CIA2ST'2022
2-3, Juin 2022
Beni Mellal, Maroc

Proceedings of the International Symposium on Artificial Intelligence and the Security of New Telecommunications Systems

CIA2ST'2022

Sultan Moulay Slimane University
Polydisciplinary Faculty
Beni Mellal, Morocco

June 2-3, 2022

Editor

Pr. Said SAFI

International Symposium on Artificial Intelligence and the Security of New Telecommunications Systems (CIA2ST'2022)

(CIA2ST'2022)

Editor

Said SAFI

Beni Mellal, Morocco **Jun 2-3, 2022**

Prologue

The main objective of this symposium is to allow interested parties to better discover current advances in fields related to Artificial Intelligence and the Security of new Telecommunications Systems. It will offer specialists the opportunity to establish privileged contacts in the scientific community, in particular with other laboratories and research teams.

In addition, this international symposium will allow French-speaking researchers to discuss their work and initiatives; and will be able to give doctoral students the opportunity to gain a broad overview of their field of research and to benefit from a first contact that is both rigorous and benevolent with all the related activities. Finally, it will also be an opportunity to share on the perspectives and projects in progress.

Honor Committee

Nabil Hmina: President of Sultan Moulay Slimane University, Beni Mellal Morocco.

Benachir El Haddadi: Dean of the Polydisciplinary Faculty, Beni Mellal Morocco.

Belaid Bouikhalene: Head of LIMATI Laboratory

Chairman

Pr. Said SAFI
LIMATI Laboratory, Polydisciplinary Faculty, Sultan Moulay Slimane University
Beni Mellal, Morocco
safi.said@gmail.com

Keynote Speakers

Mohamed Kayyali : University of California UCSB, USA.
Youssef Naimi: Hassan II university, Casablanca, Maroc.
Miloud Frikel: Caen University, Caen France.
Belaid Bouikhalene: Sultan Moulay Slimane University, Maroc
Yassine Sadqi: Sultan Moulay Slimane University, Maroc.

Organizing Committee

Abderazak Farchane
Abdessamad Ousarhane
Anouar Darif
Aziz Laaribi
Ahmed Bahlaoui
Hicham Ouchitachen
Idriss Ellahiani
Manaut Bouzid
Nadia Barj
Nouredine Falih
Rachid Bahloul
Said Hakimi
Souad Taj

Souâd Touhami
Yassine Sadqi

Scientific Committee

Abdelouhab ZEROUAL, Cadi Ayyad University, Morocco,
Abderrahim Kamel, National Engineering school of Gabes, Tunisia
Abdessamad MALAOUI, Polydisciplinary Faculty, Morocco,
Ahmed BOUMEZZOUGH, Polydisciplinary Faculty, Morocco,
Ahmed Bahlaoui, Polydisciplinary Faculty, Morocco
Ahmed NAIT-SIDI-MOH, France,
Abderazak FARCHANE, Sultan Moulay Slimane University, Morocco,
Abderrazak LFAKIR, , Sultan Moulay Slimane University, Morocco,
Belaid BOUIKHALENE, Sultan Moulay Slimane University, Morocco,
Boubekeur TARGUI, ENSICAEN School, France
Carlos Manuel Astorga ZARAGOZA, Mexico Chantal
GÜNTHER, ENSICAEN School, France,
Eric PIGEON, ENSICAEN School, France,
Jilali ANTARI, Ibn Zohr University, MoroccoJoão
Paulo COSTEIRA, Portugal
Johannes REUTER, Germany,
Hicham MOUNCIF, LIMATI Laboratory, France,
Khalid AUHMANI, ENSAT School, Morocco Mathieu
POULIQUEN, LAC Laboratory, France Mohamed FAKIR,FST
Faculty Morocco,
Miloud Frikel, LAC Laboratory, France,
Mohammed HAMZAOUI, France,
Mohamed M'SAAD, LAC Laboratory, France,
Mohamed Zidane , Ibn Zohr University, Morroco,
Naima TAIFI, Sultan Moulay Slimane University, Morocco,
Noureddine FALIH, Sultan Moulay Slimane University, Morocco,
Omar EL KHARKI, FST Faculty, Morocco,
Olivier GEHAN, ENSICAEN School, France,
Samira CHABAA , Ibn Zohr University, Maroc,
Said SAFI, Sultan Moulay Slimane University, Morocco,
Yassine SADQI, Sultan Moulay Slimane University, Morocco

CIA2ST'2022 Chairman

Biography: Prof. Said Safi received the M.Sc. and the doctorate degree from Chouaib Doukkali University and Cadi Ayyad Univerrsity, in 1997 and 2002, respectively.

He has been a Professor of information theory and telecommunication systems at the National School for Applied Sciences, Tangier, Morocco, from 2003 to 2005. Since 2006, he is a Professor of applied mathematics and programming at Polydisciplinary Faculty, Sultan Moulay Slimane University, Beni Mellal, Morocco. In 2008 he received the Ph.D. degree in Telecommunication and Informatics from the Cadi Ayyad University.

In 2015 he received the degree of Professor in Sciences at Sultan Moulay Slimane University. His general interests span the areas of communications and signal processing, estimation, time-series analysis, and system identification subjects on which he has published more than 50 journal papers and more than 70 conference papers. Current research topics focus on transmitter and receiver diversity techniques for single- and multi-user fading communication channels, and wide-band wireless communication systems

Keynote speakers

Biography : Dr. Mohamed KAYYALI Scientist awarded by UK royal science council, former researcher at University of California UCSB at image processing labs, PhD from WCU, innovation management certified from Harvard Business School, chartered fellow in the UK, he published many conferences papers, authors of 5 indexed books, entrepreneur and IEEE industrial officer, UN and EU Int. speaker, CEO and co-founder of 4D Business Consulting, CEO at KSF Space Foundation USA, president at IFGICT International federation of ICT (ITU Partner), DNBC TV an IPTV and Time Business News co-founder.

Biography : Dr. Miloud FRIKEL is an Associate Professor at National Graduate School of Engineering and Research Center (ENSICAEN), and head of SATE's Department (Embedded Systems and Control), and he is the Deputy-Director of the Systems Engineering Lab of Normandy (LIS). He was with the R&T (Networks and Telecommunications) Department of the University of Caen (Normandy University). He received his Ph.D. degree from the Center of Mathematics and Scientific Computation CNRS URA 2053, France, in array signal processing. Dr. Frikel was with the Signal Processing Lab, Institut for Systems and Robotics, Institute Superior Tecnico, Lisbon, as a researcher in the field of wireless location and statistical array processing. And he worked in the Institute for Circuit and Signal Procesing of the Technical University of Munich, Germany. M. Frikel is member of German Foundation: Alexander von Humboldt Stiftung. His research interests span several areas, including statistical signal and array processing, cellular geolocation

(wireless location), direction finding and source localization, blind channel identification for wireless communication systems and MC-CDMA systems.

Biography : Dr. Belaid BOUIKHALENE obtained a Ph.D. degree in Mathematics in 2001 and a degree of Master in Telecommunication and Computer science in 2007 from the University of Ibn Tofel Kenitra, Morocco. He is currently a professor at University Sultan Moulay Slimane, Morocco, His research focuses on mathematics and applications, decision information systems, e-learning, pattern recognition and artificial intelligence.

Biography : Dr. Yassine SADQI received his MSc and his PhD in Computer Science with a focus on computer security from Ibn Zohr University, Faculty of Sciences Agadir, Morocco, in 2012 and 2015, respectively. From 2017 to 2021, he was an Assistant Professor at Sultan Moulay Slimane University. As of February 2021, he becomes an Associate Professor. Dr. Sadqi's is the vice-president of the African Research Center of Information Technology & Cybersecurity. He is a Senior Member of the IEEE, ACM Professional Member, and a Member of OWASP. Dr. Sadqi has made contributions in the fields of Web security, authentication protocols, network security, and cybersecurity. He has published several peer reviewed research articles in international journals, book chapters, and conferences/workshops. Furthermore, He has served and continues to serve on executive and technical program committees and as a reviewer of numerous international conferences and journals.

Biography : Dr. Youssef NAIMI Born in 1966 in Casablanca. He obtained a bachelor's degree in experimental sciences in 1985 at the Moulay Abdellah high school in Casablanca. He obtained a license in Chemistry at the Faculty of Sciences of Ain Chock Casablanca in 1989. Then, he obtained a DEA in electrochemistry at the Pierre and Marie Curie University in Paris. In the same university, he obtained a doctorate (Ph.D.), thesis: Optimization study of the hydrogen electrode of fuel cells with proton exchange membrane. In 1996, he joined university teaching at the Faculty of Sciences Ben M'sik, Casablanca. He continued his research work on H2/O2 fuel cells in collaboration with the Cnam (Conservatoire National des Arts et Métiers) in Paris and defended a State doctorate in 2006. He coordinated several research projects, supervised twenty doctoral theses. He is the coordinator of the Specialized Master in Renewable Energies and Materials. He is head of the Chemistry Department at the Ben M'sik Faculty of Sciences, since January 2020.

Special session chairman biography

Biography : Prof. **Abdessamad MALAOUI** d'Enseignement Supérieur au sein de l'Université Sultan Moulay Slimane (Maroc). Docteur de l'Université de Provence (France) et Docteur de l'Université Cadi Ayyad de Marrakech (Maroc). Spécialité et axes de recherche : Electronique, Automatique, Informatique Industrielle, Systèmes embarqués, systèmes photovoltaïques et Ingénierie des Énergies Renouvelables.

Biography : Prof. **Said Belaaouad** d'enseignement supérieur à la faculté des sciences Ben M'Sick de l'université Hassan II de Casablanca. Directeur du Laboratoire de chimie Physique des Matériaux (LCPM). Ses axes de recherche portent sur l'étude des propriétés physiquo-chimiques, structures cristallines et valorisations de nouveaux phosphates condensés, ainsi que sur l'ingénierie de formation et didactique des sciences et techniques.

Contents

Keynote Talks

Oral and poster sessions

The IoS, Taking IoT Further With A.I. Via Small Satellites

Mohamed KAYYALI
University of California UCSB
Image Processing laboratory, USA
mullderjob2@gmail.com

Abstract:

It is not about artificial intelligent and decision-making support on Earth and all activities on ground but also it's about artificial intelligence and support your decision from space through small satellites, this is what we called internet of satellites the next revolution of Internet of things IOT in space. Satellites that are orbiting in the lower orbits such as nano satellites and microsatellites the future will be relaying on trajectory decision maneuvering decision telecommunication and transmission and detecting objects and many things with a real time decision from space.

Future industry and semiconductors oriented to Internet of things to relay on internet of satellites for a wider coverage and precision data that can be gathered from space while processing the real data from space and taking a decision is one of the most oriented topic targeting future industry such as navigation systems aviation transportation logistics Marine industry tracking systems motion detection and much more.

The mathematical concept behind deep learning

Belaid Bouikhalene
LIMATI Laboratoty, Sultan Moulay
Slimane University, Beni Mellal , Morocco.
b.bouikhalene@usms.ma

Abstract :

Several mathematical branches such as analysis, linear algebra, differential calculus, optimization and probability and statistics are behind deep learning. Indeed, modeling and learning based on these notions, from the development of the artificial neuron in 1943 to the implementation of deep learning algorithms in 2012.

From Single-Sensor Signal Processing To Multi-Sensor Signal Processing

Miloud Frikel

LIS Laboratory, ENSICAEN School,
Caen University, Caen France
miloud.frikel@ensicaen.fr

Résumé :

Depuis l'avènement des sciences, la plupart des chercheurs font ont fait du traitement du signal sans être des "traiteurs de signaux". En effet, faire de l'acquisition des données, les mettre en forme, les transformer permet d'extraire une information utile, c'est du traitement du signal. Si on rajoute la modélisation et des algorithmes rapides de calcul, on a cette discipline à part entière bien établie depuis les années 1960.

Le traitement du signal est au cœur des trois disciplines : La physique, les mathématiques et l'informatique. En effet, un signal est un support physique d'une information. Ces informations sont dans la nature ou issues de

réalisations technologiques. Les mathématiques ont pour rôle de modéliser ces signaux pour mieux les traiter et l'informatique permet un traitement rapide grâce à des algorithmes efficaces.

La discipline du traitement du signal a été reconnue comme telle depuis l'apparition d'ordinateurs et de micro-processeurs qui ont permis des traitements en temps-réel des différents signaux. L'exemple le plus parlant est le calcul de la transformée de Fourier grâce à des algorithmes rapides de la FFT.

Les différentes méthodes de traitement du signal sont, maintenant, mûres et bien reconnues. Cependant, avec la révolution numérique actuelle, il y a une quantité de données de plus en plus grandes à analyser et à traiter. En général, le traitement du signal consiste à un traitement mono-entrée/mono-sortie et par extension au traitement du signal multidimensionnel avec plusieurs entrées et sorties, les méthodes classiques ont montré certaines limites et nécessitent le développement de nouvelles techniques se basant sur l'optimisation par exemple ou de l'intelligence artificielle.

Cybersecurity challenges and issues in modern applications

Yassine Sadqi

LIMATI Laboratoty, Sultan Moulay Slimane University,

Beni Mellal , Morocco

yassine.sadqi@gmail.com

Abstract:

Currently, modern applications such as web-based apps, cloud-hosted apps, and mobile apps are without any doubt the most used means of communication and information exchange. Applications such as online banking, e-commerce, online blogs and social networking apps have become a common platform for the transmission of information and the provision of services online. However, the open nature of modern applications and their wide usage in delivering critical services made them a prime target of cyber-attacks. Moreover, the rapid evolution and advance in software technologies have made the structure and interaction between the client-side and server-side components of modern apps more and more complexes, which have led to several security issues. Hence, it

has become necessary to understand the core challenges and security issues of modern applications.

Storage Of Renewable Energies, Role Of Hydrogen

Youssef NAIMI
Laboratory of Physical Chemistry of Materials, FSBM, Hassan
II University of Casablanca:
youssefnaimi@outlook.com

Abstract:

Today, more than ever, humanity is confronted with the problems of pollution, energy management, and environmental protection linked to unbearable economic and industrial developments. Since 1945, revolutionary advances in the biological sciences had, have, and will continue to have a profound impact on the quality of life and development and several avenues have arisen to find alternative solutions. In addition, at the summit of world leaders, the notion of sustainable development, based on economic, social, and environmental development was adopted. The first definition of sustainable development (a development that meets the needs of the present without compromising the ability of future generations to meet theirs) was established in 1987. The United Nations Conference on Environment and Development, which was held in 1992 in Rio de Janeiro, Brazil (known as the "Earth Summit"), highlighted the critical role that renewable energy sources and technologies can play in helping to meet the double challenge of development and environmental protection. The United Nations General Assembly has adopted several resolutions endorsing the World Solar Program 1996-2005 as a contribution to the achievement of development goals. Within this framework, the World Summit on Sustainable Development (Johannesburg, South Africa, August 2002) set up a process aimed at promoting the use of sustainable and renewable energy sources to improve the living conditions of those who do not have access to conventional energy sources.

Today, the world community is aware of the role and importance of these

energies, particularly for sustainable development and the improvement of the living conditions of poor rural populations. However, renewable energies are intermittent; hence the need to find solutions for their storage, since their capture (production) is always delayed in time compared to their consumption (need). Energy savings and the rational use of energy (energy efficiency) are essential, particularly in the building sector (thermal insulation, solar water heaters, use of heat pumps, etc.). The greatest energy savings are possible in the transportation sector, as electric and fuel cell vehicles are much more efficient than current internal combustion engine vehicles. Oil is a major source of pollution that must be replaced or accompanied by other sources, respecting the environment. Of the many candidates, hydrogen appears to be the most promising. Hydrogen would have an important role in such an energy system, serving as a transport fuel, energy storage, and feedstock by producing methane. Methane is considered an energy carrier that could be used in existing natural gas infrastructure. Finally, in the more or less near future, hydrogen could also be used directly in all applications replacing natural gas and methane. In this work, we present the recent development of technologies renewable energies, energy storage, and the integration of the hydrogen chain.

Adequate modelling of COVID-19 pandemic based on time series analysis and Higher Order Statistics

S. Safi, M. Frikel[2], and M. Pouliquen[2]

[1] LIMATI Laboratory Polydisciplinary Faculty, Beni Mellal, Morocco, Email: safi.said@gmail.com
[2] LIS Laboratory, ENSICAEN, Caen, France Email: (miloud.frikel, mathieu.pouliquen)@unicaen.fr @ensicaen.fr

Abstract — In this paper we are focused on time series evolution describing the daily evolution of pandemic contamination detected in some countries in Europe, North America and Africa. Our goal is to elaborate a model which able to represent the pandemic data in each country. Some models have been developed based SIDARTHE dynamics model with more parameters. Other method is based on the Susceptible-Exposed-Infectious-Removed (SEIR) model to derive the epidemic curve. Another paper has presented an agent-based model for affine-grained computational simulation of the ongoing COVID-19 pandemic. In this paper we attempt to build our model able to represent the evolution of this epidemic in some countries in Europe in which the evolution is more critical in North America. Those models will give us an idea about the possibility to develop a universal model describing the evolution of the contamination caused by COVID-19. The developed models are used to simulate and predict the contamination in each country based on Higher Order Statistics and time series analysis technics.

Keywords — Pandemic contamination, COVID-19, Modelling, Time series analysis, Higher Order Statistics (HOS)

1. Introduction

The coronavirus COVID-19 pandemic is the defining global health crisis of our time and the greatest challenge we have faced since world war two. Since its emergence in Asia late last year, the virus has spread to every continent [1-6]. Cases are rising daily in Africa the Americas, and Europe. Countries are racing to slow the spread of the virus by testing and treating patients, carrying out contact tracing, limiting travel, quarantining citizens, and cancelling large gatherings such as sporting events, concerts, and schools. The pandemic is moving like a fire one that may yet destroy the countries least able to cope or take some erratic decisions like confinement and separation [7-13].

This paper deals with time domain statistical models [14-20] by analyzing time series representing the number of person infected by COVID-19 evolution of the pandemic in a given country [13]. In order to do this study we analyses the data at: stationary and nonstationary models, no seasonal and seasonal models, intervention and outlier models, transfer function models, regression time series models [21-28]. We discuss the process of time series analysis [14,17] including model identification, parameter estimation, diagnostic checks, forecasting. We also discuss how we can represent this pandemic by a generalized model representing this contamination regardless of the country from which the data is collected. The rest ofm this paper is organized as follow: in section 2 we represent the data base representing the pandemic in six countries collected between February and May.

2. Data analysis

2.1. Data evolution

The data representing the daily number of person contaminated by COVID-19 pandemic are collected from the website: https://www.coronavirus-statistiques.com/, in which they represent the daily cumulative infected person in different countries in the world.

In the Figure 1 the daily infected person, respectively, in (France, Germany, Italy, United Kingdom, United State of America and Morocco). Note that we will consider the data collected between February to May 2020.

The first step in the analysis of a time series representing the collected data is the selection of a suitable mathematical model (or class of models) for the data. For this we plot the data, if there are an apparent discontinuities in the series such as a sudden change of level which may be advisable, to analyse the series by first breaking it into homogeneous segments. Inspection of the graphs represented in the Figure 1 describing the number of person infected by COVID-19 in countries (France, Germany, Italy, United-Kingdom, United States of America and Morocco) suggests the possibility of representing the collected data $x(t)$ as a realisation of a process with two components:

$$x(t) = s(t) + r(t) \tag{1}$$

where $s(t)$ is a slowly changing function known as a "trend component" and $r(t)$ is a "random noise component" which is stationary.

The second step in the data analysis is to calculate the autocorrelation function $ACF(t)$ and the 2-D diagonal slice of the fourth order cumulant$C(t,t,t)$. If the (ACF) and 2-D diagonal slice of the fourth order cumulant ($C(t,t,t)$) represent the low decrease, we can conclude that the time series is non-stationary.

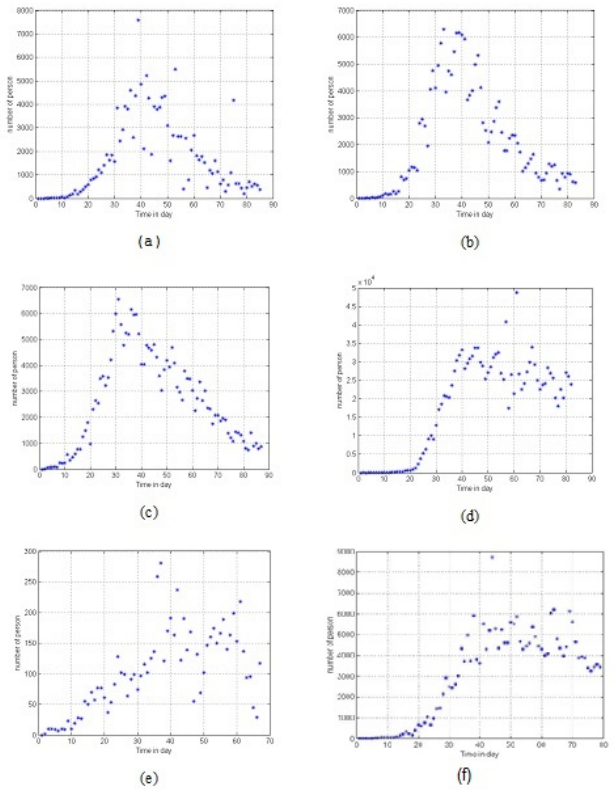


Fig. 1. Collected data, between February and May, representing the COVID-19 pandemic in five countries: (a): France, (b): Germany, (c) Italy, (d) United State of America, (e) Morocco and (f) United Kingdom.

2.2. French data analysis

In the Figure 2, we represent the autocorrelation function $ACF(t)$ (a) and the 2-D diagonal slice of the fourth order cumulants ($C(t,t,t)$ (b) of the time series $x_F(t)$ representing the data collected in France. As we can see on the two curves

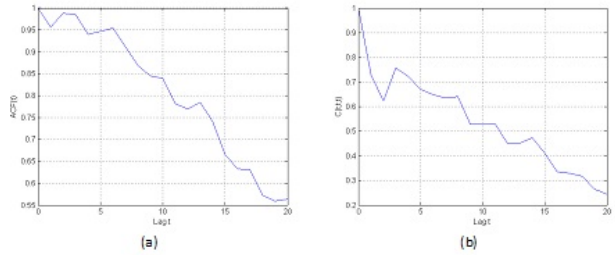


Fig. 2. The autocorrelation function (a) and the fourth cumulant diagonal slice (b) of the time series $x_F(t)$

plotted in Figure 2, the slow decrease confirms that the data represented by the time series $x_F(t)$ is non stationary [14, 17]. Therefore, it is necessary to perform a transformation on $x_F(t)$ to have a stationary process. The elimination of the trend component (to obtain a stationary process) can be performed using the differentiation operator ∇ repeatedly to the times series $x_F(t)$ as follows:

$$\begin{aligned} \nabla^d x_F(t) &= (1-B)^d x_F(t) \\ &= y_F(t) \end{aligned} \quad (2)$$

Where B is the backward shift operator and d is the order of differentiation.
The transformed times series $y_F(t)$ is plotted in Figure 3. In

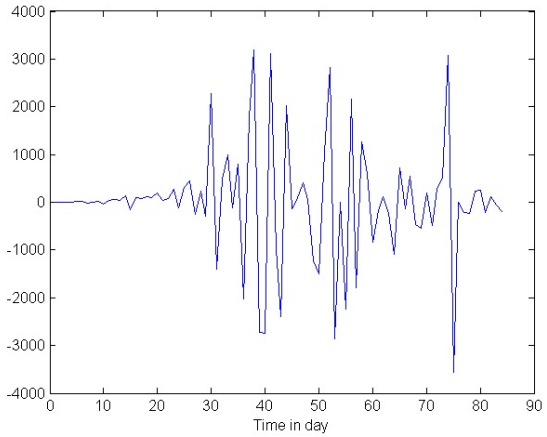


Fig. 3. The transformed time series $y_F(t)$

the Figure 4, we represent the $ACF(t)$ and the cumulant diagonal slice $C(t,t,t)$ of the transformed time series $y_F(t)$, in the case of $d = 1$.

As we can see in the Figure 4, both of the curves of the

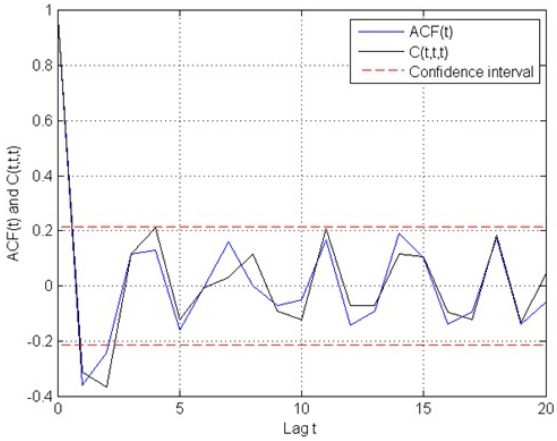


Fig. 4. The autocorrelation function (a) and the fourth cumulant diagonal slice (b) of the transformed time series $x_F(t)$

Autocorrelation function ($ACF(t)$) and the cumulant diagonal slice of fourth order ($C(t,t,t)$) decreases rapidly this confirms that the transformed time series $y_F(t)$ is stationary.

2.2..1 The identification problem

Let us consider the transformed times series $y_F(t)$, described in equation (eq. 2). The problem now is to find satisfactory ARMA(p,q) (because the transformed time series is stationary) model to represent the time series $y_F(t)$. If p and q

were known in advance this would be a straightforward application of the estimation techniques. However this is not usually the case, so it is necessary to identify appropriate values for p and q. It's clear from the curve of Figure 4 that the values of the $ACF(t)$ and $C(t,t,t)$ after lag 2 are into confidence interval at 95% (given by $\pm 1.96/\sqrt{(N)}$, where N is the data length) consequently the times series $y_F(t)$, representing the transformed data, can be described by a MA model of order 1 or 2.
In the order to estimate the parameters of the "a priori" selected models ($MA(1)$ or $MA(2)$) which could represent the time series, we use the algorithm [21], described in the following section.

3. Algorithm based on third order cumulant

The time series $y_F(k)$ is considered like an output of a $FIR(q)$ channel (or $MA(q)$ model), excited by an unobservable input sequences, i.i.d zero-mean symbols with unit energy, across a channel with memory q The output time series is described by the following equation

$$y_F(k) = \sum_{i=0}^{q} h(i)e(k-i) \tag{3}$$

where $h_q = (h(0),h(1),h(2),...,h(q))$ represents the channel impulse response, $e(k)$ with variance σ_e^2.
The completely blind channel identification problem is to estimate h_q based only on the measured data $y_F(k)$ and without any knowledge of the energy of the input data, $e(k)$.
Let us suppose that: the input $e(k)$ is no-Gaussian, with variance σ_e^2, identically independent distributed (i.i.d) with the m^{th} order cumulant vanishes for $m > 2$. In addition we suppose that channel order q is supposed to be known and $h(0) = 1$.
Then the m^{th} order cumulant of the output signal is given by the following equation [18]

$$C_{my_F}(t_1,...,t_{m-1}) = \gamma_{me} \sum_{i=-\infty}^{+\infty} h(i)h(i+t_1)...h(i+t_{m-1}) \tag{4}$$

with γ_{me} represent the m^{th} order cumulant of the excitation signal $\{e(k)\}$ at origin.

