T. Peng, J. Dong, and F. Bastani, "An adaptive user interface generation framework for web services," in Proceedings of IEEE Conference on Congress on Services Part II, 2008. SERVICES-2, Beijing, 23 – 26 September 2008, pp. 175 – 182, doi 10.1109/SERVICES-2.2008.23.

[22] E. Zeeb, A. Bobek, H. Bohn, and F. Golatowski, "Service-Oriented Architectures for embedded systems using devices profile for web services," in Proceedings of IEEE 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW '07), 21-23 May 2007, Niagara Falls, Ont., pp. 956 – 963, doi 10.1109/AINAW.2007.330.

[23] J. Jormakka and J. Lucenius, "Possibilities for improving dependability of C4I2SR systems based on Service Oriented Architecture," in IEEE Transctions 2009 Computation World: Future Computing, Service Computation, Cognitive, Adaptive, Content, Patterns, pp. 315 – 324, doi 10.1109/ComputationWorld.2009.74, ISBN: 978-1-4244-5166-1.

[24] B. Farroha and D. Farroha, "SOA as a catalyst to empower the warfighter through improved enterprise data access over the GIG," 3rd Annual IEEE Systems Conference (SYSTEMS2003), 23-26 March 2009, Vancouver, BC, pp. 48 – 53, doi 10.1109/SYSTEMS.2009.4815770.

[25] Sagem: "Showcasing systems for demanding infantry forces mission," www.soldiermod.com, Vol 11 Summer Autumn 2013, pp. 42 – 43.

[26] M. Phillips, "Air-to-Ground and ground-to-air communications," Military Technology, Vol XXXVII, Issue 6/2013, pp. 66 – 67.

[27] P. Donaldson, D-P. Merklinghaus, and S. Nitschke, "Technology enablers for global special forces capability," Military Technology, Vol XXXVII, Issue 5/2013, pp. 65 – 74.

[28] P. Cordwainer, "Grid computing on the battlefield," Military Technology, Vol XXXVII, Issue 4/2013, pp. 74 – 75.

[29] J. Antal, "Enhancing squad communications – seven considerations for the militarisation of COTS smartphones," Military Technology, Vol XXXVII, Issue 5/2013, pp. 58 – 61.

[30] D. Alexander, "Toward NATO forces 2020: Combat Multiplication Technologies for Smart Defence," Military Technology, Vol XXXVII, Issue 10/2012, pp. 84 –86.

[31] C. Colakoglu, "Tactical command and control systems and network centric warfare," Journal of Military and Information Science Vol 2, No 3/2014, pp.70 – 76.

[32] K. Rein, U. Schade, and M. R. Hieb, "Battle Management Language (BML) as an enabler," IEEE International Conference on Communications, pp. 1 – 5, ISBN 978-1-4244-3435-0, doi 10.1109/ICC.2009.5198779.

[33] T. Saarelainen, "Improving the performance of a dismounted future force warrior by means of C4I2SR," Doctoral Dissertation, ISBN 978-951-25-2457-0.

[34] T. Saarelainen and J. Timonen, "Tactical management in near real-time systems," IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support, (CogSIMA2011), Miami Beach, Florida, 22.–24 February 2011, U.S.A., pp. 240 – 247, 10.1109/COGSIMA.2011.5753452.

[35] T. Saarelainen and J. Jormakka, "Interfacing collaboration and command tools for crises management military command and control systems," International Journal of Electronic security and Digital Forensics, vol 3, No. 3, 2010, pp. 249–264.

[36] J. Jormakka and T. Saarelainen, "UAV-based sensor networks for future force warriors," International Journal On Advances in Telecommunications, vol 4, numbers 1 and 2, 2011, ISSN: 1942-2601, pp. 58–71.

[37] L. Navarovong, Z. Kan, J.M. Shea, and W. E. Nixon, "Formation reconfiguration for mobile robots with network connectivity constraints," IEEE Network, July/August 2012, pp. 18 – 24.

[38] T. Saarelainen, "Targeting situational awareness beyond the event horizon by means of sensor element munition," ICDT 2012, The Seventh International Conference on Digital Telecommunications, April 29 – May 4, 2012, Chamonix / Mont Blanc, France, pp.8 – 14.