If $m = 3$ the equation (4) yield to

$$C_{3y_F}(t_1,t_2) = \gamma_{3e} \sum_{i=0}^{P} h(i)h(i+t_1)h(i+t_2) \tag{5}$$

the same, if $m = 2$ the equation (4) becomes

$$C_{2y_F}(t_1) = \sigma^2 \sum_{i=0}^{q} h(i)h(i+t_1) \tag{6}$$

the Fourier transformation of the equations (5) and (6) gives us the spectre and bispectra respectively

$$S_{3y_F}(\omega_1,\omega_2) = \gamma_{3e}H(\omega_1)H(\omega_2)H(-\omega_1-\omega_2) \tag{7}$$

$$S_{2y_F}(\omega) = \sigma^2 H(\omega)H(-\omega) \tag{8}$$

if we suppose that $\omega = (\omega_1+\omega_2)$, the equation (8) becomes

$$S_{2y_F}(\omega_1+\omega_2) = \sigma^2 H(\omega_1+\omega_2)H(-\omega_1-\omega_2) \tag{9}$$

then, from the equations (7) and (9) we obtain the following equation

$$H(\omega_1+\omega_2)S_{3y_F}(\omega_1+\omega_2) = \varepsilon H(\omega_1)H(\omega_2)S_{2y_F}(\omega_1+\omega_2) \tag{10}$$

with $\varepsilon = (\frac{\gamma_{3x}}{\sigma^2})$. The inverse Fourier transformation of the equation (8) demonstrates that the 3^{rd} order cumulant, the Auto-Correlation Function (ACF) and the impulse response channel parameters are combined by the following equation

$$\sum_{i=0}^{q} h(i)C_{3y_F}(t_1-i,t_2-i) = \varepsilon \sum_{i=0}^{q} h(i)h(i+t_2-t_1)C_{2y_F}(t_1-i) \tag{11}$$

if we use the property of the ACF of the stationary process, such as $C_{2y_F}(t) \neq 0$ only for $(-q \leq t \leq q)$ and vanish elsewhere. In addition if we take $t_1 = -q$, the equation (11) takes the form

$$\sum_{i=0}^{q} h(i)C_{3y_F}(-q-i,t_2-i) = \varepsilon h(0)h(t_2+q)C_{2y_F}(-q) \tag{12}$$

else if we suppose that $t_2 = -q$, the equation (11) becomes

$$C_{3y_F}(-q,-q) = \varepsilon h(0)C_{2y_F}(-q) \tag{13}$$

using the equation (12) and (13) we obtain the following relation

$$\sum_{i=0}^{q} h(i)C_{3y_F}(-q-i,t_2-i) = h(t_2+q) \tag{14}$$

else if we suppose that the system is causal, i.e. that $h(i) = 0$ if $i < 0$. So, for $t_2 = -q,...,0$, the system of equations (12) can be written in matrix form as

$$\begin{pmatrix} C_{3y_F}(-q-1,-q-1) & \dots & C_{3y_F}(-2q,-2q) \\ C_{3y_F}(-q-1,-q)-\alpha & \dots & C_{3y_F}(-2q,-2q+1) \\ \vdots & \ddots & \vdots \\ C_{3y_F}(-q-1,1) & \dots & C_{3y_F}(-2q,-q)-\alpha \end{pmatrix} \times \begin{pmatrix} h(1) \\ h(2) \\ \vdots \\ h(q) \end{pmatrix} = \begin{pmatrix} 0 \\ -C_{3y_F}(-q,-q+1) \\ \vdots \\ -C_{3y_F}(-q,0) \end{pmatrix} \tag{15}$$

where $\alpha = C_{3y_F}(-q,-q)$
the above equation (15) can be written in compact form as

$$Mh_q = d \tag{16}$$

with M the matrix of size $(q+1)\times(q)$ element, h_q a column vector constitute by the unknown impulse response parameters $h_q(n) : n = 1,...,q$ and d is a column vector of size $(q+1)\times(1)$ as indicated in the equation (15). The Least Squares solution (*LS*) of the system of equation (16), permit an identification of the parameters $h_q(n)$ blindly and without any 'information' of the input selective channel. So, the solution will be written under the following form

$$h_q = (M^T M)^{-1} M^T d \tag{17}$$

■ using the presented algorithm, described above (17), we have estimated the parameters of the selected model MA(1) and MA(2) able to represent the transformed time series $y_F(t)$, this result are showed in Table 1.

Table 1
Estimation of the model parameters MA(1) and MA(2)

Parameters of models	$\widehat{h}(1)$	$\widehat{h}(2)$	$\widehat{\sigma^2}$
MA(1)	-0.453		1.012
MA(2)	-0.423	0.216	1.296

■ *Model selection*

After a time series model has been specified and its parameters have been estimated, a check must be made to test whether or not the original specification was correct. This process of selection usually involves two steps:

1) the autocorrelation function and diagonal cumulant slice of fourth order for the simulated series $\widehat{y_F}(t)$ are compared with those of the original, $y_F(t)$, series). If the two-autocorrelation functions (i.e $C(t,t,t)$) seems very different then the validity of the model is doubtful. If the two ACFs (i.e the $C(t,t,t)$) are not markedly different, then we go to the, second, following step,

2) we make a quantitative analysis of the measured and simulated data by the model.

In the following Figures (5) we have plotted the $ACF(t)$ and $C(t,t,t)$ calculated from simulated data using the $MA(1)$ and $MA(2)$. The model can be finally accepted if the errors between the measured and simulated data are not very different. From the curve of Figure 5, we remark that: the simulated and measured data have the same behaviour. So, the model MA(1) is statistically accepted for representing the time series. Finally, in order to generate the original data we proceed as follows:

1) Generate a purely random variable $e(t)$ (white non-Gaussian noise with variance $\sigma^2 = 1.012$), with no correlated samples and with zero mean;

2) The process $e(t)$ was filtered trough the $MA(1)$ to obtain the output $\widehat{y_F(t)}$.

3) The trend component was estimated by the interpolation method in which we have selected a polynomial $p_F(t) = a_4t^4 + a_3t^3 + a_2t^2 + a_1t + a_0$ (where $a_4 = 0.003$ $a_3 = -0.439$,$a_2 = 22.183$, $a_1 = -275.991$, $a_0 = 699.294$ with 4th degree presented in the Figure 6;

4) We generate finally the simulated data $\widehat{x_F}(t) = \widehat{p}(t) + \widehat{y_F(t)}$ (see Figure 7).

In conclusion we can see that the simulated data have the approximately the same behaviour like the collected data representing the time series of COVID-19 in France between

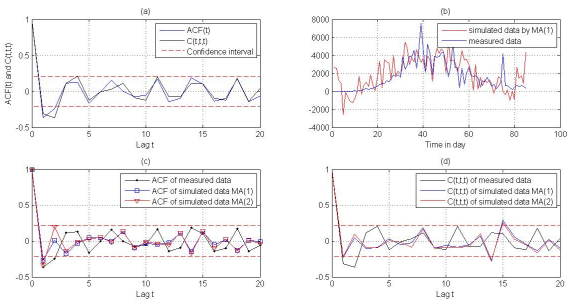


Fig. 5. (a): The $ACF(t)$ and $C(t,t,t)$ of measured data, (b): simulated and measured data, (c) $ACF(t)$ of simulated $\widehat{y_F}(t)$ using the model $MA(1)$ and $MA(2)$ and (d) $C(t,t,t)$ of simulated $\widehat{y_F}(t)$ using the model $MA(1)$ and $MA(2)$

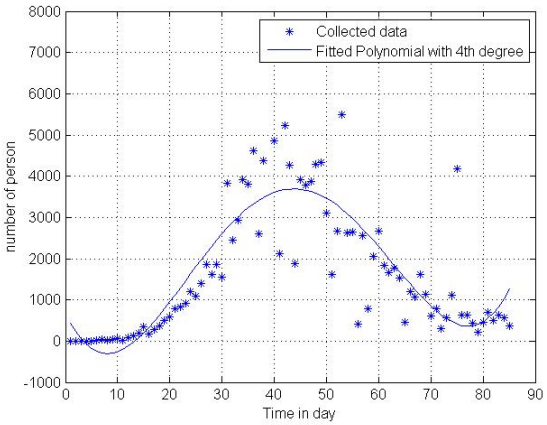


Fig. 6. Fitted polynomial with 4th degree representing the trend component in the time series $x_F(t)$.

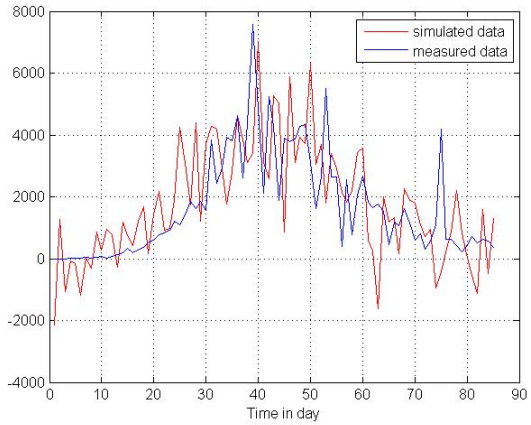


Fig. 7. Simulated and collected data representing the COVID-19 in France.

February and May. The quantitative analysis was made by calculating the relative error between the measured values and the estimated values by the following formula:

$$error_F = \frac{max(y_F(t) - x_F(t))}{max(x_F(t))} \tag{18}$$

The values of this error is 0.37, this error is acceptable because we have a small data number. If the collected data is more importent this error well be more small, because in this case the estimation problem will be taken with minimum bias.

3.1. Germany data analysis

In this part we will consider the data, representing the number of person contaminated by COVID-19 pandemic, collected in Germany. In Figure 1 (b) we have presented the data representing the number of person contaminated by COVID-19 in Germany between February and May 2020.
By proceeding, in the same manner, in the following we describe the model able to generate the time series $x_G(t)$ representing the person contaminated by COVID-19 in Germany.
■ *model selection*
Like the French data, we apply the differentiation operator to the time series $x_G(t)$ as follow

$$\begin{aligned} \nabla^d x_G(t) &= (1-B)^d x_G(t) \\ &= y_G(t) \end{aligned} \tag{19}$$

By taking $d = 1$ we obtain a stationary process $y_G(t)$. After plotting the $ACF(t)$ and $C(t,t,t)$ we observe that the process $y_G(t)$ can be modelled by MA model of order 1 or 2. So, based on the algorithm described above (equation 17) we represent the estimated the parameters in Table 2. In Figure 8 we have summarized the obtained results using the two selected 'a priori' models, in which we have plotted the different $ACF(t)$ and $C(t,t,t)$ of simulated data using the models MA(1) and MA(2) and transformed data $y_G(t)$. By analysing the results represented in Figure 8, we remark that the model MA(1) is more suitable to represent the data $y_G(t)$, so by adding the trend component using a polynomial with fourth degree $P_G(t) = a_4t^4 + a_3t^3 + a_2t^2 + a_1t + a_0$ (where $a_4 = 0.004$ $a_3 = -0.557$,$a_2 = 25.609$, $a_1 = -248.341$, $a_0 = 380.423$. As see in the Figure 8 (b) that the simulated data have approximately the same behaviour like the collected data representing the time series representing the COVID-19 pandemic in Germany between February and May. In order to validate the obtained results, we have calculated the relative error between the measured values and the estimated values by the formula describing in equation 18 which is 0.31, this value is inferior to those obtained in the case of French data.

Table 2
Estimation of the model parameters MA(1) and MA(2) selected to represent the data $y_G(t)$

Parameters of models	$\widehat{h}(1)$	$\widehat{h}(2)$	$\widehat{\sigma^2}$
MA(1)	-0.313		1.050
MA(2)	1.4852	-0.8872	3.573

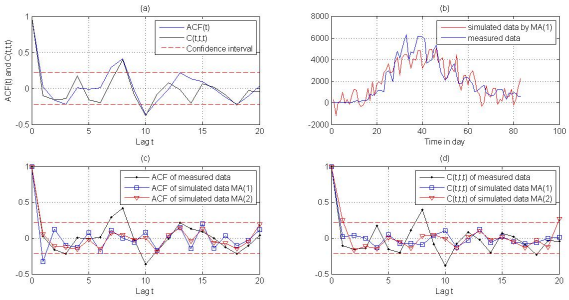


Fig. 8. (a):The $ACF(t)$ and $C(t,t,t)$ of measured data, (b): simulated and measured data, (c) $ACF(t)$ of simulated $\widehat{y_G}(t)$ using the model $MA(1)$ and $MA(2)$ and (d) $C(t,t,t)$ of simulated $\widehat{y_G}(t)$ using the model $MA(1)$ and $MA(2)$

3.2. United Kingdom data analysis

In this section we will consider the data collected in United Kingdom representing the person infected by COVID-19 between February and May, 2020 (see Figure 1). The first step is to calculate the ACF(t) and C(t,t,t), of the time series $x_{UK}(t)$ representing the data collected in United Kingdom, which they are plotted in figure 9.
The slowly decay of the $ACF(t)$ and $C(t,t,t)$ function (Figure 9), demonstrate that the time series $x_{UK}(t)$ representing the collected data of COVID-19 pandemic is not stationary.

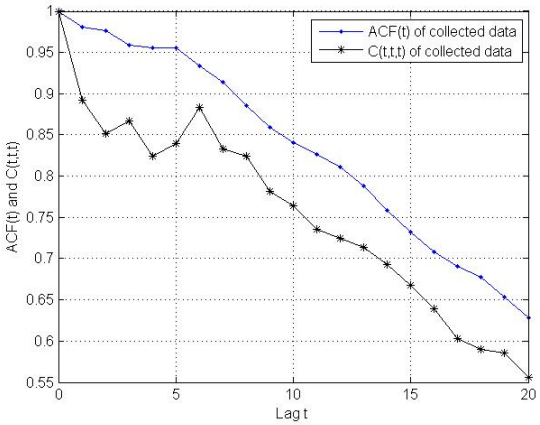


Fig. 9. The $ACF(t)$ and $C(t,t,t)$ of measured data

■ *Data analysis and model selection*
Like the French and Germany data, we apply the differentiation operator to the time series $x_{UK}(t)$ as follow

$$\begin{aligned} \nabla^d x_{UK}(t) &= (1-B)^d x_{UK}(t) \\ &= y_{UK}(t) \end{aligned} \tag{20}$$

By taking $d = 1$ we obtain a stationary process $y_{UK}(t)$. After plotting the $ACF(t)$ and $C(t,t,t)$ we observe that process $y_{UK}(t)$ can be modelled by MA model of order 1 or 2. So,

based on the algorithm described above (equation 17) we represent the estimated parameters in Table 3.

In figure 10 we have summarized the obtained results using

Table 3
Estimation of the model parameters MA(1) and MA(2) selected to represent the data $y_{UK}(t)$

Parameters of models	$\widehat{h}(1)$	$\widehat{h}(2)$	$\widehat{\sigma^2}$
$MA(1)$	-6.855		15.475
$MA(2)$	0.259	0.151	0.743

the two selected 'a priori' models, in which we have plotted the different $ACF(t)$ and $C(t,t,t)$ of simulated data using models $MA(1)$ and $MA(2)$ and transformed data $y_{UK}(t)$. By analysing the results represented in Figure 10, we remark that the model $MA(2)$ is more suitable to represent the data $y_{UK}(t)$, so by adding the trend component using a polynomial with fourth degree $P_{UK}(t) = a_4t^4 + a_3t^3 + a_2t^2 + a_1t + a_0$ (where $a_4 = 0.00$ $a_3 = -0.365$,$a_2 = 21.299$, $a_1 = -300.312$, $a_0 = 380.423$. As see in the figure 10 (b) that the simulated data have approximately the same behaviour like the collected data representing the time series of COVID-19 in United Kingdom between February and May. We have calculated the relative error, in order to make a quantitative analysis, between the measured values and the estimated values by the formula describing in equation 18 which is, in the case of United Kingdom data is equal to 0.32, this value is inferior to those obtained in the case of French data and approximately equal to the obtained value in the Germany data.

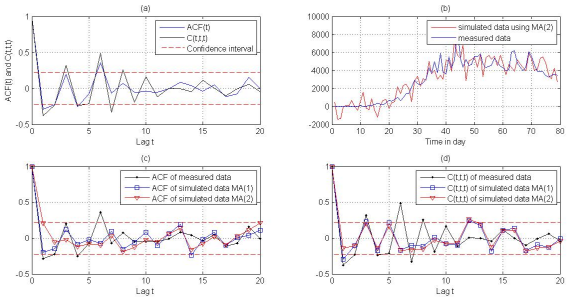


Fig. 10. (a):The $ACF(t)$ and $C(t,t,t)$ of measured data, (b): simulated and measured data, (c) $ACF(t)$ of simulated $\widehat{y_{UK}}(t)$ using the model $MA(1)$ and $MA(2)$ and (d) $C(t,t,t)$ of simulated $\widehat{y_{UK}}(t)$ using the model $MA(1)$ and $MA(2)$

3.3. Italy data analysis

In this section we will consider the fourth country in Europe which it has the first more infected country by COVID-19, its Italy. In figure 11 we represent the collected data of infected people in Italy between February 21 and May 17, 2020. Like the previous data, collected in France, Germany and United Kingdom, we represent the $ACF(t)$ and the $C(t,t,t)$ in order to see if the data is stationary or not. Its clear from the curve (Figure 12) of the $ACF(t)$ and $C(t,t,t)$ of the original data $x_I(t)$ is not stationary.

■ *Model selection* Like the French, Germany and United Kingdom data, we apply the differentiation operator to the time series $x_I(t)$ as follow

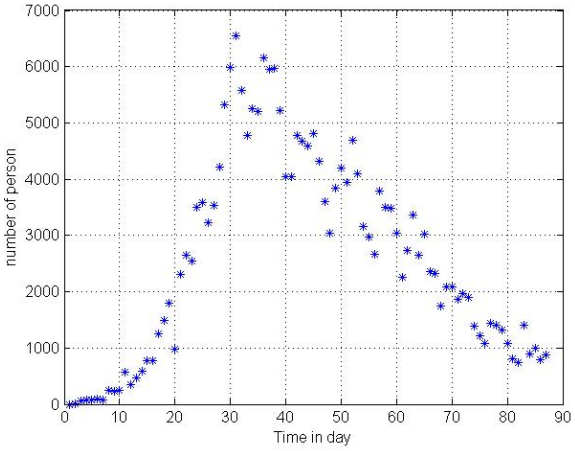


Fig. 11. The collected data in Italy

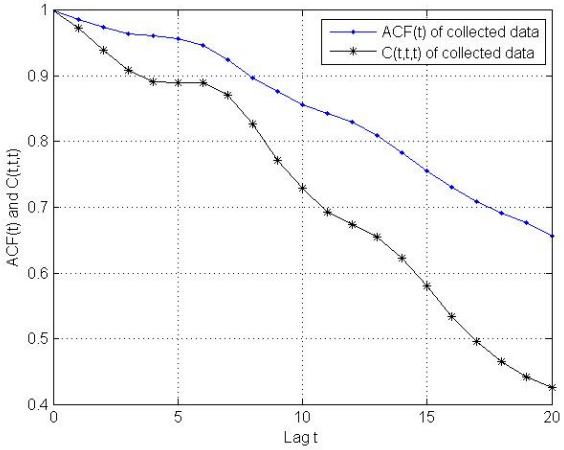


Fig. 12. $ACF(t)$ and $C(t,t,t)$ of measured data

$$\begin{aligned} \nabla^d x_I(t) &= (1-B)^d x_I(t) \\ &= y_I(t) \end{aligned} \tag{21}$$

By taking $d = 1$ we obtain a stationary process $y_I(t)$. After plotting the $ACF(t)$ and $C(t,t,t)$, we observe that the process $y_I(t)$ can be modelled by MA model of order 1 or 2. In the figure 13 we have summarized all $ACF(t)$ and $C(t,t,t)$ of collected and transformed data $y_I(t)$.

Based on the obvious figure 13 (a), we observe that the process $y_I(t)$ can be modelled by MA model of order 1 or 2. So, based on the algorithm described above (equation 17) we represent the estimated parameters in Table 4.

In the Figure 13 (b), we remark that the model $MA(2)$ is more suitable to represent the time series $y_I(t)$, and by adding the trend component using a polynomial with fourth degree $P_I(t) = a_4t^4 + a_3t^3 + a_2t^2 + a_1t + a_0$ (where

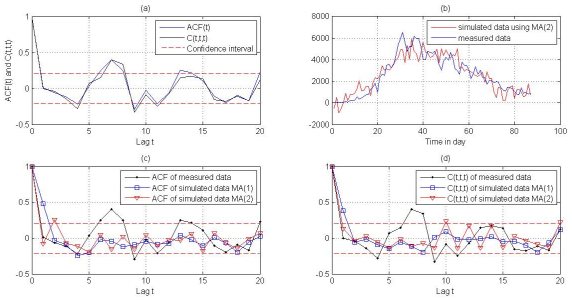


Fig. 13. (a):The $ACF(t)$ and $C(t,t,t)$ of measured data, (b): simulated and measured data, (c) $ACF(t)$ of simulated $\widehat{y_I}(t)$ using the model $MA(1)$ and $MA(2)$ and (d) $C(t,t,t)$ of simulated $\widehat{y_I}(t)$ using the model $MA(1)$ and $MA(2)$

Table 4
Estimation of the model parameters MA(1) and MA(2) selected to represent the data $y_I(t)$

Parameters of models	$\widehat{h}(1)$	$\widehat{h}(2)$	$\widehat{\sigma^2}$
$MA(1)$	0.969		1.221
$MA(2)$	-0.031	0.345	0.726

$a_4 = 0.002$ $a_3 = -0.296$,$a_2 = 12.116$, $a_1 = 5.981$, $a_0 = -465.551$). In the Figure 13 (b) we can see that the simulated data have approximately the same behaviour like the collected data representing the time series of COVID-19 in Italy between February and May. We have calculated the relative error between the measured and estimated values by the formula describing in equation 18 which is, in the case of Italy data is equal to 0.21, this value is inferior to those obtained of all obvious presented data (from: France, Germany and United Kingdom).

3.4. USA data analysis

Now we will consider the data collected in more infected country in the world, the data representing the infected person by COVID-19 in United States of America (USA) between February and May 2020. The time series $x_{Usa}(t)$ representing the data collected in USA was presented in Figure 14.

As the previous data, collected in France, Germany and United Kingdom, we represent, in the first time, the $ACF(t)$ and the $C(t,t,t)$ of the collected data in USA in order to see if the data is stationary or not.

Its clear from the curve (Figure 15) of the $ACF(t)$ and $C(t,t,t)$ that the original data $x_{Usa}(t)$ is not stationary.

■ *Model selection*

In order to eliminate the trend component, we apply the differentiation operator to the time series $x_{Usa}(t)$ as follow

$$\begin{aligned} \nabla^d x_{Usa}(t) &= (1-B)^d x_{Usa}(t) \\ &= y_{Usa}(t) \end{aligned} \tag{22}$$

By taking $d = 1$ we obtain a stationary process $y_{Usa}(t)$. After plotting the $ACF(t)$ and $C(t,t,t)$ we observe that process $y_{Usa}(t)$ can be modelled by MA model of order 1 or 2. So,

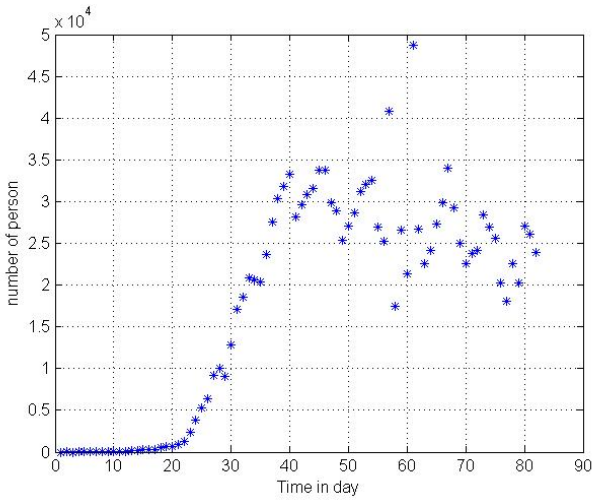


Fig. 14. Daily number of affect person in USA

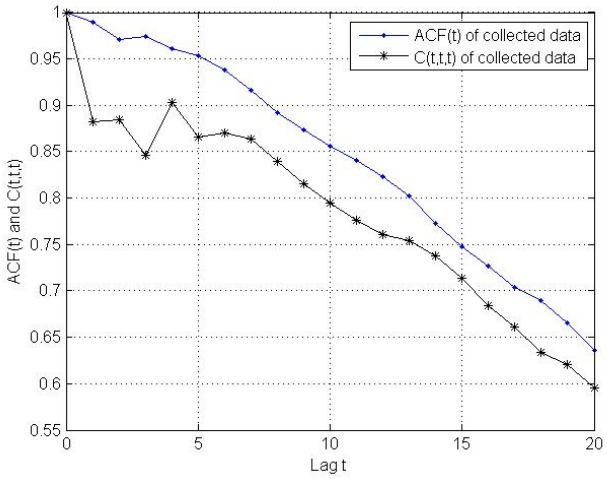


Fig. 15. $ACF(t)$ and $C(t,t,t)$ of the time series $x_{USA}(t)$

based on the algorithm described above (equation 17) we represent the estimated parameters in Table 5.

In Figure 16 we have summarized the obtained results using

Table 5
Estimation of the model parameters MA(1) and MA(2) selected to represent the data $y_{Usa}(t)$

Parameters of models	$\widehat{h}(1)$	$\widehat{h}(2)$	$\widehat{\sigma^2}$
MA(1)	-3.510		6.311
MA(2)	0.119	0.569	1.050

the two selected 'a priori' models, in which we have plotted the different $ACF(t)$ and $C(t,t,t)$ of simulated data using models MA(1) and MA(2) and transformed data $y_{Usa}(t)$. By analysing the results represented in Figure 16, we remark that the model MA(1) is more suitable to represent the data $y_{Usa}(t)$, so by adding the trend component using a polynomial with fourth degree $P_Usa(t) = a_4t^4 + a_3t^3 + a_2t^2 + a_1t + a_0$ (where $a_4 = 0.014$ $a_3 = -2.697$,$a_2 = 25.609$, $a_1 = -248.341$, $a_0 = 380.423$). As we can see in the Figure

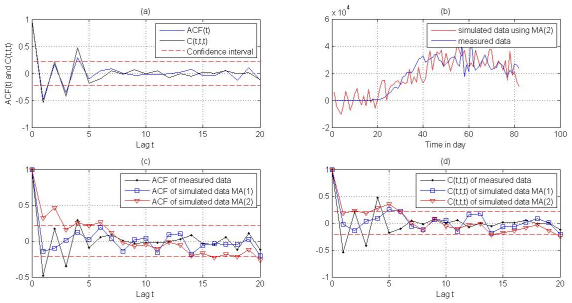


Fig. 16. (a):The $ACF(t)$ and $C(t,t,t)$ of measured data, (b): simulated and measured data, (c) $ACF(t)$ of simulated $\widehat{y_{Usa}}(t)$ using the model $MA(1)$ and $MA(2)$ and (d) $C(t,t,t)$ of simulated $\widehat{y_{Usa}}(t)$ using the model $MA(1)$ and $MA(2)$

16 (b) that the simulated data have approximately the same behaviour like the collected data representing the time series of COVID-19 in USA between February and May.
Finally we make the quantitative analysis we have calculated the relative error between the measured values and the estimated values by the formula describing in equation 18 which is 0.2836, this value is comparable to those obtained in the previous data, so the model is acceptable to represent the collected data in USA.

3.5. Moroccan data analysis

In this section we will consider the last data presented in this paper. This data was collected, between February and May 2020, in Morocco, one of the countries representing a minimum of person infected by COVID-19. The time series $x_M(t)$ representing the data collected in Morocco was presented in Figure 17.

The same processing will be applied to the time series

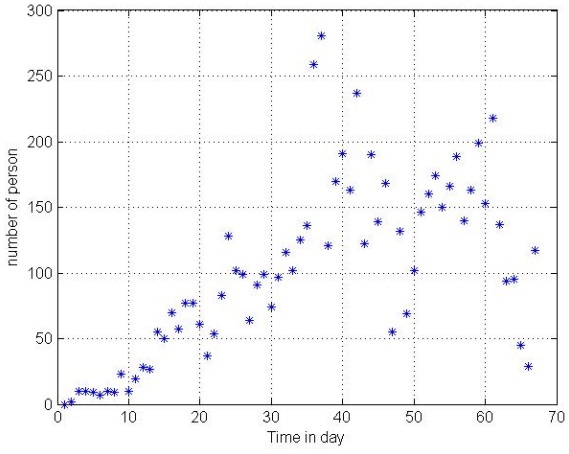


Fig. 17. Daily number of infected person in Morocco

$x_M(t)$, in the first time we calculate the $ACF(t)$ and $C(t,t,t)$, these functions have been presented in Figure 18.

The curves presented in Figure 18 indicate that the slowly decreasing of the two functions ($ACF(t)$ and $C(t,t,t)$) demonstrate that the time series $x_M(t)$ is not stationary due

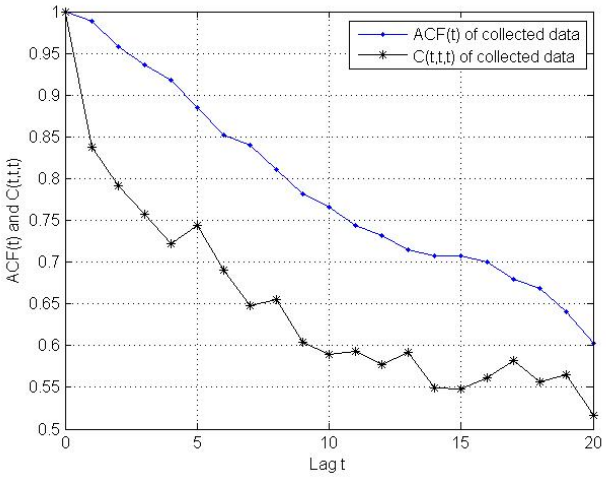


Fig. 18. $ACF(t)$ and $C(t,t,t)$ of the collected data $x_M(t)$

to the presence of trend component. So, in order to eliminate this component we apply the differentiation operator to the time series $x_M(t)$ as

$$\begin{aligned} \nabla^d x_M(t) &= (1-B)^d x_M(t) \\ &= y_M(t) \end{aligned} \tag{23}$$

By taking $d=1$ we obtain a stationary process $y_M(t)$. After plotting the $ACF(t)$ and $C(t,t,t)$, Figure **??** (a), we observe that the process $y_M(t)$ can be modelled by MA model of order 1 or 2. So, based on the algorithm described above (equation 17) we represent the estimated the parameters in Table 6.

By analysing the results represented in Figure 19, we remark that the model MA(1) is more suitable to represent

Table 6
Estimation of the model parameters MA(1) and MA(2) selected to represent the data $y_M(t)$

Parameters of models	$\widehat{h}(1)$	$\widehat{h}(2)$	$\widehat{\sigma^2}$
MA(1)	0.031		0.5953
MA(2)	0.027	−0.772	0.782

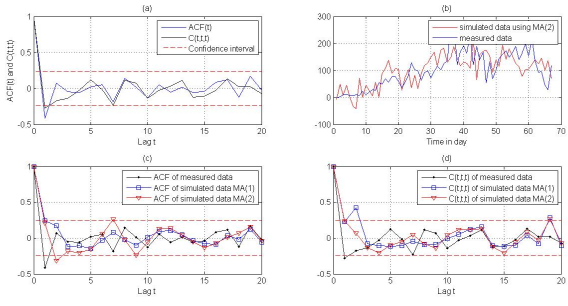


Fig. 19. (a):The $ACF(t)$ and $C(t,t,t)$ of measured data, (b): simulated and measured data, (c) $ACF(t)$ of simulated $\widehat{y_M}(t)$ using the model $MA(1)$ and $MA(2)$ and (d) $C(t,t,t)$ of simulated $\widehat{y_M}(t)$ using the model $MA(1)$ and $MA(2)$

the data $y_M(t)$, so by adding the trend component using a polynomial with third degree $P_M(t) = a_3t^3 + a_2t^2 + a_1t + a_0$ (where $a_3 = -0.003$,$a_2 = 0.170$, $a_1 = 1.147$, $a_0 = -2.169$. As see in the Figure 19 (b) that the simulated data have approximately the same behaviour like the collected data representing the time series of COVID-19 in Morocco between February and May. We have calculated the relative error between the measured values and the estimated values by the formula describing in equation 18 which is 0.412, this value is not good comparatively to those obtained in the previous data but statistically is accepted.

4. Discussion: COVID-19 pandemic modelling

In this section we will discuss the problem of modeling COVID-19 data by one model independent of the country in which the data was collected. By analysing the result obtained in the previous that all studied time series are composed of two components : the trend component and the random component. The trend component was modelled by fourth order polynomial except the moroccan data which was represented by third order polynomial, and the random component was modeled by a moving average MA with the first order or second order.
So, the question here is : it's possible to find a suitable model 'generalized model' able to represent the collected data, representing the infected person by COVID-19, regardless of the country in which the data was collected?
In order to give an approximate response of this question we proceed as follow

4.1. Random component modelling

In the same manner, and because all random component can be represented by a moving average with order 1 ($MA(1)$), we will consider only that all previous data, the random component, are modelled by $MA(1)$. The coefficient of the 'mean' $MA(1)$ model, will be considered as the average of the parameter values of the $MA(1)$ models representing the previous data in each country.
The obtained result are summarized in the following table 7.
In the Figure 20 we have presented the $ACF(t)$ and $C(t,t,t)$

Table 7
The 'mean' model parameters MA(1) to represent the random part data

Parameters of models	$\widehat{h}(1)$	$\widehat{\sigma^2}$
MA(1)	-1.688	0.5953

of the random component for different data (from different countries) and the mean model.

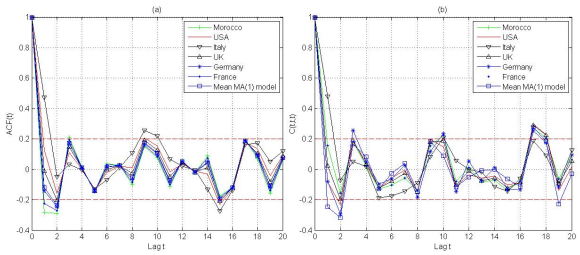


Fig. 20. (a): The $ACF(t)$ of measured and simulated data using mean model, (b): The $C(t,t,t)$ of measured data and simulated data using mean model

4.2. Trend component

Because the propagation of the COVID-19 depend on the population number in each country. In general we have a concave polynomial but with different values. So, the component value depend on the collected data in each country. This component is defined for each data. In the Figure 21 we represent the simulated data, respectively in France, Germany, Italy, United Kingdom, USA and Morocco , using the mean MA(1) model.

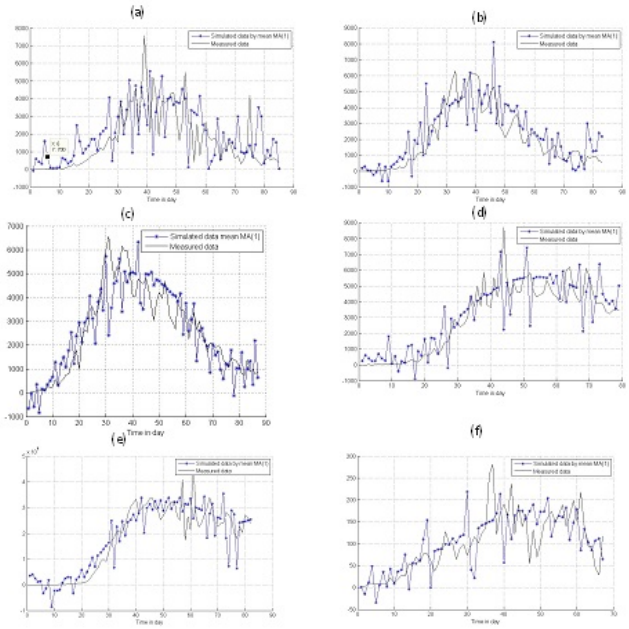


Fig. 21. Simulated and measured data representing the COVID-19 contamination in six countries :(a): France , (b): Germany, (c): Italy, (d): United Kingdom, (e): United States of America and (f): Morocco

5. conclusion

In this paper we have studied the identification problem of the evolution of COVID-19 contamination in six country in three continent : Europe, North America and Africa. The identification problem of the collected data, in each country, was represented by two component: the trend component

representing the slowly variation of the data and the random component representing random variation 'stochastic value'. By analysing the time series representing the collected data of COVID-19 pandemic, we have eliminate the trend component by using the differentiation operator to obtain a stationary process, and the random component 'stationary process' is modeled by a moving average with order 1 or 2. The identified model can represent the able to represent the collected data in each countries.

In the last we have proposed that the random component can be modeled by a $MA(1)$ with known parameters.

References

[1] Hsu A, Lane N. Impact of COVID-19 on residents of Canada's long-term care homesongoing challenges and policy response. London, United Kingdom: International Long-Term Care Policy Network; 2020.

[2] He X, Lau EHY, Wu P, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. National Library of Medecine, May;26(5), p. 672-675, 2020.

[3] Ziwei Dai and Jason W Locasale. Cooperative Virus Propagation Underlies COVID-19 Transmission Dynamics. May 7;2020.05.05.

[4] Kimberly A. Prather, Chia C. Wang, Robert T. Reducing transmission of SARS-CoV-2. Science, 27 May 2020.

[5] Ying Liu, Albert A Gayle, Annelies Wilder-Smith, Joacim Rocklv. The reproductive number of COVID-19 is higher compared to SARS coronavirus, ournal of Travel Medicine, Volume 27, Issue 2, March 2020.

[6] Moritz U. G. Kraemer, Chia-Hung Yang4, Bernardo Gutierrez,Chieh-Hsi Wu6, Brenna. The effect of human mobility and control measures on the COVID-19 epidemic in China. Science 01 May 2020, Vol. 368, Issue 6490, pp. 493-497.

[7] Kenji Mizumoto, Katsushi Kagaya2, Alexander Zarebski5 , Gerardo Chowe. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. Eurosurveillance - Volume 25, Issue 10,p.1-6, 12 March 2020.

[8] Giulia G., Franco B., Raffaele B., Patrizio C., Alessandro D. F., Angela D. M. and Marta C. Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. Nature Medicine (2020).

[9] Sheng Z., Meng, Y. D., Wenbo Y. Lei P., Zhaofen L. Dechang C. Estimation of the reproductive number of novel coronavirus (COVID-19) and the probable outbreak size on the Diamond Princess cruise ship: A data-driven analysis. International Journal of Infectious Diseases Volume 93, p. 201-204, April 2020.

[10] Zifeng Y., Zhiqi Z., Ke W., Sook-S. W., Wenhua L., Mark Z., Peng L., Xudong C., Zhongqiang G., Zhitong M., Jingyi L., Xiaoqing L., Shiyue L., Yimin L., Feng Y., Weijie G., Yifan Y., Fei L., Shengmei L., Yuqi X., Bin L., Zhoulang W., Shaobo Z., Yaonan W., Nanshan Z., Jianxing H.. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. J Thorac Dis. vol. 12, issue 3, p. 165174, March, 2020.

[11] Simiao C., Juntao Y., Weizhong Y., Chen W. and Till B. COVID-19 control in China during mass population movements at New Year. The Lancet, Volume 395, Issue 10226, pp. 764-766, 2020.

[12] Shu Y., Peihua C., Peipei D., Ziting W., Zian Z., Lin Y., Xuan Y., Qi Z., Xixi F., Xiaohui W., Weiguo L., Enmei L., Ju C., Yaolong C., Daihai H. Early estimation of the case fatality rate of COVID-19 in mainland China: a data-driven analysis. Ann Transl Med. vol. 8, issue 4, pp. 1-6, February, 2020.

[13] Davit G., Ekaterine M and Mikheil B. Research on COVID-19 Virus Spreading Statistics based on the Examples of the Cases from Different Countries. Electronic Journal of General Medicine, vol. 17, issue. 4. pp. 1-4, 2020.

[14] Brillinger D. R. and Rosenblatt M. Computation and interpretation of kth-order spectra, in: B. Harris, ed., Spectral Analysis of time series, Wiley, New York, pp. 189-232, 1987.

[15] Lacoume J. L., Amblard P. O., Comon P., Statistiques d'ordre Suprieur pour la traitement du signal, Masson, Paris, 1997.

[16] Box G. E. P. and Jenkins G., Time series Analysis, Forecasting and Control. Holden-Day, San Fransisco (1970).

[17] Brokwell P.J., Davis R.A., Time Series Theory and Methods, Second Edition Springer-Verlag. (1991).

[18] Safi S. and Zeroual A. Blind parametric identification of non-Gaussian FIR systems using higher order cumulants. International Journal of Systems Science, Taylor & Francis Group, vol. 35, issue 15, pp. 855-867, 2004.

[19] Safi S. and Zeroual A. Blind identification in noisy environment of nonminimum phase finite impulse response (FIR) system using higher order statistics. Systems Analysis Modelling Simulation, Taylor & Francis Group, vol. 35, issue 15, pp. 671-681, 2003.

[20] Safi S., Frikel M., Zeroual A. and M'Saad M. Higher order cumulants for identification and equalization of multicarrier spreading spectrum systems. Journal of telecommunications and information technology, vol.1, pp. 74-84, 2011.

[21] Safi S., Frikel M., M'Saad M and Zeroual A. Blind impulse response identification of frequency radio channels: application to BRAN a channel. Int. J. Sig. Proces, vol. 4, issue 1, pp. 201-206, 2007.

[22] Zidane M., Safi S., Sabri M. Measured and estimated data of non-linear BRAN channels using HOS in 4G wireless communications. Data in brief, Elsevier, vol. 17, pp. 1136-1148, 2018.

[23] Zidane M., Safi S., Sabri M. and Frikel M. Using Least Mean p-Power Algorithm to Correct Channel Distortion in MC-CDMA Systems. Journal of telecommunications and information technology, vol.3, pp. 23-30, 2018.

[24] Zidane M., Safi S., Sabri M. Extending HOC-based methods for identifying the diagonal parameters of quadratic systems. Signal, Image and Video Processing, Springer London, vol. 12, issue 1, pp. 125-132, 2018.

[25] Elkassimi S., Safi S., Manaut B. Blind Channel Equalization by Adaptive Filter Algorithms. Universal Journal of Applied Mathematics. vol. 5, issue 3, pp. 45-52, 2017.

[26] Ling X. and Feng D. Parameter estimation for control systems based on impulse responses. International Journal of Control, Automation and Systems. vol. 15, pp. 24712479, 2017.

[27] Meihang L. and Ximei L. The least squares based iterative algorithms for parameter estimation of a bilinear system with autoregressive noise using the data filtering technique. Signal Processing, Vol. 147, pp. 23-34, 2018.

[28] Feng D., Ling X. Alsaadi F. Tasawar H. Iterative parameter identification for pseudo-linear systems with ARMA noise using the filtering technique. IET Control Theory & Applications,vol. 12, issue 7, pp. 892 899, 2018.

Detection of Wideband Sources Propagating on a Small Antenna Array

Hassan Ougraz
LIMATI Laboratory
Polydisciplinary Faculty
Sultan Moulay Slimane University
Beni Mellal, Morocco
Email: hassan.ougrazfpb@usms.ac.ma

Said Safi
LIMATI Laboratory
Polydisciplinary Faculty
Sultan Moulay Slimane University
Beni Mellal, Morocco
Email: said.safi@usms.ac.ma

Miloud Frikel
LIS Laboratory
UNICAEN, ENSICAEN
Normandie University
Caen, France
Email: miloud.frikel@ensicaen.fr

***Abstract*—Wideband transmission is now preferred over narrowband communication due to recent technological advancements. With fewer parameters, wideband transmission can estimate more accurately the Direction-of-Arrival (DOA). This paper presents an analytical study for the localization of wideband sources to test the performance of DOA estimation on a small antenna array, under various levels of noise and different angular distances. Two incoherent wideband methods are presented to prove the accuracy of DOA estimation, the first is the Test of Orthogonality of Projected Subspaces (TOPS) method, and the second is the Test of Orthogonality of Frequency Subspaces (TOFS) method, using a Uniform Circular Array (UCA) geometry. Finally, the results prove that TOFS performs well.**

I. Introduction

A variety of fields employ antenna arrays, including radar [1]–[3], sonar, medical imaging, wireless communication systems [4], [5], and others. Antenna arrays are recommended since they can detect the DOA of multiple signals which is not possible with a single antenna.

A different types of strategies have been proposed to address the challenge of estimating the Direction-of-Arrival (DOA) of numerous wideband signals. The incoherent signal subspace algorithms estimate the wideband DOAs by employing multiple narrowband signals that have been independantly decomposed from a wideband signal [6]. In particular, these methods use narrowband DOA estimation approaches. Although the incoherent approaches improve the estimation efficiency in high Signal-to-Noise Ratio (SNR) areas, it suffers when the SNR of certain frequency ranges is low. In other words, the effectiveness of the final estimation will be effected by poor estimations from certain frequency bins. However, the coherent signal subspace algorithms have been proposed to solve these limitations and enhance DOA estimation efficiency [7]. The idea of this algorithms is how to focus covariance matrices, and many techniques for obtaining a proper focusing matrix have been proposed [8], [9]. However, every focusing approach depends on initial values, which are the predetermined directions of arriving signals, and the effectiveness of the coherent signal subspace approaches requires these initial values [10].

In this research, we interest in two incoherent approaches to localize the directions on a small antenna array, under various scenarios. The first is the Test of Orthogonality of Frequency Subspaces (TOFS) technique [11] that produces the searching steering vectors of every conceivable DOA and every frequency. This method estimates correctly when the SNR is large, since it is an incoherent technique [11], and suffers in noisy environments.

The second known as Test of Orthogonality of Projected Subspaces (TOPS) technique [12] employs the signal and noise subspaces of different frequency ranges to provide high DOA estimation efficiency without the need for initial values. This approach does not require focusing angles or a beamforming matrix, does not suffer from bias at high SNR, and integrates frequency bins better than other incoherent approaches at low SNR. However, the method has a drawback that the spectra calculated by TOPS contains some spurious peaks, making it difficult to estimate the real DOA of targets.

In this research, we concentrate on the comparison of those algorithms described above, using a small antenna array, in a different values of SNR and for different angular distance, in order to offer better comparison in term of advantage and disadvantage of each technique.

The reminder of the paper is structured as follows. Section II presents the model of wideband signal and the fundamentals of direction estimation, and the wideband methods, followed by Section III that describes the simulation outcomes. Finally, Section IV involves the conclusion and future work.

II. Problem Formulation and Wideband Methods

A. Wideband Signal Model

Two dimensional array consists of sensors that are placed in a circle where the distance between two consecutive sensors may vary, the simplest example is the uniform circular array of M elements, where the distance is equal to half of the wavelength. We consider that the L wideband sources $(L < M)$ are identifiable or can be computed [13], [14], with identical bandwidth that impinging on the number of array sensors from directions $(\theta_1, \theta_2, \ldots, \theta_L)$. The Fig. 1 shows the UCA configuration for broadband DOA estimation.

It is supposed that all signals are uncorrelated and exist within the bandwidth between the lowest frequency of signal sources (f_L) and the highest frequency of signal sources (f_H).

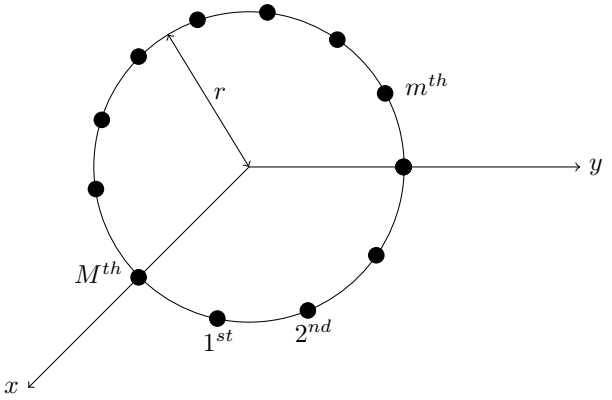


Fig. 1: UCA structure of M elements

The received signal in wideband scenario at the m^{th} sensor is:

$$x_m(f) = \sum_{l=1}^{L} s_l(f) \exp\big(jf\frac{2\pi r}{\lambda}\cos(\theta_l - \varphi_m)\big) + n(f), \quad (1)$$

where $s_l(f)$ is the l^{th} wideband signal, $n(f)$ is an additive white gaussian noise at the m^{th} element, and φ_m is the angular location of the m^{th} sensor:

$$\varphi_m = 2\pi\frac{(m-1)}{M} \quad (2)$$

The received broadband signals are then split into K narrowband signals. Then, the output signals, in Fourier domain, can be written in vector form as follows:

$$x(f_i) = A(f_i,\theta)s(f_i) + n(f_i), \quad 1 \le i \le K, \quad (3)$$

where $f_i \in]f_L, f_H[$,

$$A(f_i,\theta) = \big[a(f_i,\theta_1),\ a(f_i,\theta_2),\ \ldots,\ a(f_i,\theta_L)\big], \quad (4)$$

$$a(f_i,\theta_l) = \begin{bmatrix} \exp\left(jf_i\frac{2\pi r}{\lambda}\cos(\theta_l-\varphi_1)\right) \\ \vdots \\ \exp\left(jf_i\frac{2\pi r}{\lambda}\cos(\theta_l-\varphi_{M-1})\right) \end{bmatrix} \quad (5)$$

The covariance matrix, in wideband case, can be calculated by the formula:

$$\begin{aligned} R_{xx}(f_i) &= E[x(f_i)x^H(f_i)] \\ &= A(f_i,\theta)R_{ss}(f_i)A^H(f_i,\theta) + \sigma_n^2 I, \end{aligned} \quad (6)$$

where $R_{ss}(f_i) = E[s(f_i)s^H(f_i)]$, and I is an $M \times M$ unit matrix. Assuming that the L signal sources are uncorrelated, $R_{ss}(f_i)$ has full order, then the signal subspaces matrix $F_s(f_i)$ and noise subspaces array $F_n(f_i)$ at frequency f_i may be generated using the covariance matrix's eigen-values decomposition as:

$$F_s(f_i) = [\lambda_1(f_i),\ \lambda_2(f_i),\ \ldots,\ \lambda_L(f_i)], \quad (7)$$

$$F_n(f_i) = [\lambda_{L+1}(f_i),\ \lambda_{L+2}(f_i),\ \ldots,\ \lambda_M(f_i)], \quad (8)$$

where $\lambda_1(f_i),\ \ldots,\ \lambda_M(f_i)$ are the perpendicular eigenvectors of $R_{xx}(f_i)$, sorted in degrading rank by their eigenvalues.

B. TOFS method

In this study, the TOFS method proposed in [11] is used to detect the direction of the incident signals, as well as to demonstrate the performance of the ULA and UCA geometries. The test of orthogonality of frequency subspaces (TOFS) technique employs the noise subspace derived from the eigen-values decomposition of the correlation matrix for each frequency as noticed in [11]. The orthogonality between the steering vector and the noise subspaces is used to estimate the DOA of each incoming wideband signal source. The vector $Q_{TOFS}(\theta)$ is achieved as follows, [15]:

$$\begin{aligned} Q_{TOFS}(\theta) &= \big[a^H(f_1,\theta)F_n(f_1)F_n^H(f_1)a(f_1,\theta) \\ &\quad a^H(f_2,\theta)F_n(f_2)F_n^H(f_2)a(f_2,\theta) \\ &\quad \ldots a^H(f_K,\theta)F_n(f_K)F_n^H(f_K)a(f_K,\theta)\big] \end{aligned} \quad (9)$$

The estimated DOA of sources can be obtained by judging the extent that each element near to zero, DOA estimation could be achieved using the following equation:

$$\hat{\theta} = arg\ \max_{\theta}\ \frac{1}{\sigma_{min}(\theta)} \quad (10)$$

where $\sigma_{min}(\theta)$ is the shortest singular value of $Q_{TOFS}(\theta)$.

C. TOPS Method

TOPS computes the DOA of arriving wideband signals by utilizing both the signal and noise subspaces of every frequency range [12]. Then, we extract the signal subspace $F_s(f_i)$ and the noise subspace $F_n(f_i)$ from the eigen-decomposition of the covariance matrix of every frequency range. TOPS beats coherent signal subspaces in that it eliminates the need for initial DOA predictions for the frequency transform process.

The m^{th} term on the diagonal of the frequency transformation matrix $\Psi(f_i,\theta)$, used by TOPS, is defined as:

$$[\Psi(f_i,\theta)]_{(m,m)} = \exp\big(jf_i\frac{2\pi r}{\lambda}\cos(\theta_l - \varphi_m)\big) \quad (11)$$

The signal subspace $F_s(f_i)$ of the frequency range f_i is converted into another frequency group f_j using $\Psi(f_i,\theta)$, where the transformed signal subspace $U_{ij}(\theta)$ is defined as follows [12]:

$$U_{ij}(\theta) = \Psi(\Delta f,\theta)F_s(f_i),\ i \ne j, \quad (12)$$

where $\Delta f = f_j - f_i$. The Eq. (12) can be expressed as:

$$\begin{aligned} U_{ij}(\theta) &= \Psi(\Delta f,\theta)A(f_i,\theta)G(f_i) \\ &= A(f_j,\hat{\theta})G(f_i), \end{aligned} \quad (13)$$

where $\hat{\theta}$ is the transformed θ by using the frequency transform matrix $\Psi(f_i,\theta)$, and $G(f_i)$ is a full order square matrix that verifies $F_s(f_i) = A(f_i,\theta)G(f_i)$.

An array manifold at any frequency and DOA can be altered into another array manifold at a different frequency applying this transformation process. As a result, the converted matrix is a complete order matrix and can be utilized for the following orthogonality test between converted matrix and noise subspaces, which is discussed in details in [12].

Considering that the frequency range of interest is f_1, the estimator matrix $Q_{TOPS}(\theta)$ is given as [16]:

$$\begin{aligned} Q_{TOPS}(\theta) \quad = \quad & \left[U_{12}^H(\theta)F_n(f_2),\ U_{13}^H(\theta)F_n(f_3),\right. \\ & \left.\ldots,\ U_{1K}^H(\theta)F_n(f_K)\right] \end{aligned} \tag{14}$$

The effectiveness of the computed covariance matrix, which is basically determined by the number of snapshots and the SNR of the received signal, proves the efficiency of the DOA estimation. The subspace projection technique is applied in TOPS signal processing to minimize signal subspace component leakage in the estimated noise subspace. Then, the projection matrix $P_i(\theta)$ is calculated as [16]:

$$P(f_i,\theta) = I - \left(a^H(f_i,\theta)a(f_i,\theta)\right)^{-1}a(f_i,\theta)a^H(f_i,\theta), \tag{15}$$

where I is an $M \times M$ identity matrix. Then, we acheive the noise robust matrix $Q'_{TOPS}(\theta)$ substituting the component $U_{ij}(\theta)$ of the Eq. (14) by a new converted signal subspace matrix $V_{ij}(\theta)$.

$$V_{ij}(\theta) = P_j(\theta)U_{ij}(\theta) \tag{16}$$

$$\begin{aligned} Q'_{TOPS}(\theta) \quad = \quad & \left[V_{12}^H(\theta)F_n(f_2),\ V_{13}^H(\theta)F_n(f_3),\right. \\ & \left.\ldots,\ V_{1K}^H(\theta)F_n(f_K)\right] \end{aligned} \tag{17}$$

TOPS performs well when applying the equation (17) with $Q'_{TOPS}(\theta)$ because the projection matrix $P_i(\theta)$ eliminates subspace estimation errors.

$$\hat{\theta} = arg\ \max_{\theta} \frac{1}{\sigma_{min}(\theta)} \tag{18}$$

where $\sigma_{min}(\theta)$ is the shortest singular value of $Q'_{TOPS}(\theta)$.

III. Simulation outcomes

To examine the effectiveness of these approaches on a small antenna array of $M = 5$ elements, four various examples using Matlab software $(2020a)$ are demonstrated. The received signals are divided into 256 blocks, and the number of snapshots is assumed to be 500 samples. The signals have the same range of frequencies. The noise level is lower in the first scenario, and the sources are close. We utilize a far signals and a low noise level in the second example. The noise level is high in the third scenario, and the sources are close together. In the last example, the noise level is high, and the sources are thought to be distant.

A. Wideband DOA estimation in noiseless case with nearest signals

The Fig. 2 presents the simulation outcomes of TOPS and TOFS approaches, under a low level of noise (SNR$= 20dB$). Three angles are choosen $\{\theta_1 = 15°, \theta_2 = 75°, \theta_3 = 80°\}$, with the last two signals are considered closely placed by $\Delta\theta = 5°$.

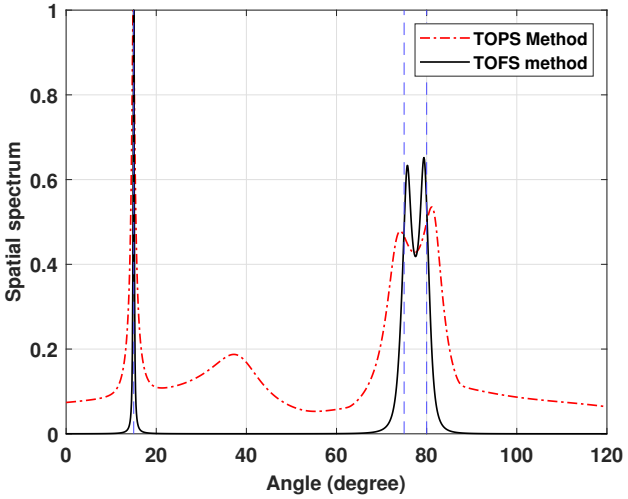


Fig. 2: Wideband DOA estimation using TOPS and TOFS methods with SNR$= 20dB$ and $\Delta\theta = 5°$.

From Fig. 2, we remark that both methods are capable to detecte the DOAs correctly. However, TOFS outperforms TOPS in this case, and can estimate the real angles more precisely. The TOPS method generates false peaks that can effect on the performance of DOA estimation.

B. Wideband DOA estimation in noiseless case with farest signals

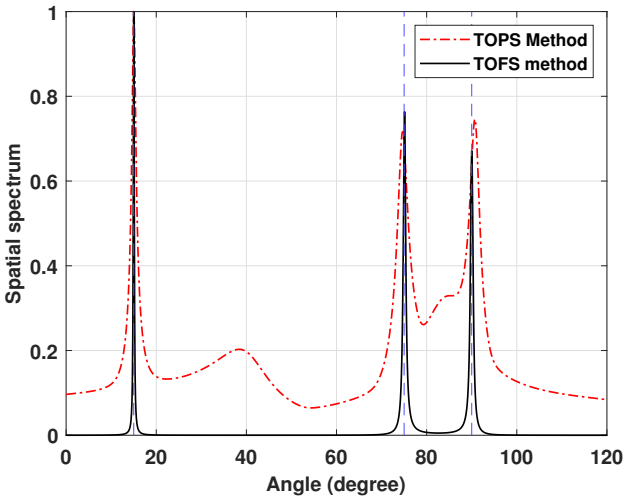


Fig. 3: Wideband DOA estimation using TOPS and TOFS methods with SNR$= 20dB$ and $\Delta\theta = 15°$.

The simulation results of TOPS and TOFS algorithms are depicted in Fig. 3, under a low level of the noise (SNR= $20dB$). In this case the DOAs are $\{\theta_1 = 15°, \theta_2 = 75°, \theta_3 = 90°\}$, the last two signals are spaced by $\Delta\theta = 15°$.

This figure proves that the approaches are able to estimate the actual DOAs correctly. More undesired angles are observable in the spectrum of TOPS, while the resolution accuracy of TOFS becomes more effective and performant, due to the signals are separated well.

C. Wideband DOA estimation in noisy case with nearest signals

The Fig. 4 depicts the wideband DOA estimation of TOPS and TOFS techniques, under a high noise level (SNR= $3dB$). The three angles are $\{\theta_1 = 15°, \theta_2 = 80°, \theta_3 = 85°\}$. The last two signals are separated by $\Delta\theta = 5°$.

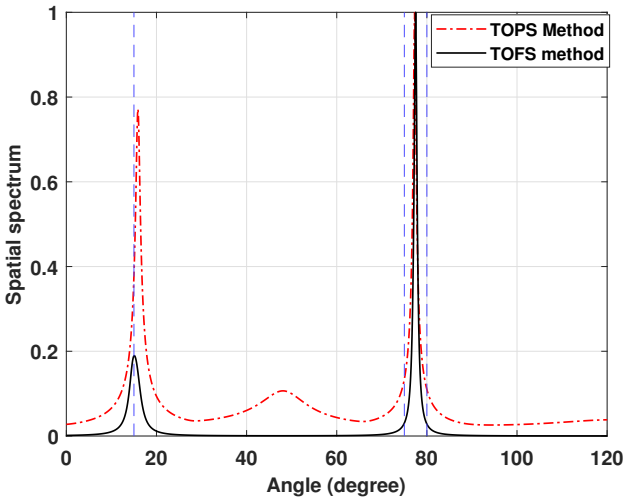


Fig. 4: Wideband DOA estimation using TOPS and TOFS methods with SNR= $3dB$ and $\Delta\theta = 5°$.

From Fig. 4, it is evident that the algorithms are not able to calculate the two closest DOAs effectively, due to these methods suffers when the floor is noisy and also when the signals are close.

D. Wideband DOA estimation in noisy case with farest signals

The Fig. 5 presents the simulations outcomes of TOPS and TOFS approaches, under a high level of the noise (SNR= $3dB$). The three DOAs are $\{\theta_1 = 15°, \theta_2 = 75°, \theta_3 = 90°\}$, with the last two signals are spaced by $\Delta\theta = 15°$.

From this figure, we observe that only TOPS can separate between the angles correctly and can estimate the DOAs, due to TOPS work better in a noisy environment, while TOFS is the best when the SNR is mid-to-high.

IV. Conclusion

This paper presents several simulations to show the efficiency of two incoherent wideband methods on a small antenna array, under different levels of noise and distance between adjacents antenna elements. The number of antenna

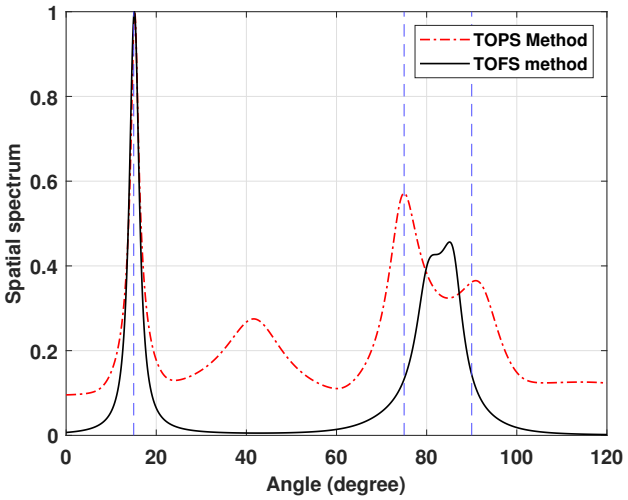


Fig. 5: Wideband DOA estimation using TOPS and TOFS methods with SNR= $3dB$ and $\Delta\theta = 15°$.

elements have a great impact on the type of array configuration that will be selected. For example, the performance of a circular array is affected more than other geometries when the number of elements N are reduced due to the nature of the arrays are arranged. As a results, we used a small array of 5 antenna elements. The simulation outcomes show that the wideband DOA estimation accuracy of TOFS is better than the TOPS approach when the SNR is mid-to-high, while TOPS surpasses TOFS in a noisy environment. It is also proves that TOFS can eliminate all unwanted peaks, while TOPS cannot.

References

[1] S. Ebihara, Y. Kimura and T. Shimomura, "Coaxial-fed circular dipole array antenna with ferrite loading for thin directional borehole radar sonde", *IEEE Transactions on Geoscience and Remote Sensing*. vol. 53, no. 4, pp. 1842–1854, 2015.

[2] U. Nielsen and J. Dall, "Direction-of-arrival estimation for radar ice sounding surface clutter suppression", *IEEE Transactions on Geoscience and Remote Sensing*. vol. 53, no. 9, pp. 5170–5179, 2015.

[3] R. Takahashi, T. Inaba and T. Takahashi, "Digital monopulse beamforming for achieving the CRLB for angle accuracy" *IEEE Transactions on Aerospace and Electronic Systems*. vol. 54, no. 1, pp. 315–323, 2018.

[4] Z.M. Liu, Z.T. Huang and Y.Y. Zhou, "Direction-of-arrival estimation of wideband signals via covariance matrix sparse representation", *IEEE Transactions on Signal processing*. vol. 59, no. 9, pp. 4256–4270, 2011.

[5] A. Toktas and A. Akdagli, "Compact multiple-input multiple-output antenna with low correlation for ultra-wideband applications", *IET Microwaves, Antennas and Propagation*. vol. 9, no. 8, pp. 822–829, 2015.

[6] M. Wax and T. Kailath, "Spatio-temporal spectral analysis by eigenstructure methods", *IEEE Transactions on Acoustics, Speech, and Signal Processing*. vol. 32, no. 4, pp. 817–827, 1984.

[7] H. Wang and M. Kaveh, Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources. IEEE Trans. Acoust. Speech Signal Process, **33**(4), 823–831 (1985).

[8] M.A. Doron and A.J. Weiss, "On focusing matrices for wide-band array processing", *IEEE Transactions on Signal Processing*. vol. 40, no. 6, pp. 1295–1302, 1992.

[9] F. Sellone, "Robust auto-focusing wideband DOA estimation", *Signal Processing*. vol. 86, no. 1, pp. 17–37, 2006.

[10] D. Swingler and J. Krolik, "Source location bias in the coherently focused high-resolution broadband beamformer", *IEEE Transactions on Acoustics, Speech, and Signal Processing*. vol. 37, no. 1, pp. 143–145, 1989.

[11] H. Yu, J. Liu, Z. Huang, Y. Zhou and X. Xu, "A new method for wideband DOA estimation", *in 2007 International Conference on Wireless Communications, Networking and Mobile Computing*. pp. 598–601, 2007.

[12] Y.S. Yoon, L.M. Kaplan and J.H. McClellan, "TOPS: New DOA estimator for wideband signals", *IEEE Transactions on Signal processing*. vol. 54, no. 6, pp. 1977–1989, 2006.

[13] P.J. Chung, J.F. Bohme, C.F. Mecklenbrauker and A.O. Hero, "Detection of the number of signals using the Benjamini-Hochberg procedure", *IEEE Transactions on Signal Processing*. vol. 55, no. 6, pp. 2497–2508, 2007.

[14] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria", *IEEE Transactions on Acoustics, Speech, and Signal Processing*. vol. 33, no. 2, pp. 387–392, 1985.

[15] H. Hayashi and T. Ohtsuki, "DOA estimation for wideband signals based on weighted Squared TOPS", *EURASIP Journal on Wireless Communications and Networking*. vol. 2016, no. 1, pp. 1–12, 2016.

[16] A. Asadzadeh, S.M. Alavi, M. Karimi and H. Amiri, "Coherent wideband signals DOA estimation by the new CTOPS algorithm", *Multidimensional Systems and Signal Processing*. vol. 31, no. 3, pp. 1075–1089, 2020.

K-plane algorithm for identification of switched linear Systems

A. Bourzik[1], S. Safi[1], B. Bouikhalene[1], M. Pouliquen[2], M. Frikel[2]
[1] *LIMATI Laboratory, Polydisciplinary Faculty*
Sultan Moulay Slimane University, Beni Mellal, Morocco
[2] *LIS laboratory, ENSICAEN*
Caen Normandie Univesity, Caen, France
Email: *abdelaaatibourzik@gmail.com*

***Abstract*—This paper focuses on identification of switched linear Systems. Our idea is to formulate this identification problem as a subspace clustering problem. More precisely, in this paper we are going to view the K-plane algorithm as a solution for the identification of switched linear systems. The K-plane algorithm is an extension of K-means algorithm: while the K-means algorithm clusters points using centroid, the K-plane algorithm cluster them using planes. This algorithm is based on assigning each point to the nearest plane, and then estimate the parameters of each plane using an iterative way, based on the euclidean distance between points and planes. Some simulation results are presented for different configurations in order to see how these configurations influence the estimation of subsystem parameters.**

***Index Terms*—System Identification, switched linear systems, subspace clustering, K-plane**

I. Introduction

This paper addresses the problem of the identification of switched linear systems using a subspace clustering approach especially the K-plane algorithm. This identification problem combines two important research areas in a single one, it combines the identification problem and the subspace clustering problem.

The class of switched linear systems is an important and popular class of hybrid systems, because of its ability to model numerous complex phenomena. Whereas the identification of this type of systems is strongly needed in the application area, hence the identification of switched linear systems has several interesting presented methods using different approaches, including : clustering-based approach (ex. [1], [2] and [3]), algebraic and geometric approach like in ([4], [5]), recursive approach (ex. [6]). The idea in this paper is to formulate the identification problem as a clustering problem where the observation vector is interpreted as a point in space. A such approach is already used in [7], that uses the supervised learning algorithm such as Support Vector Machines (SVM) to formulate the identification problem as a classification problem. In our approach we are going to use unsupervised learning algorithm such as subspace clustering. There exist several subspace clustering algorithms in literature, including iterative methods (ex. K-means, K-planes and K-subspaces [8]), algebraic methods (ex. Generalized principal component analysis GPCA [9]), statistical methods (ex. Random sample consensus (RANSAC) [10]) and spectral methods [11][12]. In the presented work we use the clustering approach for identification of switched linear systems, using an iterative based subspace clustering method, especially the K-plane algorithm that is considered as an extension of K-means.

The K-plane algorithm generalizes the K-means algorithm. The K-means clustering algorithm assigns N data points into n_s clusters so that similar data points can be grouped together around the nearest center. The K-plane algorithm assigns N data points into n_s hyper-planes so that similar data points can be grouped together on the same hyperplane. The K-plane algorithm has been first introduced in [13]. The K-plane algorithm have a lot of applications especially in clustering and identification problems, also used for more purposes like for piece-wise linear regression [14], which can learn continuous as well as discontinuous piece-wise linear functions, the main idea is to repeatedly partition the data and learn a linear model in each partition. Before that it is used for identification of discrete time hybrid systems in the Piece-Wise Affine (PWA) systems [3], this problem involves the estimation of both the parameters of the affine sub-models and the partition of the PWA map from data. Also some of the newest researches using K-plane to design methodologies for style clustering on stylistic data, where each cluster depends on both the similarities between data samples and its latently or apparently distinguishable style [15]. In [16] a solution is proposed for eigenvalues problems especially while the data points are almost on the same hyperplane, where normal K-plane clustering fails. Therefore, the K-plane is one of iterative methods of subspace clustering, that are largely used and had some advantages: their simplicity since they alternates between assigning points to subspaces and estimating the subspace parameters, it can handle both linear and affine subspaces explicitly and converges to a local optimum in a finite number of iterations. However, iterative methods are suffering from some drawbacks: it requires the number of subspaces and the dimensions to be known beforehand, either it depends on an initialization step.

The paper is organized as follows: the considered identi-

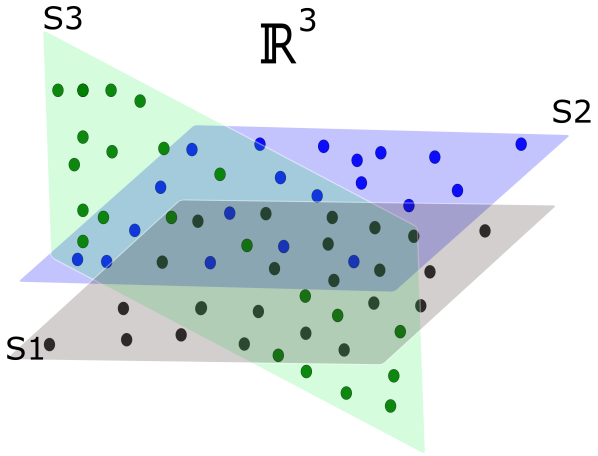


Fig. 1: A set of sample points $\{(y_t;\phi_t)\}_{t=1}^N$ in $\mathbb{R}^3$ drawn from an union of three subspaces.

fication problem and proposed algorithm are formulated in Section II, especially the considered problem and notation in subsection II-A and the proposed algorithm in subsection II-B. Experimental results and discussion are given in section III. Section IV concludes the paper and gives some perspectives.

II. Notation and proposed algorithm

A. Notation and considered identification problem

Consider a discrete time switched system parameterized as follows:

$$y_t = \phi_t^T \theta_{\lambda_t} + \mu_{\lambda_t} + e_t \tag{1}$$

where

- $\{y_t\}_{t=1}^N$ is the output sequence of the system and N represents the number of samples.
- $\phi_t \in \mathbb{R}^{d+1}$ is the observation vector at time t defined as:

$$\phi_t = \begin{pmatrix} u_t \\ \vdots \\ u_{t-d} \end{pmatrix} \tag{2}$$

  with $\{u_t\}_{t=1}^N$ the input sequence of the system and d is the order of the system.
- $\theta_j \in \mathbb{R}^{d+1}$ is the parameter vector associated to the j$^{\text{th}}$ subsystem with $j \in \{1,\dots,n_s\}$ and n_s is the number of subsystems.
- $\lambda_t \in \{1;...;n_s\}$ is the label (also known as the discrete state) of the active subsystem at time t.
- μ_j is the intercept of the subsystem j.
- $\{e_t\}_{t=1}^N$ is the noise sequence which is assumed to be a zero mean sequence uncorrelated with $\{u_t\}_{t=1}^N$ and $\{\lambda_t\}_{t=1}^N$.

Fig.1 represents such a data collection with $d=1$, $n_s=3$ and where the basis represents components of $\{\phi_t\}_{t=1}^N$ and the vertical axis represents $\{y_t\}_{t=1}^N$.

In the following we consider $\mu_j = 0 \ \forall \ j$ for convenience and λ_t is assumed to be unknown. Given d and n_s, the **objective** of this paper is to realize the estimation of the parameter vectors $\{\theta_j\}_{j=1}^{n_s}$ using $\{(y_t;\phi_t)\}_{t=1}^N$.

Such an identification problem corresponds to a subspace clustering problem. The main goal of subspace clustering is to estimate the number of subspaces n_s, the dimension d, the parameter vectors $\{\theta_j\}_{j=1}^{n_s}$ of each subspace and the label λ_t. It is clear that, if $n_s = 1$, then this estimation problem can easily be solved. However, when $n_s > 1$, the subspace clustering problem becomes significantly more difficult due to a number of challenges as discussed in [17].

Among the various subspace clustering algorithms, the K-plane algorithm can solve the two following problems: (1) this algorithm can estimate, for each point $(y_t;\phi_t)$, the subspace membership, i.e. the active subsystem, (2) this algorithm can estimate the equation of each subspace, i.e. the parameter vectors in our case. Therefore the K-plane is applied in the next subsection on our identification problem.

B. The proposed algorithm

The K-plane algorithm is an iterative algorithm: several steps are iterated until convergence of the algorithm. Let consider the following two-step process:

step 1

In this first step we assume that we know estimates $\widehat{\theta}_j$ for $j \in \{1,\dots,n_s\}$.

For each time t it is possible to estimate the output of each subsystem. This output is noticed $\widehat{y}_{t/j}$ and is defined by:

$$\widehat{y}_{t/j} = \phi_t^T \widehat{\theta}_j \tag{3}$$

From these estimates it is possible to estimate the active subsystem at each time t, i.e. it is possible to estimate λ_t. This estimate corresponds to the label of the nearest subsystem, among subsystems parameterized by $\widehat{\theta}_j$, from y_t. $\widehat{\lambda}_t$ is thus defined by

$$\widehat{\lambda}_t = argmin_{j\in[1,n_s]} d_{t/j} \tag{4}$$

where $d_{t/j}$ is the distance between y_t and the subsystem parameterized by $\widehat{\theta}_j$:

$$d_{t/j} = |y_t - \widehat{y}_{t/j}| \tag{5}$$

step 2

In this second step we assume that we know estimates $\widehat{\lambda}_t$ for $t \in [1;N]$.

The estimation of the parameter vector θ_j can be done using the least squares algorithm gathering

all the points $(y_t; \phi_t)$ which belongs to the j^{th} subsystem:

$$\widehat{\theta}_j = argmin_{\theta_j} \sum_{t/\widehat{\lambda}_t=j} |y_t - \phi_t^T \theta_j|^2 \qquad (6)$$

It might be noticed that the first step depends on estimates provided in the second step, and second step depends on estimates provided in the first step. It follows that the implementation of these two steps requires an iterative process. This iterative process is described in Algorithm 1.

Algorithm 1 Identification algorithm

Input: The set of points $\{(y_t; \phi_t)\}_{t=1}^N$, the number of subsystems n_s and the order d.

1: Initialize the parameter vector of each subsystem j randomly.
2: For each time t, implementation of the first step in order to estimate λ_t (see (3), (4) and (5)).
3: For each j, implementation of the second step in order to estimate θ_j (see (6)).
4: Repeat 2 and 3 until convergence.

Output: Assign each point to the closest estimated subsystem and estimate of the parameter vector for each subsystem.

As discussed in the introduction, the K-plane algorithm is very sensitive to the initialization of parameters: its behavior depends on the initialization. Therefore we advise to implement this algorithm several times and to keep only the parameter vectors that are obtained with several different initialization.

III. Experimental results and discussion

In this section we give some numerical results in order to evaluate performance of proposed identification algorithm. We choose the order of the system as $d = 1$. Different configurations are tested: different values of Signal to Noise Ration (SNR), different values for n_s, different values for N.

For each numerical experiment, the input sequence is $\{u_t\}_{t=1}^N$ is a random integers sequence drawn from the discrete uniform distribution on the interval $[-10, 10]$.

A. First experiment

In this first experiment we consider a switched linear system such that $d = 1$, $ns = 2$. The parameter vectors θ_1 and θ_2 are given by:

$$\theta_1 = \begin{pmatrix} 0.948 \\ -0.316 \end{pmatrix} \quad , \quad \theta_2 = \begin{pmatrix} 0.820 \\ 0.533 \end{pmatrix} \qquad (7)$$

The sequence $\{\lambda_t\}_{t=1}^N$ is randomly generated. In this first experiment $N = 100$ and we choose the SNR values as 3dB, 10dB , 30dB, 50dB, 100dB and 300dB.

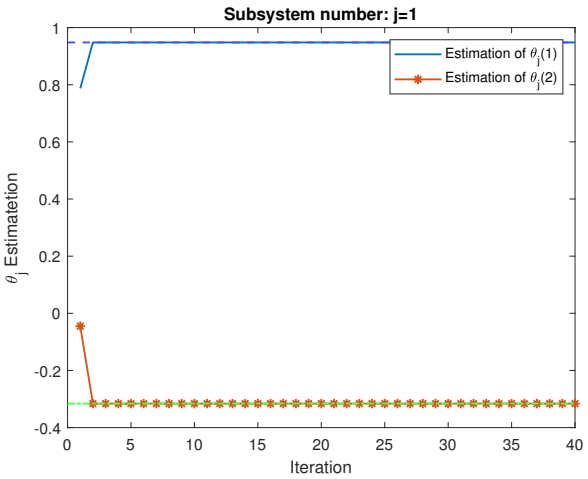


Fig. 2: First experiment: estimate of θ_1 as function of iterations for $N = 100$ and SNR$= 300dB$.

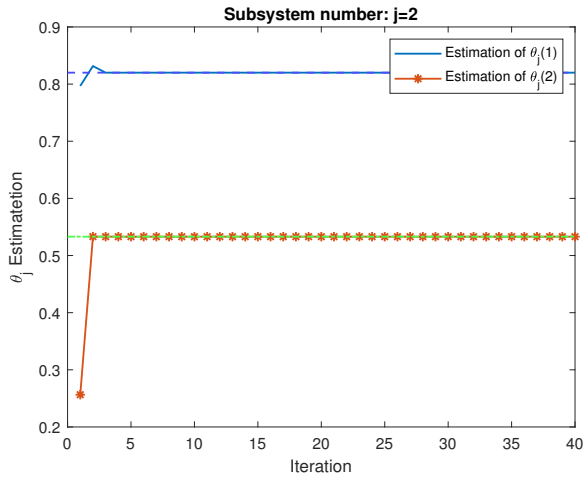


Fig. 3: First experiment: estimate of θ_2 as function of iterations for $N = 100$ and SNR$= 300dB$.

In the case SNR$=$ 300dB, Fig.2 shows the estimated parameters of the first subsystem as a function of number of iterations, and Fig.3 shows the estimated parameters of the second subsystem. These results show that the proposed algorithm allows the estimation of parameter vectors.

We keep the same previous system and the number of points at $N = 100$ and we test the presented algorithm with different SNR. Tab I gives estimates of θ_1 and θ_2 for different levels of SNR. These results show that the algorithm is efficient even in presence of noise.

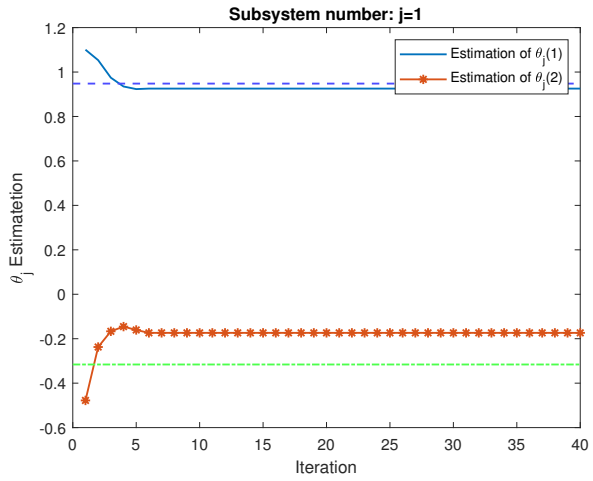


Fig. 4: Third experiment: estimate of θ_1 as function of iterations for $N = 100$ and SNR$= 300dB$.

	$\theta_1(1)$	$\theta_1(2)$	$\theta_2(1)$	$\theta_2(2)$
Real value	0.948	-0.316	0.820	0.533
SNR =3dB	0.994	-0.445	0.783	0.565
SNR =10dB	0.931	-0.331	0.795	0.584
SNR =30dB	0.948	-0.311	0.822	0.534
SNR =50dB	0.947	-0.315	0.820	0.532
SNR =100dB	0.948	-0.316	0.820	0.533
SNR =300dB	0.948	-0.316	0.820	0.533

TABLE I: First experiment: estimate of θ_1 and θ_2 as function of SNR.

B. Second experiment

In this second experiment we consider the same system than in the first experiment and we increased the number of points to $N = 1000$. We directly test the algorithm for different noise level. Tab II gives estimates of θ_1 and θ_2 for different levels of SNR. We can see that for $N = 1000$ the algorithm performs better than for $N = 100$. We can say that if there are more available points, then the estimation is better and fast.

	$\theta_1(1)$	$\theta_1(2)$	$\theta_2(1)$	$\theta_2(2)$
Real value	0.948	-0.316	0.820	0.533
SNR =3dB	0.985	-0.437	0.818	0.727
SNR =10dB	0.948	-0.307	0.836	0.578
SNR =30dB	0.949	-0.316	0.820	0.536
SNR =50dB	0.948	-0.316	0.820	0.533
SNR =100dB	0.948	-0.316	0.820	0.533
SNR =300dB	0.948	-0.316	0.820	0.533

TABLE II: Second experiment: estimate of θ_1 and θ_2 as function of SNR.

C. Third experiment

We consider another switched linear system such that $d = 1$, $N = 100$ and $n_s = 10$. The two first subsystems have the same parameters than in the first experiment (see (7)), parameters of the other subsystems have been randomly generated. We consider here the noise free case.

The estimated parameters for the first subsystem are given in Fig.4. Compare with Fig.2, it appears that the estimation is more difficult with an increasing number of subsystems.

D. Fourth experiment

In order to test the influence of the number of available data and the noise level we have realized different estimations of parameter vector, for each subsystem with different values of N, n_s and SNR. As shown in the figure (Fig.5) for every configuration (a) and (b) as described in the caption of the figure, we change the values of SNR and see how many iterations takes the algorithm to converge to the right subsystem parameters or near to it, in order to test the convergence of the algorithm versus the noise level.

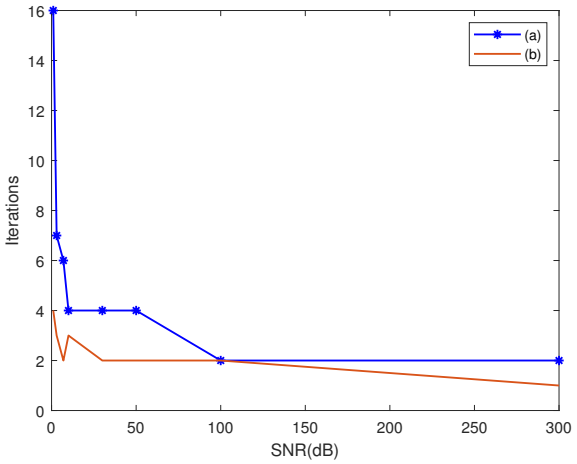


Fig. 5: Number of convergence iterations as a function of SNR, for (a): $N = 1000$ and $n_s = 2$, (b): $N = 10$ and $n_s = 2$.

The figure (Fig.5) indicate that the k-plane is not very sensitive to noise, especially when the number of subsystems is small it can estimate the subsystem parameters even if the SNR ratio is important like in (b), anyway the algorithm keeps converging in a finite number of iterations, and in a small period of time. Generally K-plane and other extensions of iterative methods like K-means, K-subspaces, are considered as a very simple way of improving the performance of algebraic algorithms in the case of noisy data [17]. The K-plane algorithm is very good when dealing with noisy data sequence, but steel suffer from a big drawback added to its sensitive to initialization is that it needs to give the number of planes (subsystems) in advance.

E. Fifth experiment

In all the previous experimental results, in all the cases, the algorithm converge in a finite number of iteration that is not

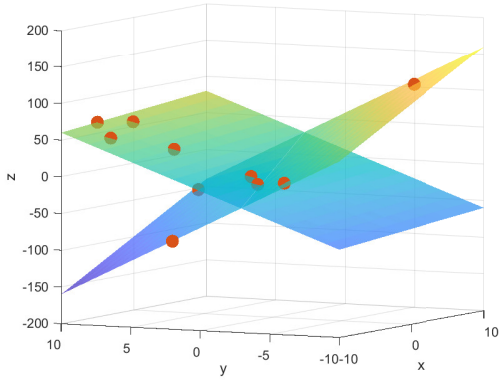


Fig. 6: Fourth experiment: clustering of $N = 10$ points characterizing $n_s = 2$ subsystems in the noise free case.

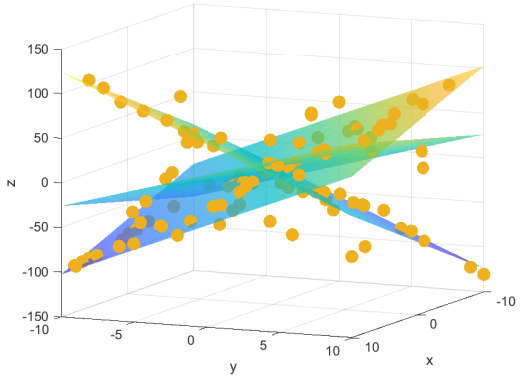


Fig. 7: Fourth experiment: clustering of $N = 100$ points characterizing $n_s = 3$ subsystems in the noisy case (SNR=30dB).

very big (less than 20 iterations), this demonstrates the good performance of the K-plane algorithm in the identification of linear system of order $d = 1$.

After assigning each point to the nearest subsystem based on the euclidean distance in the equation (5), and after estimating the parameters of each subsystem we can use these parameters to show these subsystems as clusters, which are plotted as surfaces that considered as subspaces, as figured in (Fig. 6) and (Fig. 7). From these figures it is clear that points belong to hyperplane characterizing the identified subsystems. Each plane is considered as a subsystem that has two parameters $\theta_j(1)$ and $\theta_j(1)$, that are estimated using the previous algorithm.

IV. Conclusions and perspective

In this paper we used the K-plane algorithm as a solution for identification of switched linear systems problem. This algorithm assigns each point to the closest estimated subsystem and estimate the parameter vector for each subsystem. In addition, an evaluation of this algorithm is done with different configurations: different number of points, different number of subsystems and different level of noise. The K-plane algorithm appears to be useful, especially in the case of noisy data due to it's iterative refinement. Future work will be localized on the comparison of the present algorithm with other identification algorithms, in addition trying to find an indicator about the number of subsystems without offering it in advance.

References

[1] Hayato Nakada, Kiyotsugu Takaba, and Tohru Katayama. Identification of piecewise affine systems based on statistical clustering technique. *Automatica*, 41(5):905–913, 2005.

[2] Giancarlo Ferrari-Trecate, Marco Muselli, Diego Liberati, and Manfred Morari. A clustering technique for the identification of piecewise affine systems. *Automatica*, 39(2):205–217, 2003.

[3] Mojtaba Tabatabaei-Pour, Karim Salahshoor, and Behzad Moshiri. A modified k-plane clustering algorithm for identification of hybrid systems. In *2006 6th World Congress on Intelligent Control and Automation*, volume 1, pages 1333–1337. IEEE, 2006.

[4] Yi Ma and René Vidal. Identification of deterministic switched arx systems via identification of algebraic varieties. In *International Workshop on Hybrid Systems: Computation and Control*, pages 449–465. Springer, 2005.

[5] Fabien Lauer, Gérard Bloch, et al. Identification of linear hybrid systems: A geometric approach. In *2013 American Control Conference*, pages 830–835. IEEE, 2013.

[6] Yasmin Hashambhoy and René Vidal. Recursive identification of switched arx models with unknown number of models and unknown orders. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 6115–6121. IEEE, 2005.

[7] Abdelhak Goudjil, Mathieu Pouliquen, Eric Pigeon, Olivier Gehan, and Mohammed M'Saad. Identification of systems using binary sensors via support vector machines. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 3385–3390. IEEE, 2015.

[8] Brady Lund and Jinxuan Ma. A review of cluster analysis techniques and their uses in library and information science research: k-means and k-medoids clustering. *Performance Measurement and Metrics*, 2021.

[9] Rene Vidal, Yi Ma, and Shankar Sastry. Generalized principal component analysis (gpca). *IEEE transactions on pattern analysis and machine intelligence*, 27(12): 1945–1959, 2005.

[10] Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[11] Andrew Ng, Michael Jordan, and Yair Weiss. On spectral clustering: Analysis and an algorithm. *Advances in neural information processing systems*, 14, 2001.

[12] Ulrike Von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, 17(4):395–416, 2007.

[13] Paul S Bradley and Olvi L Mangasarian. K-plane clustering. *Journal of Global optimization*, 16(1):23–32, 2000.

[14] Naresh Manwani and PS Sastry. K-plane regression. *Information Sciences*, 292:39–56, 2015.

[15] Suhang Gu, Yusuke Nojima, Hisao Ishibuchi, and Shitong Wang. Fuzzy style k-plane clustering. *IEEE Transactions on Fuzzy Systems*, 29(6):1518–1532, 2021. doi: 10.1109/TFUZZ.2020.2979676.

[16] Yuan-Hai Shao, Lan Bai, Zhen Wang, Xiang-Yu Hua, and Nai-Yang Deng. Proximal plane clustering via eigenvalues. *Procedia Computer Science*, 17:41–47, 2013.

[17] René Vidal. Subspace clustering. *IEEE Signal Processing Magazine*, 28(2):52–68, 2011.

Positive Definite Kernels for Minimum Phase Channel Identification from Binary Output

Rachid Fateh*, Anouar Darif and Said Safi
LIMATI Laboratory, Polydisciplinary Faculty, Sultan Moulay Slimane University
PO Box 592, Beni Mellal, Morocco.
Email: fateh.smi@gmail.com (R.F)

***Abstract*—In several scientific fields, it is very important to have a representation of the system in order to control or optimize its function. In this paper, we consider the identification problem of finite impulse response (FIR) for single-input single-output nonlinear systems under perturbations and with binary-valued measurements. In the one hand, we have used an algorithm based on the framework of reproducing kernel Hilbert Spaces (RKHS) to identify the impulse response of the minimum phase channel. Then, the convergence analysis of this algorithm is presented in the presence of noise. Some examples of experimental illustrative result are provided in this work to demonstrate the effectiveness of the kernel affine projection algorithm (KAPA) in nonlinear system identification from output binary measurements.**

***Index Terms*—Finite impulse response, Nonlinear systems, Identification, Reproducing kernel Hilbert spaces, Positive defined kernel**

I. Introduction

Over the last two decades, the area of signal processing has experienced a significant improvement due to the development of new techniques for the handling of non-linear problems. Among the machine learning applications, we can cite pattern recognition [1], acoustic echo cancellation [2], system identification [3], [4], channel equalization [5], time series analysis [6], as well as decisional issues such as classification and nonlinear regression [7]. The data available in these application fields, obtained from natural systems, are very complex and cannot be explained by traditional linear models. Consequently, the researchers have experienced the need to to propose non-linear algorithms that could handle a wide class of problems.

The non-linearity of engineering applications often motivates the use of control techniques based on a non-linear model. As in literature, numerous forms of input/output type nonlinear framework presented; Commonly, the Volterra-Wiener model [8], [9], the Hammerstein model [10], [11], the NARMAX model [12], and so on. In recent years, a lot of attention has been focused on kernel methods [13], [14]. Indeed, many studies on the employment of the kernel approach, essentially in modeling complex nonlinear systems, and channel identification have been issued (see e.g. [5], [15]–[25]). This kind of modeling is performed by the representation of the considered process by scalar products, which represent a mathematical model based on the architecture and the operation of the non-linear system. Kernel methods are characterized by their ability to adapt and process information, which makes them able to model the most common real systems. These techniques exploit the theory of reproducing kernels. The main idea is the kernel trick, allowing to transform the data by the help of a nonlinear application, in a high dimensional space known as the Reproducing Kernel Hilbert Space (RKHS) [13], [26], where linear methods can be enforced. Many adaptive kernel filtering algorithms were developed so far in the scientific literature. Among them are the kernel least mean squares (KLMS) [27], and kernel recursive least squares (KRLS) [28]. To enhance the reliability of these approaches, a number of their subtypes were also proposed, including the normalized kernel least-mean square (KNLMS) [29], complex kernel LMS (CKLMS) [30], quantized KLMS (QKLMS) [31], fixedbudget QKLMS (QKLMS-FB) [32], set-membership NLMS (SM-KNLMS) [33], extended KRLS (Ex-KRLS) [34], sliding-window KRLS (SW-KRLS) [35], and so on. In this paper, we concentrate on finite impulse response nonlinear system identification with binary values output observations. Firstly, we use an algorithm based on the positive definite kernel to estimate recursively the channel parameters. Secondly, the mean square error (MSE) will be employed to evaluate the precision of the estimated values.

The remainder of the paper is organized as follows: It starts in Section II with a statement of the identification problem. A thorough description of the kernel affine projection algorithms is given in Section III, and in Section IV it is implemented to a finite impulse response nonlinear system with binary-valued measurements. Finally, in Section V, we give the conclusion.

II. The considered identification problem

In this section, we describe the structure of the non-linear model that will be used for the identification problem of finite impulse response (FIR) single-input single-output (SISO) communication channels. It is obtained by cascading a static nonlinear block and a linear dynamic model. The general Hammerstein is a popular model to represent the nonlinear system's behavior. The Figure 1 illustrates that this model consists of two blocks.

The following equations are used to represent the system

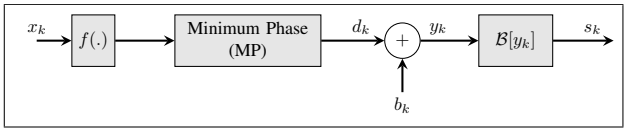


Fig. 1: Block of the Hammerstein system with binary outputs and noises.

described in Fig. 1:

$$y_k = \sum_{i=0}^{L-1} h_i f(x_{k-i}) + b_k \tag{1}$$

where x_k and y_k designate the input signal and the output system at time k respectively, $\{h_i\}_{i=0}^{L-1}$ is the impulse response coefficients (FIR filter), b_k is a random noise and $f(.)$ the nonlinear characteristic.
The system output y_k, is detected by the binary sensor $\mathcal{B}[.]$, with a finite threshold $C \in \mathbb{R}$, as in the following mathematical formula:

$$s_k = \mathcal{B}[y_k] = \begin{cases} 1 & \text{if } y_k \geq C, \\ -1 & \text{otherwise}. \end{cases} \tag{2}$$

Therefore, to improve the evaluation of the system and to achieve a successful minimum phase channel identification result, we suppose that:

(**A1**) The input sequence x_k, is independent and identically distributed bounded random process with zero mean.
(**A2**) The measurement noise b_k is Gaussian and independent of the system input x_k and output y_k
(**A3**) There is no delay in the system, i.e. $h_0 \neq 0$.
(**A4**) The C-value is available (i.e. known).

III. Kernel Affine Projection Algorithm

In this section, we describe the kernel affine projection algorithm (KAPA) [38]. It belongs to the class of stochastic gradient algorithms. Their main idea is to perform the linear affine projection algorithms (APA) [38] in the kernel feature space, which is associated to the positive definite kernel κ, through the characteristic map $\Psi(.)$ as defined by the following equation:

$$\begin{aligned} \Psi \quad : \quad & \mathcal{X} \longrightarrow \mathcal{H} \\ & x \longrightarrow \kappa(x, .) \end{aligned} \tag{3}$$

where $\kappa(x, .)$ is a reproducing kernel, denoting the inner product $\langle \Psi(x), . \rangle$, $\mathcal{X}$ is the input space, and $\mathcal{H}$ is the higher-dimensional space (nonlinear Hilbert space). Different kernel functions are available, such as polynomial kernel, linear kernel, Laplace kernel and radial based kernel (Gaussian). However, among these different types of kernels, the Gaussian kernel is the most popularly used Mercer kernel [13], due to its universal approximation property and its high stability in numerical terms:

$$\kappa(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right), \quad \forall x_i, x_j \in \mathcal{X}. \tag{4}$$

where $\sigma > 0$ is the kernel width.

The optimization problem of the KAPA algorithm, without the regularization term, can be expressed as :

$$\min_{\Psi \in \mathcal{H}} \sum_{i=1}^{n} (d_i - \Psi(x_i))^2 \tag{5}$$

where d_i is the desired response of the model at time i and $\Psi(x_i)$ is the response of the model at the i^{th} observation x_i. Based on the representation theorem [37], the solution of the optimization problem (5) is as follows:

$$\Psi(.)_n = \sum_{i=1}^{m} \alpha_{n,k} \kappa(., x_k) \tag{6}$$

where $\alpha_n = (\alpha_{n,1}, \alpha_{n,2}, ..., \alpha_{n,n})^\top$ is the vector of the model parameters. The combination of (6) and (5) leads to the following problem:

$$\alpha_n = \arg\min_{\alpha} \|\mathbf{d}_n - \mathbf{H}_n \alpha\|^2 \tag{7}$$

where $\mathbf{d}_n = (d_1, d_2, ..., d_n)^\top$ is the desired response vector until time n and $\mathbf{H}_n$ is an $n \times m$ matrix whose $(i, j)^{th}$ element was $\kappa(x_i, x_j)$. If we suppose that $(\mathbf{H}_n^\top \mathbf{H}_n)^{-1}$ is available, the resolution of the problem (7) is obtained by :

$$\alpha_n = (\mathbf{H}_n^\top \mathbf{H}_n)^{-1} \mathbf{H}_n^\top \mathbf{d}_n \tag{8}$$

The fundamental idea of the Affine Projection Algoritm (APA) is to take into consideration the last p observations of the model only $\{x_n, ..., x_{n-p+1}\}$. In this case, $\mathbf{d}_n$ is now a p-element column vector, $\mathbf{d}_n = (d_n, ..., d_{n-p+1})^\top$, and $\mathbf{H}_n$ a matrix of size $p \times m$ where the element (i, j) is $\kappa(x_{n-i+1}, x_j)$, given by :

$$\mathbf{H}_n = \begin{pmatrix} \kappa(x_n, x_1) & \cdots & \kappa(x_n, x_m) \\ \vdots & \ddots & \vdots \\ \kappa(x_{n-p+1}, x_1) & \cdots & \kappa(x_{n-p+1}, x_m) \end{pmatrix}$$

The optimal coefficient vector α_{n+1} is identified at time $n+1$, by minimizing $\|\alpha_{n+1} - \alpha_n\|^2$ under the constraint of posterior error nullification for the last p observations.

$$\begin{cases} \min_{\alpha_{n+1}} \|\alpha_{n+1} - \alpha_n\|^2, \\ \text{s.t } \mathbf{d}_{n+1} = \mathbf{H}_{n+1} \alpha_{n+1}, \end{cases} \tag{9}$$

where $\mathbf{d}_{n+1} = (d_{n+1}, ..., d_{n-p+2})^\top$, $\alpha_n = (\alpha_{n+1,1}, ..., \alpha_{n+1,m})^\top$ and $\mathbf{H}_{n+1}$ is always a matrix $p \times m$ whose (i, j) element is $\kappa(x_{n-i+2}, x_j)$.

Note that m indicates the model order. At time $n+1$, when a new data is received (x_{n+1}). The update of the model coefficients is performed as follows :

$$\alpha_{n+1} = \alpha_n + \eta \mathbf{H}_{n+1}^\top (\varepsilon I + \mathbf{H}_{n+1} \mathbf{H}_{n+1}^\top)^{-1} (\mathbf{d}_{n+1} - \mathbf{H}_{n+1} \alpha_n) \tag{10}$$

As recommended in [36], a control parameter for the convergence step (η) and the regularization term (εI) have been introduced into this equation.

IV. Experimental Results

In this section, we present some numerical examples to illustrate the performance of the kernel affine projection algorithm for the SISO nonlinear system, by employing a minimum phase model. The numerical data were generated according to (1) and (2) with $C = 0.5$. The kernel algorithm was applied in different scenarios. For each scenario, the estimated FIR model order is $L = 3$. To implement the kernel algorithm, we selected the Gaussian kernel width $\sigma = 0.5$, the step-size parameter η equal to 0.2 and $\varepsilon = 0.001$. The input sequence x_k is a random sequence of zero mean with a uniform distribution in $[-1; 1]$. A Gaussian noise of zero mean is added to the output. The noise amplitude is adjusted to obtain a desired value of the signal-to-noise ratio (SNR). We define the SNR by the following relationship:

$$SNR = 10 \log_{10} \left[\frac{E[d_n^2]}{E[b_n^2]} \right] \tag{11}$$

where $E[.]$ is the mathematical expectation.
To measure the estimation quality, the following mean square error (MSE) was used:

$$MSE = \frac{1}{M} E \left[\| s_n - \alpha_{n-1}^{\top} \Psi(x_n) \|^2 \right] \tag{12}$$

where M being the total amount of Monte Carlo iteration.

A. Minimum Phase Channel Identification

In this model, we consider a third order channel, described by the RIF-PM model and given by the following equation :

$$\begin{cases} d_k = f(x_k) - 0.860 f(x_{k-1}) + 0.740 f(x_{k-2}) \\ \text{Zeros: } z_1 = 0.4300 + 0.7451i, z_2 = 0.4300 - 0.7451i, \end{cases} \tag{13}$$

The characteristics of this channel are represented in figure 2. It has two zeros, all of them inside the unit circle (i.e. minimum phase channel).

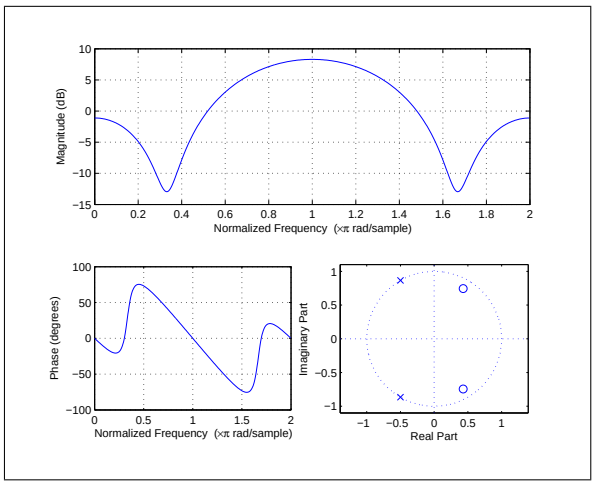


Fig. 2: Minimum phase channel characteristics

As a first simulation experiment, for the purpose of examining the performance of the KAPA algorithm, we select a data length $N = 1000$ and $SNR = 10dB$. The simulation results are shown in Figure 4.4. From this figure 3, we can conclude that the KAPA algorithm is able to find the minimum phase channel impulse response that is similar to the true impulse response.

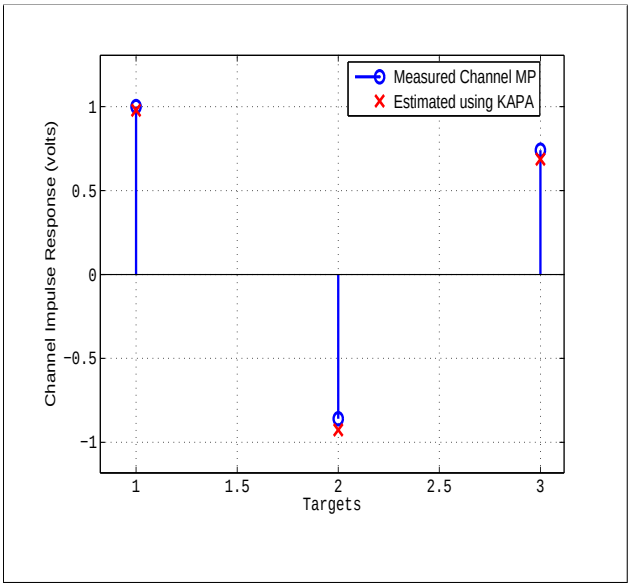


Fig. 3: Estimation of the minimum phase channel impulse response for $N = 1000$ and $SNR = 10dB$.

The next simulation experiment focuses on the influence of the data length N on the kernel affine projection algorithm performance. Note that N is a data length that has an impact on the estimated channel parameters and the level of the evaluated mean square error.

In order to evaluate the influence of the data length on the estimation results, a Monte-Carlo simulation with 50 iterations for $SNR = 10dB$ was performed. The algorithm was applied for different values of N (N varies from 400 to 1200). The Figure 4 show the estimation of the amplitude and phase of the minimum phase channel, using the KAPA algorithm for different numbers of data length and for SNR ($SNR = 10dB$). From these results, we observe that the estimated amplitude and phase patterns, via the kernel algorithm, have the same shape as those of the measured data. On the one hand, we notice that the amplitude and phase estimates are following the real model in perfect accordance to the measured data for a higher data length $N = 1200$, at $SNR = 10dB$. On the other hand, we notice that even if the number of data length is relatively small ($N = 400$, and $N = 800$), the kernel algorithm provides a good estimate of the model with a small difference between the estimated and measured values. These results are very important because we are in the minimum phase channel case. Under troublesome conditions ($N = 400$), the performance of the kernel affine projection algorithm (KAPA) is deteriorated on the phase estimation and

we observe a big difference between the estimated impulse response of the minimum phase channel and the measured phase. Basically, the Gaussian noise has a clear effect on the phase estimation and a minor effect on the amplitude estimation.

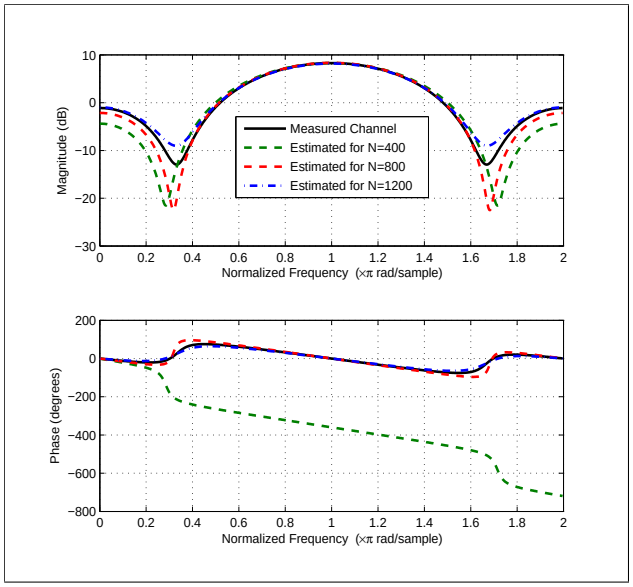


Fig. 4: Estimation of the amplitude and phase of the minimum phase channel, for different data length N and $SNR = 10dB$.

In the third experiment, to observe the noise level impact on the estimation quality, a Monte-Carlo simulation with 50 iterations ($N = 1000$) has been carried out for different SNR values. The algorithm was applied for the following SNR levels: $4dB$, $8dB$, and $12dB$. In this simulation, the frequency domain is considered. The estimation of the amplitude and phase of the minimum phase channel is shown in Figure 5 for different SNR and for a data length $N = 1000$ using the kernel algorithm. From this figure, we remark that the estimated amplitude and phase patterns, in the case of $SNR = 12dB$ have the most common form of the measured data. Compared to the amplitude and phase patterns estimated in the case of $SNR = 4dB$ and $SNR = 8dB$, we notice a small deviation between the estimated and measured amplitude and phase patterns. Figure 5 shows that the kernel algorithm is able to estimate the amplitude and the phase, with good precision, even in a highly noisy environment ($SNR = 4dB$).

Finally, in a fourth simulation experiment, we implement the KAPA algorithm on a nonlinear system under perturbations and with binary output measurements (see Section II) and we compare its performances to those obtained with other algorithms proposed in the literature, namely the least mean squares (LMS) and kernel least mean squares (KLMS) [27] algorithms. To demonstrate the advantage of the KAPA algorithm compared to LMS and KLMS, we test them on an illustrative example of a Monte-Carlo simulation with 50 executions for different SNR.

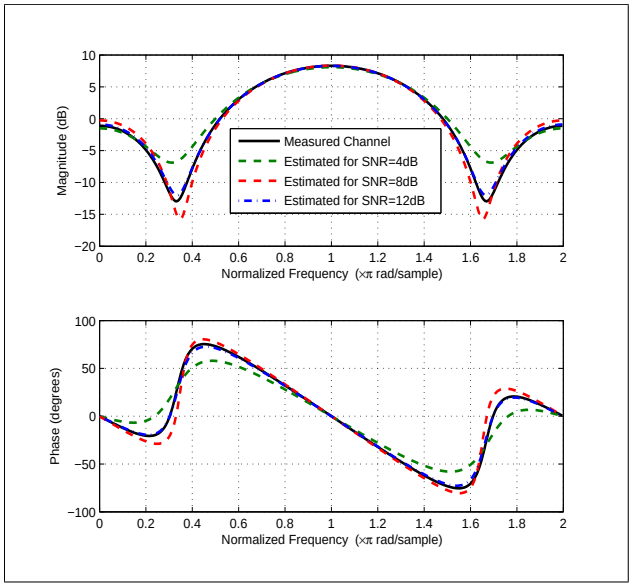


Fig. 5: Estimation of the amplitude and phase of the minimum phase channel, for different SNR and $N = 1000$.

The variations of the mean square error (MSE) as a function of different SNR values are illustrated in Figure 6. As expected, the performance improves with the increase of SNR. From these results, we can observe that the impact of SNR is obvious, which is related to the regularity of the evaluated mean square error. We can clearly observe that the KAPA algorithm achieves the best performance and it is also statistically significant. From Figure 6, we see that the MSE values provided by the KAPA algorithm are very small, compared to those given by other algorithms in the literature (LMS and KLMS), for different SNR and for a data length $N = 1000$, which implies that the estimated parameters are very close to the exact values. For example, when the $SNR = 20dB$, we observe that the MSE value obtained via the KAPA algorithm is 0.015, but using the LMS and KLMS algorithms we have an MSE equal to 0.04 and less than 0.04 respectively.

We discuss the choice of the convergence step value (η). This constant has a significant influence on the convergence speed and stability of the adaptive algorithm. Substituting the correct value (typically a small positive constant) for η is necessary for the proper implementation of the KAPA algorithm:

- If the value of η chosen is too small, the time needed to find the optimal solution by an adaptive filter is too long.
- If the chosen value of η is too high, the adaptive filter becomes unstable and the output yields deviations.

Now we will try to evaluate the influence of the convergence step on the behavior of the three algorithms (LMS, KLMS and KAPA). For this reason, we perform a Monte-Carlo simulation with 50 iterations of data (on the data length $N = 1000$). The MSE evolution as a function of η is shown in Figure

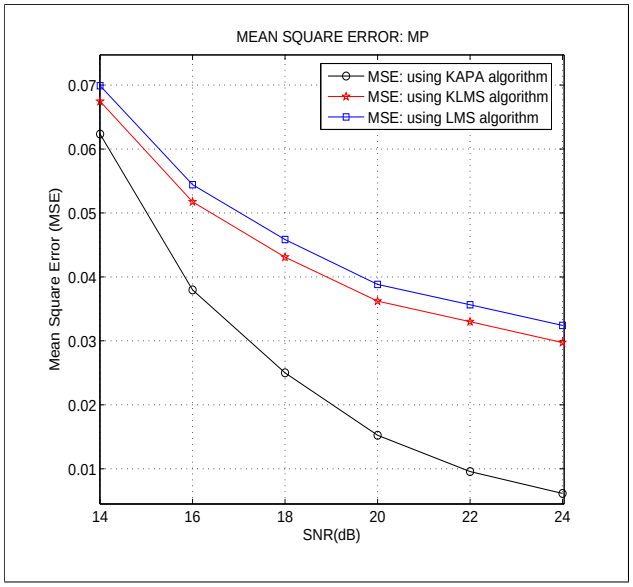


Fig. 6: MSE for different SNR values and for a data length $N = 1000$.

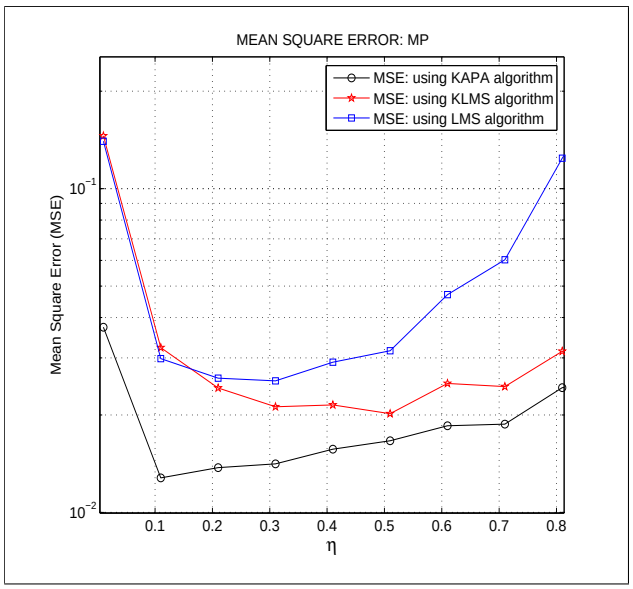


Fig. 7: MSE for different η values and for $N = 1000$ with $SNR = 20dB$.

7. The results demonstrate that, when $\eta = 0.3$, the MSE of three algorithms will increase with an augmentation of η. Nonetheless, the KAPA algorithm is still best than the LMS and KLMS algorithm over any η. The improved performance of the KAPA algorithm, with smallest MSE, is attained by the step size parameter $\eta = 0.1$.

V. Conclusion

In this paper, we have reformulated the nonlinear system identification problem with a binary output as an estimation problem. An estimation algorithm based on positive definite kernels has been presented for the identification of the amplitude and phase of the minimum phase system impulse response. Simulation results show that the KAPA algorithm is appropriate for the identification of systems from binary measurements on the output in the presence of the measurement noise, even if a relatively small data length ($N \leq 800$) was considered. The step size setting η has an important influence on the learn performance of all the three algorithms (LMS, KLMS and KAPA). In fact, for the minimum phase channel impulse response, the affine kernel projection algorithm with $\eta = 0.1$ is more successful in the low noise condition and is more robust compared to other step sizes η.

Future works, will focus on an improved adaptation of this algorithm to more complicated systems.

References

[1] J. Shawe-Taylor, and N. Cristianini, Kernel methods for pattern analysis. Cambridge university press, Cambridge, UK, 2004.

[2] S. Sankar, A. Kar, S. Burra, M. N. S. Swamy, and V. Mladenovic, "Nonlinear acoustic echo cancellation with kernelized adaptive filters", *Applied Acoustics*, vol. 166, pp. 107329, 2020.

[3] L. Ljung, System identification : theory for the user. Upper Saddle River (NJ): Prentice Hall PTR, 1999.

[4] F. Ding, System identification-performances analysis for identification methods. Science Press, Beijing, 2014.

[5] R. Fateh, A. Darif, and S. Safi, "Performance Evaluation of MC-CDMA Systems with Single User Detection Technique using Kernel and Linear Adaptive Method", *Journal of Telecommunications & Information Technology*, no. 4, pp. 1–11, 2021.

[6] N. I. Sapankevych, R. Sankar, "Time series prediction using support vector machines: a survey", *IEEE computational intelligence magazine*, vol. 4, no. 2, pp. 24–38, 2009.

[7] K. Zhang, W. Li, Y. Han, Z. Geng, and C. Chu, "Production capacity identification and analysis using novel multivariate nonlinear regression: Application to resource optimization of industrial processes", *Journal of Cleaner Production*, vol. 282, pp. 124469, 2021.

[8] S. Gupta, A. K. Sahoo, and U. K. Sahoo, "Volterra and Wiener model based temporally and spatio-temporally coupled nonlinear system identification: A synthesized review",*IETE Technical Review*, vol. 38, no. 3, pp. 303–327, 2021.

[9] G. Favier, A. Y. Kibangou, and T. Bouilloc, "Nonlinear system modeling and identification using Volterra-PARAFAC models", *International Journal of Adaptive Control and Signal Processing*, vol. 26, no. 1, pp. 30–53, 2012.

[10] J. Vörös, "Parameter identification of discontinuous Hammerstein systems", *Automatica*, vol. 33, no. 6, pp. 1141–1146, 1997.

[11] I. W. Hunter, and M. J. Korenberg, "The identification of nonlinear biological systems: Wiener and Hammerstein cascade models", *Biological cybernetics*, vol. 55, no. 2, pp. 135–144, 1986.

[12] H. T. Pham, and B. S. Yang, "A hybrid of nonlinear autoregressive model with exogenous input and autoregressive moving average model for long-term machine state forecasting", *Expert Systems with Applications*, vol. 37, no. 4, pp. 3310–3317, 2010.

[13] N. Aronszajn, "Theory of reproducing kernels", *Transactions of the American mathematical society*, vol. 68, no. 3, pp. 337–404, 1950.

[14] W. Liu, J. C. Principe, and S. Haykin, "Kernel adaptive filtering: a comprehensive introduction", Vol. 57. Hoboken, NJ, USA: John Wiley & Sons, 2011.

[15] R. Fateh, and A. Darif, "Mean Square Convergence of Reproducing Kernel for Channel Identification: Application to Bran D Channel Impulse Response", *In International Conference on Business Intelligence*, Springer, Cham, pp. 284–293, 2021.

[16] R. Fateh, A. Darif, and S. Safi, "Kernel and Linear Adaptive Methods for the BRAN Channels Identification", *In : International Conference on Advanced Intelligent Systems for Sustainable Development*, Springer, Cham, pp. 579–591, 2020.

[17] R. Fateh, A. Darif, and S. Safi, "Identification of the Linear Dynamic Parts of Wiener Model Using Kernel and Linear Adaptive", *In : International Conference on Advanced Intelligent Systems for Sustainable Development*, Springer, Cham, pp. 387–400, 2020.

[18] R. Fateh, A. Darif, and S. Safi, "Channel Identification of Non-linear Systems with Binary-Valued Output Observations Based on Positive Definite Kernels", *In E3S Web of Conferences*, vol. 297, EDP Sciences, 2021.

[19] M. Zidane, S. Safi, M. Sabri, and A. Boumezzough, "Impulse response identification of minimum and non minimum phase channels", *In 4th Workshop on Codes, Cryptography and Communication Systems (WCCCS'13)*, pp. 07–08, 2013.

[20] M. Zidane, S. Safi, M. Sabri, and M. Frikel, "Using Least Mean p-Power Algorithm to Correct Channel Distortion in MC-CDMA Systems", *Journal of Telecommunications and Information Technology*, vol. 3, pp. 23–30, 2018.

[21] M. Zidane, S. Safi, M. Sabri, and A. Boumezzough, "Higher order statistics for identification of minimum phase channels", *International Journal of Mathematical, Computational, Physical and Quantum Engineering*, vol. 8, no. 5, pp. 831–836, 2014.

[22] S. Safi, A. Jilali, and M. Frikel, "NonLinear System Identification Using Cumulants", *In International Conference on Signals, Automatic control and Telecommunications (ICSAT'2018)*, May, 2018.

[23] S. Safi, M. Frikel, M. M'Saad, and A. Zeroual, "Blind Identification of Frequency Selective Channel using Higher Order Statistics", *In : 2007 IEEE International Conference on Signal Processing and Communications*. IEEE, pp. 173–176, November, 2007.

[24] S. Safi, and A. Zeroual, "MA system identification using higher order cumulants application to modelling solar radiation", *Journal of Statistical Computation and Simulation*, vol. 72, no. 7, pp. 533–548, 2002.

[25] S. Safi, and A. Zeroual, "Blind identification in noisy environment of nonminimum phase finite impulse response (FIR) system using higher order statistics", *Systems Analysis Modelling Simulation*, vol. 43, no. 5, pp. 671–681, 2003.

[26] J. W. Xu, A. R. Paiva, I. Park, and J. C. Principe, "A reproducing kernel Hilbert space framework for information-theoretic learning", *IEEE Transactions on Signal Processing*, vol. 56, no. 12, pp. 5891–5902, 2008.

[27] W. Liu, P. Pokharel, and J. C. Principe, "The kernel least-mean-square algorithm", *IEEE Transactions on signal processing*, vol. 56, no. 2, pp. 543–554 2008.

[28] Y. Engel, S. Mannor, and R. Meir, "The kernel recursive least-squares algorithm", *IEEE Transactions on signal processing*, vol. 562, no. 8, pp. 2275–2285, 2004.

[29] A. A. Khan, S. M. Shah, M. A. Z. Raja, N. I. Chaudhary, Y. He, and J. A. Machado, "Fractional LMS and NLMS algorithms for line echo cancellation", *Arabian Journal for Science and Engineering*, vol. 46, no. 10, pp. 9385–9398, 2021.

[30] P. Bouboulis, and S. Theodoridis, "The complex Gaussian kernel LMS algorithm", *In International Conference on Artificial Neural Networks*, Springer, Berlin, Heidelberg, pp. 11–20, 2010.

[31] B. Chen, S. Zhao, P. Zhu, and J. C. Príncipe, "Quantized kernel least mean square algorithm", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 1, pp. 22–32, 2011.

[32] S. Zhao, B. Chen, P. Zhu, and J. C. Príncipe, "Fixed budget quantized kernel least-mean-square algorithm", *Signal Processing*, vol. 93, no. 9, pp. 2759–2770, 2013.

[33] A. Flores, and R. C. de Lamare, "Set-membership kernel adaptive algorithms", *In 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2676–2680, Mar. 2017.

[34] W. Liu, I. Park, Y. Wang, J. C. and Principe, "Extended kernel recursive least squares algorithm", *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 3801–3814, 2009.

[35] S. Van Vaerenbergh, J. Via, and I. Santamaría, "A sliding-window kernel RLS algorithm and its application to nonlinear channel identification", *in Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'06)*, vol. 5, pp. 789–792, Toulouse, France, May 2006.

[36] A. Sayed, Fundamentals of adaptive filtering. Wiley-IEEE Press, NY, USA, June 2003.

[37] B. Schölkopf, R. Herbrich, A. J. Smola, "A generalized representer theorem", *International Conference on Computational Learning Theory*, Springer, Berlin, Heidelberg, pp. 416–426, 2001.

[38] W. Liu, J. C. Principe, "Kernel affine projection algorithms", *EURASIP Journal on Advances in Signal Processing*, vol. 2008, no. 1, pp. 1–12, 2008.

Algorithmes de décodage des codes MDPC quasi-cycliques pour le cryptosystème de McEliece

Abdellatif Kichna
Laboratoire LIMATI
Faculté Polydisciplinaire
Université Sultane Moulay Slimane
Béni Mellal, Maroc
abdellatif.kichna@usms.ac.ma

Abderrazak Farchane
Laboratoire LIMATI
Faculté Polydisciplinaire
Université Sultane Moulay Slimane
Béni Mellal, Maroc
a.farchane@gmail.com

***Abstract*—Le développement des ordinateurs quantiques constitue une menace majeure sur les cryptosystèmes de chiffrement modernes basés sur la factorisation des entiers et le logarithme discret. Cette situation a suscité la recherche de bons cryptosystèmes qui permettent d'assurer la sécurité contre les menaces quantiques ainsi que classiques. Certains cryptosystèmes se révèlent résistants à la puissance des ordinateurs quantiques, dont l'un est basé sur les codes correcteurs d'erreurs, introduit en 1978 par McEliece. Une variante de ce cryptosystème est basée sur les codes de contrôle de parité à densité modérée quasi-cyclique (QC-MDPC), caractérisés par leur taille de clé compacte, contrairement au cryptosystème original de Goppa-McEliece. Dans ce travail, les algorithmes de décodage utilisés pour le Cryptosystème McEliece basé sur les codes QC-MDPC et leurs caractéristiques ont été présentés.**

***Index Terms*—Cryptographie post quantique, cryptographie basée sur les codes, codes QC-MDPC, Décodage Bit-Flipping.**

I. Introduction

La cryptographie est la science qui utilise les mathématiques pour chiffrer et déchiffrer des données afin de les protéger. Il est utilisé pour stocker des informations importantes ou pour les transmettre via des canaux de communication non fiables afin qu'elles ne puissent être lues par personne d'autre que le destinataire légitime. Il existe deux principaux types de cryptographie : la cryptographie symétrique qui nécessite que l'expéditeur et le destinataire aient des clés identiques pour chiffrer et déchiffrer les données, tandis que la cryptographie asymétrique, ou cryptographie à clé publique, utilise une clé publique pour chiffrer les messages et une clé privée pour un récepteur pour les déchiffrer.

Les techniques cryptographiques ont été utilisées pour fournir une communication sécurisée entre les utilisateurs depuis l'époque classique. Cependant, les progrès dans le domaine de la physique quantique et les calculateurs quantiques pourraient donner naissance à de puissants ordinateurs quantiques au cours de la prochaine décennie. un ordinateur quantique est fondamentalement différent des ordinateurs traditionnels. Les grands ordinateurs quantiques peuvent logiquement résoudre certains problèmes beaucoup plus rapidement que les ordinateurs classiques en utilisant les meilleurs algorithmes connus à ce jour. Contrairement aux ordinateurs classiques, qui utilisent des bits qui ne peuvent prendre que les valeurs 0 ou 1, les machines quantiques utilisent des qubits qui peuvent représenter simultanément différents états possibles intermédiaires entre 0 et 1, on parle de superposition.

Ce développement quantique ainsi que les résultats de Peter Shor [1] sur l'existence d'algorithmes polynomiaux pour résoudre les problèmes de logarithme discret et de factorisation des nombres sur les ordinateurs quantiques, ont été un stimulant en cryptographie pour rechercher d'autres problèmes de calcul difficiles sur lesquels des cryptosystèmes puissants peuvent être construits. Il s'est avéré que les méthodes de chiffrement à clé publique qui peuvent être considérées comme sûres dans le monde de l'informatique quantique ont été créées à la fin des années 1970 [2], c'est-à-dire avant même que les chercheurs pensent aux menaces possibles des ordinateurs quantiques sur la cryptographie.

L'un des moyens de se défendre contre un ordinateur quantique consiste à modifier les systèmes de chiffrement des données pour passer à la cryptographie quantique, qui utilise des dispositifs quantiques pour transférer les informations. Le problème de cette transition est qu'il est nécessaire de modifier les systèmes de chiffrement au niveau physique. Une autre solution est la cryptographie post-quantique, qui utilise des systèmes et des problèmes classiques, mais si complexes que même un ordinateur quantique ne peut pas les résoudre.

Il s'agit du développement de nouveaux types de systèmes cryptographiques qui peuvent être appliqués avec les ordinateurs classiques d'aujourd'hui, mais qui ne seront pas affectés par les méthodes quantiques de demain. En principe, il existe de multiples domaines prometteurs en cryptographie post-quantique, dans ce travail nous allons nous concentrer sur l'un d'entre eux qui est la cryptographie basée sur les codes correcteurs d'erreurs, qui est basée sur la difficulté de décoder un code en présence d'erreurs aléatoires. En particulier, nous mettons l'accent sur les algorithmes de décodage pour la version de McElice qui utilise des codes QC-MDPC. Ceux-ci constituent une alternative intéressante aux codes Goppa généralement utilisés.

Dans la section suivante, nous présentons la cryptographie basée sur les codes correcteurs d'erreurs et les définitions nécessaires. Dans la troisième section, nous décrivons le

système de chiffrement QC-MDPC McEliece, sa génération de clés, son chiffrement et déchiffrement. Dans la quatrième section, nous présentons les algorithmes de décodage du QC-MDPC McEliece.

II. Cryptographie basée sur les codes

La cryptographie basée sur des codes utilise des codes de correction d'erreurs pour construire une cryptographie à clé publique. Nous donnons un aperçu des définitions pertinentes de la théorie des codes avant de passer au cryptosystèmes basés sur les codes.

A. La théorie des codes

La théorie des codes de détection et de correction d'erreurs traite de la transmission et du stockage fiables des données. Dans la pratique, les supports d'information ne sont pas totalement fiables, dans le sens où le bruit (toute forme d'interférence) provoque fréquemment une distorsion des données (figure 1). Pour faire face à cette situation indésirable mais inévitable, une certaine forme de redondance est intégrée aux données d'origine. Grâce à cette redondance, même si des erreurs sont introduites (jusqu'à un certain degré), l'information originale peut être récupérée, ou du moins la présence d'erreurs peut être détectée. Cependant, ce type de redondance ne corrige pas l'erreur. C'est exactement ce que font les codes de correction d'erreurs, ils ajoutent de la redondance au message original de sorte qu'il est possible pour le récepteur de détecter l'erreur et de la corriger, en récupérant le message original. Ceci est crucial pour certaines applications où il n'est pas possible de renvoyer le message. Le problème crucial à résoudre est alors de savoir comment ajouter cette redondance afin de détecter et de corriger autant d'erreurs que possible de la manière la plus efficace.

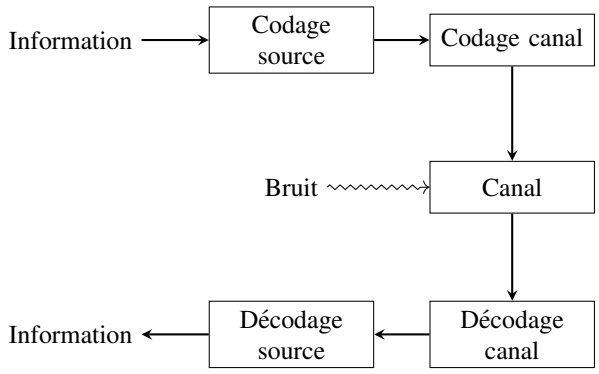


Fig. 1. un modèle de schéma de communication avec bruit

Un mot d'information m est codé par un codeur en un mot de code c. Le mot de code c est transmis sur le canal. Le canal peut transformer le mot codé et introduire des erreurs. En sortie du canal, on obtient le mot reçu y. En sortie du canal, on obtient le mot reçu y. Le décodeur tente alors de décoder y et produit un mot décodé c'. Si $c' = c$, alors nous disons que le décodage a réussi.

Définition 1: (Code Linéaire). Un code linéaire C de longueur n et de dimension k sur $\mathbb{F}_q$ est un sous-espace vectoriel de $\mathbb{F}_q^n$ de dimension k.

Définition 2: (Matrice Génératrice). Soit C un code linéaire (n, k, d) sur $\mathbb{F}_q$. Une matrice génératrice G de C est une matrice $k \times n$ dont les lignes forment une base de C.
Le codage est l'opération qui consiste à multiplier un vecteur m de $\mathbb{F}_q^k$ par la matrice génératrice G d'un code. Le vecteur du code c ainsi construit est appelé mot de code.

$$c = mG$$

On dit qu'une matrice génératrice G est sous forme systématique si elle s'écrit :

$$G = [I_k|P]$$

avec I_k est la matrice identité de taille k.
Définition 3: (Matrice de contrôle de parité). Une matrice de contrôle de parité pour un code linéaire C de paramètres (n, k) est une matrice H de taille $(n - k) \times n$ tel que :

$$cH^T = 0 \text{ si et seulement si } c \in C.$$

Si G est une matrice génératrice d'un code C et H est une matrice de contrôle de parité du même code C, alors :

$$GH^T = 0$$

Définition 4: (Syndrome). Le syndrome d'un mot $y \in \mathbb{F}^n$ par rapport à une matrice de contrôle de parité H est le vecteur:

$$s = yH^T$$

Ainsi, les mots de code d'un code linéaire sont exactement les mots pour lesquels $s = 0$ (indépendamment du choix de la matrice de contrôle de parité pour ce code).
Définition 5: (Poids et distance de Hamming). Le poids du Hamming est le nombre des éléments non nuls d'un mot de code.
La distance de Hamming d entre deux mots de code est le nombre des symboles qui diffèrent.

Définition 6: (code quasi-cyclique). Un code linéaire est dit quasi-cyclique si chaque décalage cyclique par un entier n_0 est également un mot de code.

B. Cryptosystèmes basés sur les codes

Le premier cryptosystème asymétrique basé sur un code correcteur d'erreurs est le cryptosystème de McEliece. Il a été initialement proposé sur la base des codes binaires de Goppa par Robert McEliece en 1978 [2].
Ce cryptosystème utilise la matrice génératrice G du code binaire de Goppa, une matrice de brouillage S et une matrice de permutation P. G est la clé secrète, et la clé publique est obtenue en multipliant les trois matrices SGP.
Contrairement aux systèmes conventionnels de chiffrement à clé publique tels que RSA, le cryptosystème McEliece possède des algorithmes de chiffrement et de déchiffrement de faible complexité.
Les codes Goppa sont considérés comme les meilleurs codes pour le cryptosystème McEliece pour les raisons suivantes :

- l'existence de plusieurs algorithmes de décodage rapide du code en temps polynomial.
- le code Goppa est facile à définir, mais difficile à reconnaître, car la matrice génératrice du code est presque aléatoire ; tout polynôme irréductible sur le champ $GF(2^m)$ convient pour sa compilation
- Pour l'instant, le cryptosystème McEliece basé sur les codes Goppa n'est pas susceptible d'être attaqué.

Néanmoins, ce cryptosystème présente un certain nombre de faiblesses telles qu'une taille de clé énorme et un faible taux de transmission. Mais il reste un candidat pour la cryptographie post-quantique et de nombreuses modifications ont été construites sur sa base. Les premiers pas vers l'objectif de réduire la taille de la clé sans réduire le niveau de sécurité des systèmes cryptographiques post-quantiques ont été donnés par Monico et al. au moyen de codes de contrôle de parité à faible densité (codes LDPC) [3], puis par Gaborit avec des codes quasi-cycliques [4], et enfin par Baldi et Chiaraluce par une combinaison des deux [5]. Cependant, les cryptosystèmes McEliece utilisant des codes LDPC sont susceptibles d'être attaqués sur la base de mots de code de faible poids [6]. Une sous-classe intéressante de codes LDPC est constituée par les codes MDPC (Moderate Density Parity Check) et leur variante quasi cyclique (QC-MDPC) [7]. Ces codes ont des densités suffisamment faibles pour permettre un décodage par des méthodes simples comme l'algorithme de bit flipping de Gallager. Cependant, les densités sont suffisamment élevées pour empêcher les attaques basées sur la présence de mots très épars dans le code dual.

III. Cryptosystème de McEliece basé sur QC-MDPC

Misoczki et al. [7] ont proposé une variante très prometteuse basée sur des codes MDPC quasi cycliques. Dans cette section, nous allons nous focaliser sur leur constuction.

A. Génération de clés

Supposons que nous voulons générer des clés pour le QC-MDPC de McEliece avec un niveau de sécurité λ. Soit n_0, n, r, et w les paramètres supportant ce niveau de sécurité λ (tableau I).
Nous générons d'abord un code QC-MDPC aléatoire en choisissant aléatoirement un vecteur h dans $\mathbb{F}_2^n$, tel que le poids de h soit w.
Nous divisons h en $n_0 = \dfrac{n}{r}$ parties égales:

$$h = [\ h_0 \mid h_1 \mid ... \mid h_{n_0-1}\ ]$$

Ensuite, nous construisons la matrice de contrôle de parité H comme suit :

$$H = [\ H_0 \mid H_1 \mid ... \mid H_{n_0-1}\ ]$$

où chaque H_i est la matrice cyclique avec h_i comme première ligne. Il est nécessaire que le bloc H_{n_0-1} soit inversible. S'il ne l'est pas, nous devons recommencer la procédure de génération de clé en choisissant un autre h au hasard. H est la clé privée. La clé publique G sous sa forme systématique est obtenue à partir de H comme suit : $G = (I_r \mid Q)$, où:

$$Q = \begin{pmatrix} (H_{n_0-1}^{-1}\ H_0) \\ (H_{n_0-1}^{-1}\ H_1) \\ \vdots \\ (H_{n_0-1}^{-1}\ H_{n_0-2}) \end{pmatrix}$$

TABLE I
PARAMÈTRES DU MCELIECE QC-MDPC RECOMMANDÉS PAR [7]

Niveau de sécurité λ	n_0	n	r	w	t
80	2	9602	4801	90	84
80	3	10779	3593	153	53
80	4	12316	3079	220	42
128	2	19714	9857	142	134
128	3	22299	7433	243	85
128	4	27212	6803	340	68
256	2	65542	32771	274	264
256	3	67593	22531	465	167
256	4	81932	20483	644	137

B. Chiffrement

Afin de chiffrer un message $m \in \mathbb{F}_q^k$, on génère une erreur aléatoire $e \in \mathbb{F}_q^n$ de poids $\leq t$. Le texte chiffré est le vecteur.

$$c = mG + e$$

C. Déchiffrement

Pour obtenir le texte en clair m à partir du texte chiffré reçu $c \in \mathbb{F}_2^n$, un algorithme de décodage est nécessaire. Gallager [8] a proposé l'algorithme de bit flipping pour décoder les codes LDPC. Il est basé sur l'idée que chaque bit du syndrome indique si l'équation correspondante est satisfaite ou non. Ainsi, si une position est impliquée dans de nombreuses équations non satisfaites, il s'agit très probablement d'une erreur. Le même algorithme peut être utilisé pour les codes MDPC quasi-cyclique.

Contrairement au cryptosystème McEliece standard, la variante QC-MDPC n'utilise pas la matrice de brouillage S et la matrice de permutation P, parce que la récupération de la clé privée H à partir de la clé publique G est suffisamment difficile. Notez qu'en raison de la forme systématique de G, le texte en clair apparaît dans les k premiers bits. texte en clair apparaît dans les k premiers bits du texte chiffré y avec relativement peu de bits inversés.

IV. Algorithmes de décodage du QC-MDPC

Gallager a initialement présenté les codes LDPC [8] comme ayant des matrices de contrôle de parité régulières, pour lesquelles le poids des lignes et le poids des colonnes sont constants. Avec des matrices de contrôle de parité régulières, chaque bit participe au même nombre de contrôles de parité, ce qui simplifie la description des décodeurs. Chaque ligne d'une matrice de contrôle de parité peut être considérée comme une équation de contrôle de parité. Les positions non nulles d'une ligne indiquent quels bits du mot reçu participent dans

cette équation. Les positions non nulles d'une colonne de la matrice de contrôle de parité indiquent dans quels équations de contrôle de parité participe un bit. Le syndrome d'un mot reçu indique quels équations ne sont pas satisfaites. Gallager a présenté aussi un décodeur itératif pour les codes LDPC. Cet algorithme de décodage est généralement appelé algorithme de bit flipping. L'algorithme fonctionne comme suit: il calcule le nombre d'équations de contrôle de parité non satisfaites pour chaque bit du mot reçu, puis il inverse les bits dont la valeur dépasse un certain seuil. Ce processus est itéré jusqu'à ce que le mot soit décodé.

Dans [7], afin de décoder leur code MDPC, les auteurs ont utilisé l'algorithme original bit flipping de Gallager [8] qui a été introduit pour la première fois pour les codes LDPC. L'algorithme 1 commence par calculer le syndrome du mot de code reçu. Il vérifie ensuite le nombre d'équations de contrôle de parité non satisfaites associées à chaque bit du mot de code, et inverse chaque bit qui participe dans plus de b équations non satisfaites. Ces étapes sont répétées jusqu'à ce que le syndrome devienne nul ou que le nombre maximal d'itérations soit atteint. Dans ce cas, on dit qu'une erreur de décodage s'est produite.

La différence entre l'algorithme original bit flipping de Gallagher et celui proposé par les auteurs est la manière dont le seuil est calculé dans chaque algorithme. Dans le travail original, une séquence de seuils b_i est calculée au préalable pour toutes les itérations. Dans [7], les auteurs ont utilisé un seuil différent recalculé dans chaque itération, ils l'ont pris comme étant le nombre maximum d'équations non satisfaites moins un certain $\delta \geq 0$, $seuil = Max_{upc} - \delta$. Cela peut augmenter la vitesse de convergence des algorithmes, mais peut également provoquer l'inversion d'un plus grand nombre de bits corrects, ce qui peut entraîner un échec du décodage.

Algorithm 1 Algorithme de décodage Bit flipping

Require: $H \in \mathbb{F}_2^{r \times n}$ (matrice de contrôle de parité), $y \in \mathbb{F}_2^n$ (mot reçu)
Ensure: $y \in \mathbb{F}_2^n$
for $iter = 1 \dots MaxIter$ **do**
  $s = yH^T$
  if $s = 0^r$ **then**
    break
  end if
  $th = seuil(contexte)$
  for $i = 1 \dots n$ **do**
    $upc[i] =< s, h_i >$
    if $upc[i] \geq th$ **then**
      $y[i] = 1 - y[i]$
    end if
  end for
end for
return y

Le premier tour de l'implémentation supplémentaire de la soumission BIKE [9] comprenait un décodeur appelé Black-Gray, Il fonctionne de manière similaire à l'algorithme de bit flipping de Gallager, mais crée deux ensembles supplémentaires lors de chaque itération, un ensemble noir contenant les bits inversés, et un ensemble gris contenant les bits légèrement en dessous du seuil. Après le bit flipping initial, les bits de l'ensemble noir dont plus de la moitié des équations de contrôle de parité ne sont pas satisfaites sont inversés à nouveau. Ensuite, les bits de l'ensemble gris, dont plus de la moitié des équations de contrôle de parité sont insatisfaites, sont inversés. Les deux ensembles noir et gris aident à corriger les bits qui ont été inversés par erreur et donne confiance dans l'exactitude des inversements effectués.

Dans [10], les auteurs ont observé le comportement du décodeur Black Gray et ont constaté que lorsque $wt(e)/n$ est suffisamment petit, la probabilité d'inverser des bits par erreur devient faible, ce qui élimine la nécessité de l'approche Black-Gray, ce qui a conduit les auteurs à proposer de nouvelles variantes de BG : Un décodeur Black dans lequel chaque itération consiste uniquement en l'étape du bit flipping et du masque noir. Un décodeur Black-Gray-Flip qui commence par une itération Black-Gray et le reste des itérations est le même que le décodeur bit flipping. Un décodeur Black-Gray-Black, qui commence par une itération de Black-Gray et plusieurs itérations de Black. Les résultats ont montré qu'au niveau de performance, le décodeur Black-Gray-Flip atteint un meilleur taux de décodage d'erreurs par rapport aux autres variantes.

Le travail [11] examine deux autres variantes de décodeurs bit flipping. La première est l'algorithme de bit flipping en parallèle, dans lequel un seuil spécifique pour l'inversion des bits est d'abord calculé, puis les bits à tous les endroits pertinents sont inversés en parallèle. Le second adopte une approche progressive, dans laquelle le seuil est recalculé à chaque fois qu'un bit est inversé.

les auteurs de BIKE dans le 2ème round du compétition de NIST [12] ont défini une variante de l'algorithme bit flipping appelée Backflip. Cet algorithme attribue à chaque bit inversé une valeur ttl : time-to-live, un temps de vie. Après l'expiration de ce temps, le bit est inversé de nouveau. Les bits ayant un nombre plus élevé d'équations de contôles de parité non satisfaites se voient attribuer des valeurs de durée de vie plus élevées pour indiquer qu'ils sont plus fiables. Cette méthode permet d'enregistrer des informations sur la fiabilité des bits.

V. Conclusion

Dans cet article, nous avons présenté une vue d'ensemble sur la variante de McEliece basé sur les codes MDPC quasi-cycliques, ainsi que leurs algorithmes de décodage. Diverses tentatives d'amélioration du décodage bit flipping ont été développées au cours des dernières années. Nous avons constaté que l'utilisation de différents algorithmes lors du décodage des codes qc-mdpc permet d'obtenir une augmentation significative de la probabilité de déchiffrer un message avec succès. Dans les travaux futurs, on va se concentrer sur l'amélioration des performances du décodeurs des codes MDPC quasi-cycliques, en trouvant une combinaison de décodeurs permettant de minimiser le taux d'échec

du décodage ainsi que le nombre d'itérations, et aussi en introduisant une nouvelle méthode pour le choix du seuil , et en les comparant aux décodeurs existants.

REFERENCES

[1] P. W. Shor, "Algorithms for quantum computation: discrete logarithms and factoring," Proceedings 35th Annual Symposium on Foundations of Computer Science, 1994, pp. 124-134, doi: 10.1109/SFCS.1994.365700.

[2] R. J. McEliece, A Public-Key Cryptosystem Based On Algebraic Coding Theory. Deep Space Network Progress Report, vol. 44, pp. 114–116 (1978)

[3] C. Monico, J. Rosenthal, A. Shokrollahi, Using low density parity check codes in the McEliece cryptosystem. 215. 10.1109/ISIT.2000.866513. (2000)

[4] P. Gaborit, Shorter keys for code-based cryptography. (2005).

[5] M. Baldi, and F. Chiaraluce, Cryptanalysis of a new instance of McEliece cryptosystem based on QC-LDPC codes. IEEE International Symposium on Information Theory, pp. 2591-2595 (2007)

[6] T. Fabšič, V. Hromada, P. Stankovski, P. Zajac, Q. Guo, TA. Johansson, Reaction Attack on the QC-LDPC McEliece Cryptosystem. In: Advanced Data Mining and Applications. Advanced Data Mining and Applications, pp 51–68 (2017)

[7] R. Misoczki, J. Tillich, N. Sendrier, P. Barreto, MDPC-McEliece: New McEliece variants from Moderate Density Parity-Check codes. IEEE International Symposium on Information Theory, 2069-2073 (2013)

[8] R. G. Gallager, Low Density Parity Check Codes in , Cambridge, MA:M.I.T. Press (1963)

[9] N. Ducker, S. Gueron, Additional implementation of BIKE: Bit Flipping Key Encapsulation. https://bikesuite.org/additional.html (2017)

[10] N. Drucker, S. Gueron, D. Kostic, QC-MDPC decoders with several shades of gray. IACR Cryptol. ePrint Arch. (2019)

[11] N. Sendrier, V. Vasseur, "On the Decoding Failure Rate of QC-MDPC Bit-Flipping Decoders" In: Ding J., Steinwandt R. (eds) Post-Quantum Cryptography. PQCrypto 2019. Lecture Notes in Computer Science, vol 11505. Springer, Cham. 2019

[12] N. Aragon, P. S. L. M. Barreto, S. Bettaieb, L. Bidoux, O. Blazy, J.-C. Deneuville, P. Gaborit, S. Gueron, T. Güneysu, C. A. Melchor, R. Misoczki, E. Persichetti, N. Sendrier, J.-P. Tillich, V. Vasseur, and G. Zémor. "BIKE: Bit Flipping Key Encapsulation, Round 2 Submission." 3.0. 2019

DOA Estimation Using Sparse Representation and Hight-resolution Methods in Various Noise Levels Environnements

A. Ghandouri[1], S. Safi[1], M. Frikel[2]
[1]*LIMATI Laboratory, Sultan Moulay Slimane University, Beni Mellal, Morocco*
[2]*LIS laboratory, ENSICAEN, Cean, France*

***Abstract*—The objective of this paper is to present a source separation method using a sparse representation of the signals, and to show that its performance is better than that obtained by high-resolution methods in different noise level environments, even under a closely spaced signal. Actually, so-called high-resolution methods are based on the estimation of the covariance matrix of the sensor signals. Thus, a sparse representation does not use this covariance matrix to resolve the DOA. In addition, a higher signal-to-noise ratio (SNR) scenarios displays a better resolution. In that reason, our simulations results verify the performance of each method by applying the algorithms in different noise level environments and prove the efficiency of the sparse representation.**

***Index Terms*—Array processing, direction-of-arrival (DOA) estimation, angle-of-arrival (AOA), sparse representation.**

I. Introduction

In the last decade, the problem of sparse representations has become particularly active and its applications have spread to many fields. This problem has brought new challenges in direction-of-arrival (DOA) estimation algorithm problems using the angle-of-arrival (AOA) technique. Over the past thirty years, a great deal of work has led to the development of high-resolution methods in the framework of uniformly distributed antenna arrays and for narrowband type sources [1], [2]. These are themes that give rise to many industrial applications: Geolocation [3], Radiocommunication [4], Location systems [5], etc. Also, the applications of the sparse representation of received signals by the acoustic antenna are widespread in many fields, such as blind signal separation [6], [7].

To apply the technique of sparse signal analysis to sound source localization, we will use "ill-posed" inverse linear problems [8] in which a priori information about the unknown coefficients can be incorporated. A redundant representation is considered in order to free the directional vector matrix from its dependence on the source positions. This redundant representation plays an important role in ensuring signal sparsity [9]. To reinforce the signal sparsity, an essential point of this algorithm is to use a non-quadratic term (l_1-norm in this case). A simplified method based on the Singular Value Decomposition (SVD) of the received signals by the acoustic antenna is also proposed under the name l_1-norm-based Singular Value Decomposition (l_1-SVD) algorithm [10].

This work focuses more specifically on the problems encountered in antenna processing and source localization. More precisely, we are interested in the estimation of the DOA. In addition, high-resolution methods are introduced such as beamforming, minimum variance distortionless response (MVDR), The Multiple Signal Classification (MUSIC) and Minimum norm (MIN-NORM). They are based on the estimation of the covariance matrix of the sensor signals. However, these methods require the use of a propagation model and are not very robust to model errors, which can degrade their performance.

This paper begins by presenting, in Section II, a model for the received signals by an antenna array for the case where the sources are narrowband [11], [12] and corrupted by a uniform Additive White Gaussian Noise (AWGN), impinging on a Uniform Linear Array (ULA). Subsequently, in Section III, methods that are related to the estimation of the DOA of sound sources are presented, we are interested in methods whose spatial resolution is higher than that of conventional antenna processing: MVDR, MUSIC, minimum norm and a sparse representation for estimating DOAs of narrowband sound sources. The simulations of the methods presented in this paper were then performed in Section IV. Finally, in Section V, conclusions will be drawn.

II. Signal modeling

A. Signal representation

Before presenting the data model, authors consider the same assumptions taken in [13], as shown in figure (Fig. 1). This antenna receives P ($P < N$) (P and N are the number of sources andthe number of sensors, respectively) signals from radiating sources with additive noise super imposed. It is assumed that the emitted signals by these sound sources are stationary, centered and uncorrelated with the additive noise. The P sources are placed in a far-field, thus assumed to be point-like, and the additive noise observed on the different sensors is stationary, centered and of variance σ^2.

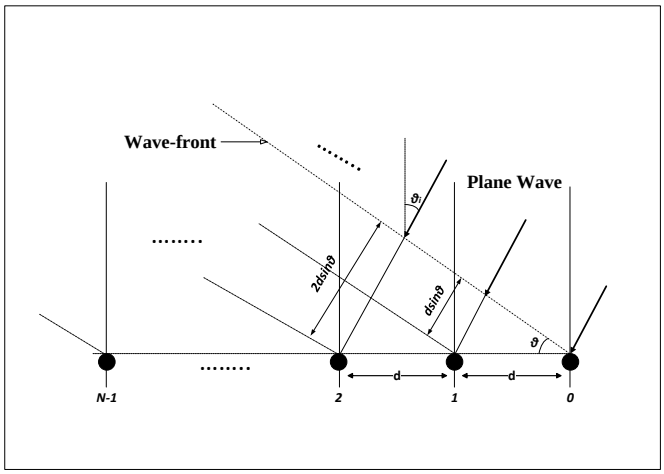


Fig. 1: Wave arrival schematic at ULA antenna.

The received signal on the n^{th} sensor at time t is:

$$x_n(t) = \sum_{i=1}^{P} s_i(t)a(\theta_i)+b_n(t), \quad n = \{0, 1, \cdots, N-1\}, \quad (1)$$

where $s_i(t)$ is complex amplitude of the i^{th} source, $a(\theta_i)$ is the directional vector or transfer vector for the i^{th} source and $b_n(t)$ is the additive noise on the n^{th} sensor. It is assumed to be white, Gaussian and with variance σ^2. We can also simplify the writing of the equation (1) by using matrices. Thus, if we pose:

$$A(\theta) = (a(\theta_1) \;\; a(\theta_2) \;\; \cdots \;\; a(\theta_p)), \quad (2)$$

$$s(t) = [s_1(t) \;\; s_2(t) \;\; \cdots \;\; s_p(t)]^T, \quad (3)$$

$$b(t) = [b_0(t) \;\; b_1(t) \;\; \cdots \;\; b_{N-1}(t)]^T, \quad (4)$$

where $A(\theta)$ is the matrix of directional vectors of sources of dimension $(N \times P)$, it contains the information on the angles of arrival, $s(t)$ is the vector of the complex amplitudes of the signals emitted by the P sources at time t and $b(t)$ is the additive noise vector on the N sensors. The equation (1) becomes [14]:

$$x(t) = A(\theta)s(t) + b(t). \quad (5)$$

B. Covariance matrix

The basic principle of DOA estimation methods is to determine the mathematical expectation between the received signals in the form of a covariance matrix R for an observation is given by:

$$R = E[x(t)x^H(t)], \quad (6)$$

$$= A(\theta) \underbrace{E[s(t)s^H(t)]}_{R_s} A^H(\theta) + \underbrace{E[b(t)b^H(t)]}_{R_B}, \quad (7)$$

$$= A(\theta)R_sA^H(\theta) + \sigma^2 I, \quad (8)$$

where $\sigma^2 I$ is the noise covariance matrix with σ^2 is the power of the noise for each element of the array and I is the identity matrix, the term H corresponds to the transpose-conjugate combination of the matrix. In practice, the covariance matrix is estimated from a finite number of time samples as:

$$R \simeq \widehat{R}_x = \frac{1}{T}\sum_{k=1}^{T}(x(k)x^H(k)), \quad (9)$$

where $x(k)$ is the signal vector sampled at time k and T is the number of samples.

III. Methods

A. Conventional beamforming

Source localization is one of the main tasks in antenna processing and the simplest method is conventional beamforming. A beamformer is a spatial filter that processes the data obtained from a sensor array in order to increase the amplitude of the signals relative to the background noise and/or interference. The principle of this classical algorithm is to dephase, weight and sum the signals at the output of each sensor.

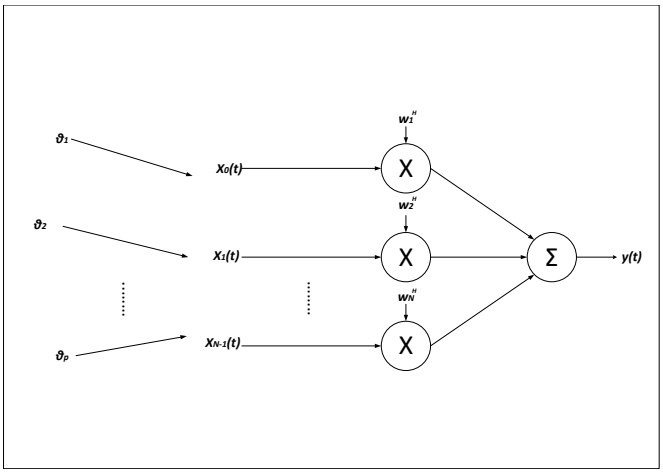


Fig. 2: Illustration of spatial filtering. Each output of the array is multiplied by a coefficient of the filter.

The output signal of an N sensor array, followed by a spatial filter, is expressed as:

$$y(t) = w^H x(t), \quad (10)$$

where $w = [w_1 w_2 \cdots w_N]^T$ is the beamforming weighting vector; it is a vector of dimension $(N \times 1)$ that acts as a spatial filter. The power at the output of the antenna is:

$$P_u = E[y(t)y(t)^*] = E[|y(t)|^2] = w^H Rw. \quad (11)$$

B. Capon's method (MVDR)

In order to cope with the limitations of the conventional beamformer presented above, which makes the interpretation of the data is difficult, in addition to the nearby sources are not discernible, such as its resolving power when locating multiple sources, Capon proposed a method [15], in 1969 and it is sometimes given the name of its author. This method is based on spatial filtering to focus the direction of the main lobe of the array while cancelling the power in other directions, this technique is also called Minimum Variance

Distortion Response (MVDR) [10]. This method is based on minimization problem modeled by the following equation:

$$\min_{w} P_u = w^H R w \quad subject\ to\ w^H a(\theta = \theta_i) = 1. \tag{12}$$

Thus, we can obtain the power at the output of the antenna by the following relation:

$$P_{MVDR}(\theta) = \frac{1}{a^H(\theta)\widehat{R}^{-1}a(\theta)}. \tag{13}$$

C. MUSIC method

The Multiple Signal Classification (MUSIC) method is an example of this type of method that was first proposed by Schimidt R.O. in 1986 [16] to estimate the DOA of signals. It assumes that the theoretical covariance matrix has the form given in the relations (6), (7) and (8) (with spatial white noise) and that the covariance matrix of the signals from the R_s sources have full rank P. It is assumed that ($N > P$) (N and P are the number of sensors and the number of sources, respectively) and that the directional vectors $\{a(\theta_1) \quad a(\theta_2) \ \cdots \ a(\theta_p)\}$ of the matrix (6) are linearly independent. The source transfer matrix A is of full rank P, which is:

$$rank(A(\theta)R_s A^H(\theta)) = P. \tag{14}$$

According to the relation (14), $A(\theta) R_s A^H(\theta)$ has P strictly positive eigenvalues, and the other $(N-P)$ eigenvalues are all equal to zero. The spectral decomposition of the covariance matrix R into eigenelements to separate the signal subspace from the noise subspace can be expressed in the following form:

$$R = AR_sA^H + \sigma^2 I = U\Lambda U^H = U_s\Lambda_s U_s^H + \sigma^2 U_b U_b^H, \tag{15}$$

and,

$$U = [U_s \ U_b], \tag{16}$$

where Λ is the diagonal matrix of the eigenvalues of the matrix R, U is the matrix formed by the eigenvectors corresponding to the eigenvalues of the matrix R (ordered by decreasing), $U_s = [u_1 \ u_2 \ \cdots \ u_P]$ and $U_b = [u_{P+1} \ u_{P+2} \ \cdots \ u_N]$ represent respectively the eigenvalue matrices associated with the signal subspace and the noise subspace. The angular spectral function obtained by the MUSIC method allows us to determine the values of θ for which this function is maximal and it is defined as follows:

$$P_{MUSIC}(\theta) = \frac{1}{a^H(\theta)U_b U_b^H a(\theta)}. \tag{17}$$

This relation (17) can also be expressed as the quadratic norm of the projections of the signal subspace onto the noise subspace:

$$P_{MUSIC}(\theta) = \frac{1}{\sum_{i=P+1}^{N} |a^H(\theta)u_i|^2}, \tag{18}$$

where u_i is the i^{th} vector of the matrix U_b from the noise subspace with $i = \{P+1, \ \cdots, \ N\}$. Note that $P_{MUSIC}(\theta)$ is not a true spectrum (it is a measure of the distance between two subspaces), it gives us peaks corresponding to the exact DOA of the waves but does not inform us about the power of the sources.

D. Minimum norm method

Unlike the MUSIC method where the directional vectors are projected onto the whole $(N \times P)$ dimensional noise subspace, the minimum norm method [17] projects the directional vectors onto a particular one-dimensional vector v which is a combination of the noise eigenvectors $[u_{P+1}, \ u_{P+2}, \ \cdots, \ u_N]$. The angular spectrum is expressed in the following form:

$$P_{MIN-NORM}(\theta) = \frac{1}{|a^H v|^2}. \tag{19}$$

Since v lies in the noise subspace, it satisfies the following relation:

$$a^H(\theta_k)v = 0, \quad k = \{1, \ 2, \ \cdots, \ P\}, \tag{20}$$

where $v = [v_1, \ v_2, \ \cdots, \ v_N]^T$. The determination of v is therefore a minimization problem. The formulation of the optimization problem is given using the first column of the identity matrix e_1 as follows:

$$\min_{v} (v^H V_b v) \quad subject\ to \quad e_1^H v = 1, \tag{21}$$

where $e_1 = [1, \ 0, \ \cdots, \ 0]^T$. Similar to the MUSIC method, the angular spectrum of the minimum norm method is expressed as follows [18]:

$$P_{MIN-NORM}(\theta) = \frac{1}{a^H(\theta)v_{m-n}v_{m-n}^H a(\theta)}. \tag{22}$$

E. Sparse representation

Recall that the objective is to find the directions of the waves emitted by the point sources. A model was established in Section II, and we consider an acoustic antenna formed by N sensors and placed in a medium comprising P point sources, the signal observed at the output of the sensors is expressed in a general way:

$$y(t) = A(\theta)\, s(t) + b(t), \quad t = \{1, 2, \cdots, T\}. \tag{23}$$

The parametric estimation problem indicated by the relation (23) is transformed into a sparse source representation problem. For this purpose, we introduce a redundant representation of A. This one serves as a redundant dictionary in terms of all possible source positions $\{\widetilde{\theta_1}, \widetilde{\theta_2}, \ \cdots, \widetilde{\theta_L}\}$, which leads to a grid, or a spatial sampling, including the maximum of potential positions of the sources to be localized. This number of source positions L is generally much larger than the number of real sources P, and even much larger than the number of sensors N. The matrix A of size $(N \times L)$ is then constructed in the following form:

$$A\left(\widetilde{\theta}\right) = \left[a\left(\widetilde{\theta_1}\right), a\left(\widetilde{\theta_2}\right), \ \cdots, a\left(\widetilde{\theta_L}\right)\right]. \tag{24}$$

Note that here the matrix $A\left(\widetilde{\theta}\right)$ is formed by the directional vectors corresponding to all the possible positions of the potential sources, which is different from the matrix of

directional vectors of the relation (23) in the sense that this matrix contains only the directional vectors of the positions of the real (or exact) sources. In this representation the transfer matrix is known and does not depend on the exact position of the sources which is provided by the set $\{\theta_1, \theta_2, \cdots, \theta_P\}$, as see in (23). On the other hand, the complex amplitudes of the signals reaching the antenna at time t can be expressed by a vector of dimension $(L \times 1)$:

$$x(t) = [x_1(t), x_2(t), \cdots, x_L(t)]^T, \quad (25)$$

where the p^{th} element $x_p(t)$ is non-zero and equal to $s_p(t)$ if $x_p(t)$ and $s_p(t)$ correspond to the same source direction while all other elements are zero. We have the representation:

$$y(t) = A\left(\widetilde{\theta}\right) x(t) + b(t), \quad t = \{1, 2, \cdots, T\}, \quad (26)$$

where T is the number of time samples. In fact, the observation vector of the signal $y(t)$ appearing in (23) or in (26) is identical. This means that $y(t)$ expressed using (26) can also be used by traditional antenna processing methods to locate sources. However, a form of redundant representation appears in (26) allowing us to transform the problem of parametric estimation of the position of the sources into a problem of spatial sparse estimation of these sources. Similar to many non-parametric source location algorithms, this algorithm seeks to locate sources by estimating the energy of the signal whose spectral representation has dominant peaks corresponding to the true source positions, in the ideal case. We need to obtain an estimate of the signal $x(t)$ via the observation vector $y(t)$ by solving (26) which is an "ill-posed" problem. The key assumption is that the number of sources is low, which ensures the sparsity. Under such assumption, the spatial spectrum is sparse, i.e., the majority of the elements of $x(t)$ are zero (or near zero).

We have in fact the problem of minimizing the objective functional $J(x)$, involving a l_2-norm attachment term to the data and a l_1-norm penalty on the coefficients. The appropriate object function for the problem takes the following form:

$$J(x) = \left|\left|y(t) - A\left(\widetilde{\theta}\right) x(t)\right|\right|_2^2 + \lambda ||x(t)||_1. \quad (27)$$

In the real-valued case, the formulation (27) corresponds to a convex quadratic problem that can be solved by various algorithms. But the data in our models are complex-valued and these two formulations are no longer a pure quadratic problem. The problem can however be formulated as a Second Order Cone Programming (SOCP) which is efficient for optimizing functions with complex, quadratic convex and linear terms and allows to approach the optimization more easily [19]. The general form of an SOCP is expressed as a minimization with respect to $x(t)$:

$$min\ c^T x \ \ such\ as \ \ Ax = b \ \ and \ \ x \in K, \quad (28)$$

where K is a quadratic cone, i.e., such that $\{(x_1, x_2) \in R \times R^{L-1}\}$ with $||x_2||_2 \leq 1\}$. For a SOCP problem, K can be expressed as: $K = \{R_+^L \times Q_1 \times \cdots \times Q_L\}$, where R_+^L is the positive orthogonal cone of dimension L and $\{Q_1, \cdots, Q_L\}$ are the second order cones also called Lorentz cones. Figure 3 illustrates a second order cone in R^3 space:

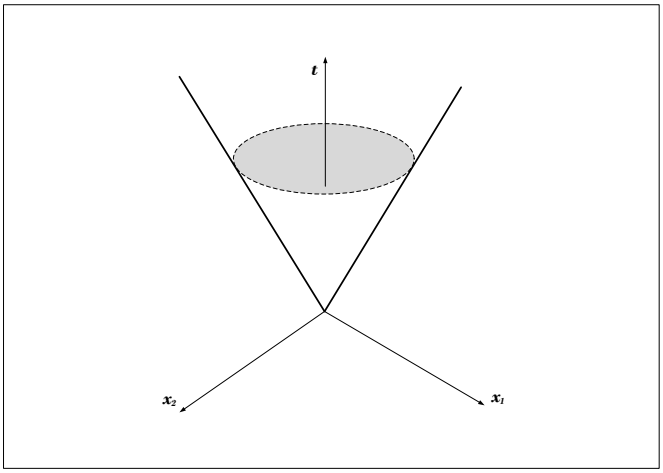


Fig. 3: Illustration of the second order cone (Lorentz) in R^3.

In this paper, the optimization problem will be solved using an optimization software programmed in MATLAB under the name of Self-Dual Minimization (Se-DuMi) [20] developed by J. Sturm [21].

F. l_1-SVD approach

In the relation (23) in order to find the sparse representation of the signal with the multiple time samples, a direct way is to solve separately T problems from $t = 1$ to $t = T$ and to obtain a set of T solutions $\hat{x}(t)$ with $t = \{1, 2, \cdots, T\}$. The main disadvantage of treating each time sample separately is that there is no relation between these subproblems for different time samples. For example, if we change the value of $x(t_1)$ by another random value, it does not induce a direct influence on $x(t_2)$. The consequence is that this approach suffers from sensitivity to the signal-to-noise ratio (SNR).

We thus, consider another approach that synthesizes the multiple temporal samples before seeking the sparse representation of the signal, we enforce the signal's sparsity only spatially but not temporally. Specifically, we first perform a l_2-norm on each spatial element of $x(t)$ that includes the multiple temporal samples and then perform a l_1-norm on this new value of x to find the spatially sparse representation. The main weakness of this approach is its huge computational cost. The size of this inverse problem increases linearly with the number of samples T and the computational cost increases "super-linearly" with T. When T is large, this approach is not reliable to solve the source location problem. In order to reduce the computational time, we now present an extremely efficient approach for problems with a large number of temporal samples, known as l_1-SVD, which was initiated by Malioutov and Alan S Willsky [22].

Let Y be the synthesized observation matrix of size $(N \times T)$ of the temporal data: $Y = [y(t_1), y(t_2), \cdots, y(t_T)]$. We define in a similar way X and B and we obtain: $Y = A\left(\widetilde{\theta}\right) X + B$.

To reduce the computation time and the sensitivity to noise, we perform a singular value decomposition (SVD) of the matrix Y and we deduce the reduced observation matrix Y_{SVD}. The principle of this l_1-SVD approach is to decompose the matrix Y into signal and noise subspaces by retaining only, the signal subspace, the problem size is considerably reduced compared to the approach presented above. Note that unlike the subspace methods, such as MUSIC and minimum norm, here it is the signal subspace that is retained and not the noise subspace.

In the ideal case of the signal without noise, the set of vectors $y(t_i)$ with $i = \{1, 2, \cdots, T\}$ is in a subspace of dimension P, where P is the number of sources. We only need one basis of this subspace to realize the sparse representation of the signals. Mathematically, we perform a singular value decomposition (SVD):

$$Y = WLV'. \tag{29}$$

We obtain a reduced observation matrix Y_{SVD} of dimension $(N \times P)$, which contains the major power by performing the following decomposition:

$$Y_{SVD} = WLD_P = YVD_P, \tag{30}$$

where $D_P = [I_P \ O]'$. Here, I_P is a unit matrix of dimension $(P \times P)$ and O represents a null matrix of dimension $P \times (T - P)$. In the same way, we have $X_{SVD} = XVD_P$ and $B_{SVD} = BVD_P$. We obtain:

$$Y_{SVD} = A\left(\widetilde{\theta}\right) \ X_{SVD} + B_{SVD}. \tag{31}$$

Considering the relation (31) column by column (each column corresponds to a singular vector of the signal subspace):

$$y_{SVD}(p) = A\left(\widetilde{\theta}\right) x_{SVD}(p) + b_{SVD}(p), \quad p = \{1, 2, \cdots, P\}. \tag{32}$$

Comparing (26) with (32) we can notice that the number of equations is reduced from T to P. Generally, the number of sources P is much smaller than the number of time samples T, so this reduction on the computational cost is very important.

IV. Simulations

This section focuses on testing the resolution capability of each algorithm mentioned in section III. We consider a uniform linear array formed by 20 sensors separated by half a wavelength of the actual narrowband source signals. Two zero-mean narrowband signals in the far-field impinge upon this array from distinct DOA. The simulation is done by taking $N = 3$ sources and the total number of temporal samples is T = 1000.

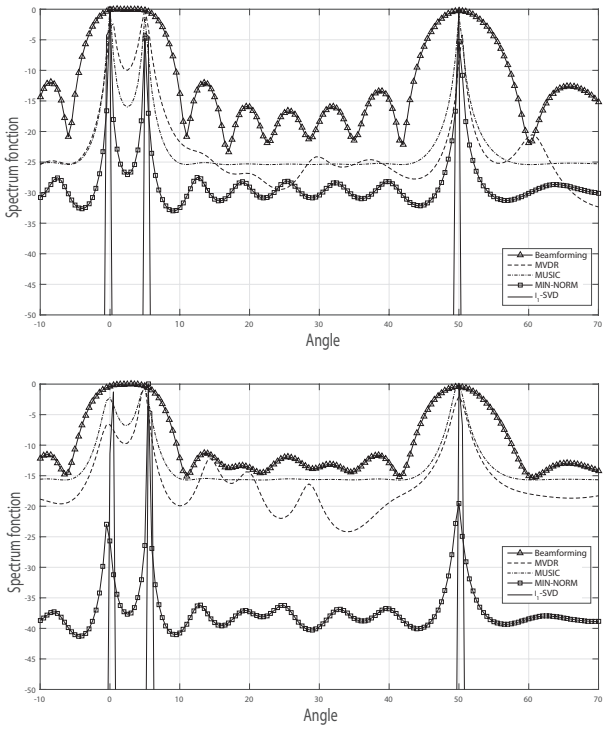


Fig. 4: Spectra for uncorrelated sources, DOAs: $0°$, $5°$ and $50°$. Top: SNR = 10 dB. Bottom: SNR = 0 dB, using the presented method.

In figure (Fig. 4), we compare the spectrum obtained using sparse representation method with those of beamforming, Capon's method, MUSIC, and minimum norm. In the top plot, the SNR is 10 dB, we have three sources, two are closely spaced ($5°$ of separation) and the third source is further away. The Sparse technique, Capon's, MUSIC and minimum norm are able to resolve the nearby sources and the other far-field, whereas beamforming method merge the sources closely spaced. In the bottom plot, we decrease the SNR to 0 dB, we can notice that the sparse technique is still able to resolve the three sources and even for high resolution methods but not with the same performance as in the top plot.

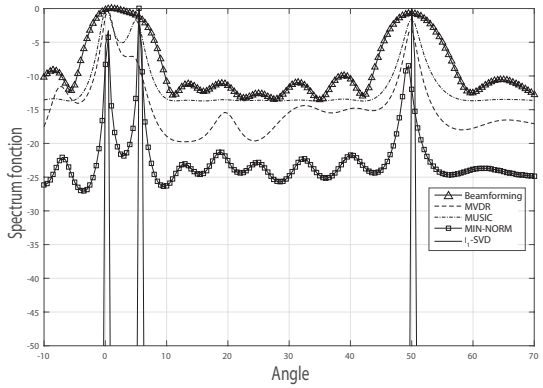


Fig. 5: Spectra for uncorrelated sources, SNR = -2dB, DOAs: $0°$, $5°$ and $50°$, using the presented method.

In figure (Fig. 5), we set the SNR to -2 dB, hence Beamforming, MVDR, MUSIC and minimum norm can locate the source $50°$, but not able to separate the closely-spaced sources, contrary to the minimum norm and sparse method are able to resolve the separation of the three sources. In the same, we descend the value of SNR to -5 dB and we can notice, in the figure (Fig .6), that the sparse technique is still successful in separating the closely spaced sources, while other localization methods have much difficulty to resolve closely-spaced sources, especially at low SNRs.

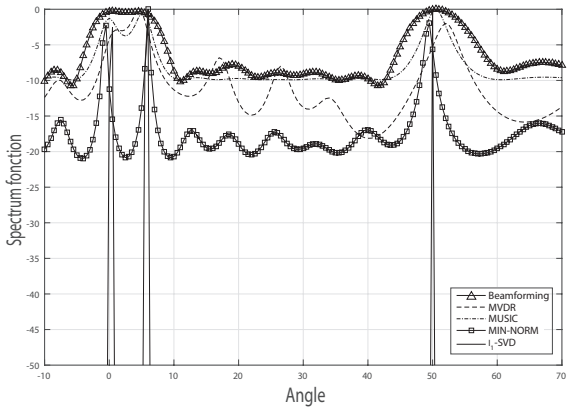


Fig. 6: Spectra for uncorrelated sources, SNR = -5dB, DOAs: $0°$, $5°$ and $50°$, using the presented method.

Continuously, in the figure (Fig. 7), we decrease the SNR to -10 dB until to -12. This improves the robustness of a sparse representation in resolving closely-spaced sources despite low SNR.

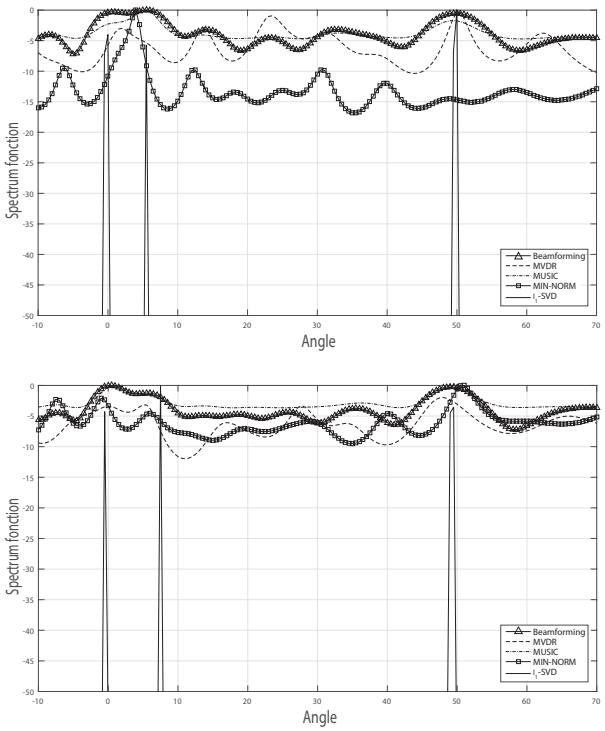


Fig. 7: Spectra for uncorrelated sources, DOAs: $0°$, $5°$ and $50°$. Top: SNR = -10 dB. Bottom: SNR = -12 dB, using the presented method.

As can be seen in Fig.1 to Fig.7, in noisy environments (low SNR), the sparse representation method performs better than all the others methods.

V. Conclusions

In this paper, we have discussed the performance analysis related to the resolution capability of DOA algorithms for localizing noise sources closely spaced. The algorithms have been simulated under different noise level environments. For each noise level, we have presented the performance of the resolution of the algo-rithms by searching to separate two closely spaced sources and one far-field. The results shown that in noisy environments, the sparse representation algorithm is the best performing one and requires fewer elements to separate the close sources. The minimum norm algorithm is more significant and in the same time is less sensitive to noise. Otherwise, in favorable environments, MUSIC performs well (hight SNR).

References

[1] SCHMIDT, Ralph. "Multiple emitter location and signal parameter estimation." IEEE transactions on antennas and propagation, 1986, vol. 34, no 3, p. 276-280.

[2] KHMOU, Youssef and SAFI, Said. Lorentzian model of spatially coherent noise field in narrowband direction finding. Int. Arab J. Inf. Technol., 2017, vol. 14, no 4A, p. 572-577.

[3] FRIKEL, Miloud, TARGUI, Boubekeur, SAFI, Said, and al. Bearing detection of noised wideband sources for geolocation. In : 18th Mediterranean Conference on Control and Automation, MED'10. IEEE, 2010. p. 1650-1653.

[4] LINHARES, Agostinho, TERADA, Marco Antonio Brasil, and SOARES, Antonio José Martins. Estimating the location of maximum exposure to electromagnetic fields associated with a radiocommunication station. Journal of Microwaves, Optoelectronics and Electromagnetic Applications, 2013, vol. 12, no 1, p. 141-157.

[5] PATWARI, Neal, O'DEA, Robert J., and WANG, Yanwei. Relative location in wireless networks. In : IEEE VTS 53rd Vehicular Technology Conference, Spring 2001. Proceedings (Cat. No. 01CH37202). IEEE, 2001. p. 1149-1153.

[6] ZIBULEVSKY, Michael and PEARLMUTTER, Barak A. Blind source separation by sparse decomposition in a signal dictionary. Neural computation, 2001, vol. 13, no 4, p. 863-882.

[7] SHUANGCAI, Luo, YING, Xiong, HAO, Cheng, and al. An algorithm of radar deception jamming suppression based on blind signal separation. In : 2011 International Conference on Computational Problem-Solving (ICCP). IEEE, 2011. p. 167-170.

[8] HADAMARD, Jacques. Lectures on Cauchy's problem in linear partial differential equations. Courier Corporation, 2003.

[9] ZHANG, Jialing, WU, Jimei, and HU, Bingbing. Dictionary Learning via a Mixed Noise Model for Sparse Representation Classification of Rolling Bearings. IEEE Access, 2020, vol. 8, p. 213416-213425.

[10] MALIOUTOV, Dmitry, CETIN, Müjdat, and WILLSKY, Alan S. A sparse signal reconstruction perspective for source localization with sensor arrays. IEEE transactions on signal processing, 2005, vol. 53, no 8, p. 3010-3022.

[11] FOUTZ, Jeffrey, SPANIAS, Andreas, and BANAVAR, Mahesh K. Narrowband direction of arrival estimation for antenna arrays. Synthesis Lectures on Antennas, 2008, vol. 3, no 1, p. 1-76.

[12] KHMOU, Youssef and SAFI, Said. DOA estimation with fourth order propagator. In : 2014 International Conference on Multimedia Computing and Systems (ICMCS). IEEE, 2014. p. 1295-1300.

[13] CHEN, Zhizhang, GOKEDA, Gopal, and YU, Yiqiang. Introduction to Direction-of-arrival Estimation. Artech House, 2010.

[14] EL OUARGUI, Ismail, SAFI, Said, and FRIKEL, Miloud. Minimum array elements for resolution of several direction of arrival estimation methods in various noise-level environments. Journal of Telecommunications and Information Technology, 2018, no 2, p. 87-94.

[15] CAPON, Jack. Hight-resolution frequency-wavenumber spectrum analysis. Proceedings of the IEEE, 1969, vol. 57, no 8, p. 1408-1418.

[16] SANTOSH, Sandeep and SHARMA, Karan. A Review on Multiple Emitter Location and Signal Parameter Estimation. International Journal of Engineering Research, 2013, vol. 2, no 3, p. 239-244.

[17] TROSCLAIR, Tyler M. and ADHIKARI, Kaushallya. Minimum Norm Method for Linear and Planar Sparse Arrays. arXiv preprint arXiv:2106.03666, 2021.

[18] GODARA, Lal C. Application of antenna arrays to mobile communications. II. Beamforming and direction-of-arrival considerations. Proceedings of the IEEE, 1997, vol. 85, no 8, p. 1195-1245.

[19] DONOHO, David L., ELAD, Michael, and TEMLYAKOV, Vladimir N. Stable recovery of sparse overcomplete representations in the presence of noise. IEEE Transactions on information theory, 2005, vol. 52, no 1, p. 6-18.

[20] PAPP, Dávid and YILDIZ, Sercan. alfonso: Matlab package for non-symmetric conic optimization. INFORMS Journal on Computing, 2022, vol. 34, no 1, p. 11-19.

[21] STURM, Jos F. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. Optimization methods and software, 1999, vol. 11, no 1-4, p. 625-653.

[22] DONOHO, David L., ELAD, Michael, and TEMLYAKOV, Vladimir N. Stable recovery of sparse overcomplete representations in the presence of noise. IEEE Transactions on information theory, 2005, vol. 52, no 1, p. 6-18.

Smart Livestock Monitoring System

Khalid EL MOUTAOUAKIL
LIMATI Laboratory
Polydisciplinary Faculty, Sultan Moulay Slimane University.
Beni Mellal, Morocco
elmoutaouakil.kh@gmail.com

Brahim JABIR
LIMATI Laboratory
Polydisciplinary Faculty, Sultan Moulay Slimane University.
Beni Mellal, Morocco
ibra.jabir@gmail.com

Noureddine FALIH
LIMATI Laboratory
Polydisciplinary Faculty, Sultan Moulay Slimane University.
Beni Mellal, Morocco
nourfald@yahoo.fr

***Abstract*— It has become necessary to adopt a set of agricultural practices based on advanced agricultural technology in accordance with the agriculture 4.0 revolution, in order to meet the challenges facing farmers in their quest to meet the growing demand for food. In this regard, this work presents a smart livestock monitoring system architecture that aims to use a set of advanced agricultural technologies adapted to the condition and needs of each animal, in order to offer livestock farm managers and veterinarians a system for monitoring and tracking the health, production and status of their livestock. We also highlight the latest advanced technologies and their use for livestock farming and digital agriculture in general.**

Keywords— Smart livestock farming, Livestock monitoring, Digital agriculture, Smart livestock management.

I. INTRODUCTION

Agriculture digitalization has an important space on the agenda of governments in their quest to address the challenges of food and nutrition insecurity, climate change, youth unemployment and overall economic growth. With the right policies, innovation and investment, the agriculture could not only be able to feed a growing population, but also to create decent jobs for millions of young people [1].

Technology, as we have seen in other sectors, is key to influencing change and driving development [2]. It brings countries together, reduces trade barriers and offers a window of opportunity for young "digital native" entrepreneurs at the forefront of innovation applied to different economic sectors. In agriculture, digitalization could be a game-changer by increasing productivity, profitability and resilience to climate change [3].

As expected, in recent years, the agricultural technology sector has evolved like other sectors. With this change, traditional production systems have transformed into modern, productive and innovative systems [4]. Thus, we are now taking about agriculture with water saving, smart agriculture, high quality, high efficiency and non-polluting agriculture. Agriculture 4.0 is the most effective and necessary approach to achieve all these transformations [5].

The techniques of transition to the agriculture 4.0 standards are numerous, namely, artificial intelligence, data science, analytics, Blockchains, Big data, IoT (Internet of Things), 3D printing [6]. This digital transformation affects a set of agricultural areas, namely livestock sector, irrigation sector, cultivation, etc [7].

After this introduction, in the following second section, we review the state of the art of the latest advanced technologies in use as part of the agriculture 4.0 in general. In the third section, we look more closely at a smart livestock monitoring system architecture that we propose for smart monitoring of livestock and the advanced technologies and tools in use for precision livestock monitoring and tracking of the health, status and production of the livestock. In the fourth section, we focus on the security issues and challenges that faces livestock monitoring and tracking systems. Then, we conclude this paper and propose some prospects for development.

II. AGRICULTURAL ADVANCED TECHNOLOGIES

In the following subsection, we present the key important advanced technologies for the agriculture 4.0 that has a large effect on the agriculture industry.

A. Big data analytics for the agriculture

Big data analytics reflects the challenges posed by data that is too large, too loosely structured, and too fast to be managed by traditional methods. Farms and agricultural activities now routinely generate data of an unprecedented scale and complexity. Organizations around the world are placing increasing importance on extracting relevant insights from massive amounts of data. Trying to extract meaningful information efficiently, quickly and easily from these data sources is a challenge. Thus, analytics has become a vital thing to realize the full value of big data to improve agricultural business performance [8]. In the following sub sections, we highlight some of the best and most used, mostly open source, big data technologies that work as a big data analytics system to analyze large amounts of unstructured data to make decisions:

1) Apache Hadoop

Hadoop, the most commonly used framework, combines hardware and open-source software. It takes incoming data streams and distributes them across blocks. It also provides tools to analyze the data. It's a framework that enables distributed processing of large data sets on a cluster of computers. It is designed to scale from a single server to thousands of machines, each providing local compute and storage [9].

2) MapReduce

Introduced by Google in 2004, MapReduce is a distributed data processing model. It consists of dividing problems into two parts: a Map function that processes source data into sufficient statistics and a Reduce function that merges all sufficient statistics into a final answer [10]. Once the Map function has been applied to all data, the Reduce function can be run to combine the results of the Map phases. MapReduce provides the fastest, most cost-effective, and scalable mechanism for returning results. Today, most of the advanced technologies for the management of Big Data are developed on MapReduce.

3) Apache Spark

Apache Spark is an open-source big data processing framework designed for speed, ease of use, and sophisticated analytics. Spark enriches MapReduce with new features such as in-memory data storage and near real-time processing. The performance can be many times better than other big data technologies [11].

B. Artificial intelligence-based agriculture

The advancement of machine learning (ML) with machine vision will make agricultural technologies accurate, robust and inexpensive. Artificial intelligence (AI) can be used to identify and clean data or use data as a way to establish contextual awareness. AI contributes to the speed of data, facilitating fast computing decisions that lead to other decisions. AI contributes to variety attenuation by capturing, structuring, understanding unstructured agricultural data generating structure data. In addition, AI enables agricultural data analysis and decision making.

Machine Learning (ML) is used to teach machines to process data more efficiently [12]. With the abundance of datasets available, the demand for machine learning is on the rise. Many industries, including the agriculture, are applying machine learning to extract relevant information.

The goal of ML is to learn from data. Many studies have been done on how to make machines learn by themselves.

ML algorithms are generally divided into two categories: supervised and unsupervised [13]. Semi-supervised algorithms represent a hybrid of the two, but have been used less often to solve problems.

Deep Learning (DL) extends classical ML by adding more "depth" (complexity) to the model based on artificial neural networks (ANN). Each neural network consists of a sequence of layers through which data passes. These layers are made up of neurons, and the neurons in one layer are connected to the ones in the next layer. These connections are defined by what we call a weight (a numerical value). Each layer also has what is called a bias. Data starts at the input layer and is transformed as it passes through subsequent layers.

There are a lot of research that uses DL to build solutions for problems in the agricultural industry.

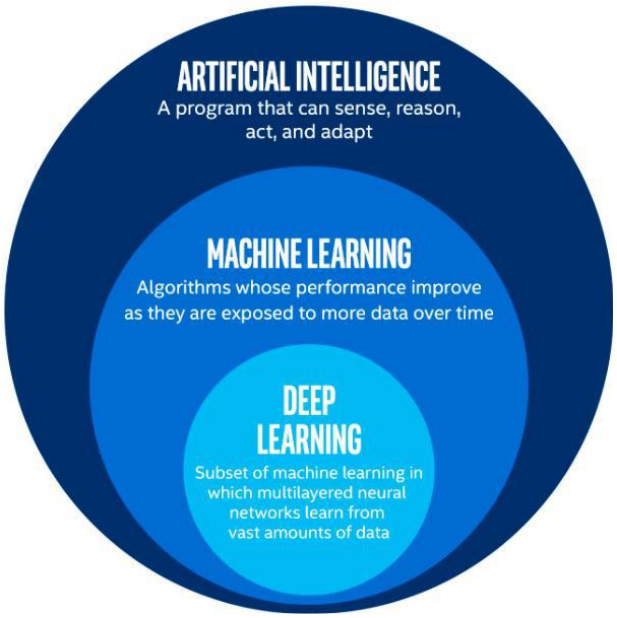


Figure 1: Relation between AI, ML and DL

C. Agricultural Internet of Things (AIOT)

Agricultural Internet of Things (AIoT) is considered as an indispensable technology for smart farms, while its development in relation to the future internet approach provides a basis for a new generation of agricultural management information systems, allowing entire smart farms to become active nodes in agricultural value chains.

Both in the agricultural sector and in the industrial sector, IOT technology can be expressed by some specific characteristics, such as heterogeneity, due to a variety of devices, interconnectivity, high scalability, object-related services, and lastly, dynamic changes since the state of devices can dynamically change at any time [14].

D. Agricultural Cloud Computing

Cloud computing is a computing method based on the Internet, in this way, shared software and hardware resources and information can be provided to various computers on demand. Agricultural environments tend to generate voluminous amounts of heterogeneous data that require huge storage spaces, while the development of sophisticated intelligent infrastructures constantly increases the demand for more complex intelligent services [15], the realization of these operations on local servers could result in lack of scalability and inefficient farm management systems, therefore the need to use cloud computing.

E. Blockchains for digital agriculture

The use of Blockchain in the agriculture improves food security by allowing information to be traced across the food supply chain. The ability of blockchains to store and manage data creates traceability, which is in use to facilitate the development and deployment of intelligent farming and index-based crop insurance advances. In the Agriculture 4.0 context, this is a significant step forward.

III. LIVESTOCK MONITORING SYSTEM ARCHITECTURE

Precision livestock farming (PLF) is the use of advanced technologies to optimize the production and welfare of each animal through continuous real-time monitoring and tracking of its health, welfare, production and environmental impact. In this context, this system comes to provide a solution for those concerned.

Because animal reactions can be very rapid, we need an ongoing and continuous monitoring and management tool. Depending on the monitored variable, the word "continuous" can mean every second (for example, for stress monitoring) or once a day (for example: for weight monitoring). When PLF technologies are installed in livestock farms, they generate a huge amount of data. This is why we use cloud storage and real-time data analysis which allows relevant insights to be calculated from the data.

To monitor and track the health, the production and the status of a livestock, we propose a smart monitoring system. The input phase of the system consists of several IoT sensor modules and devices. For processing, we use a system unit that process and sends data to a cloud server that has a dedicated web application accessible from computers and mobile devices.

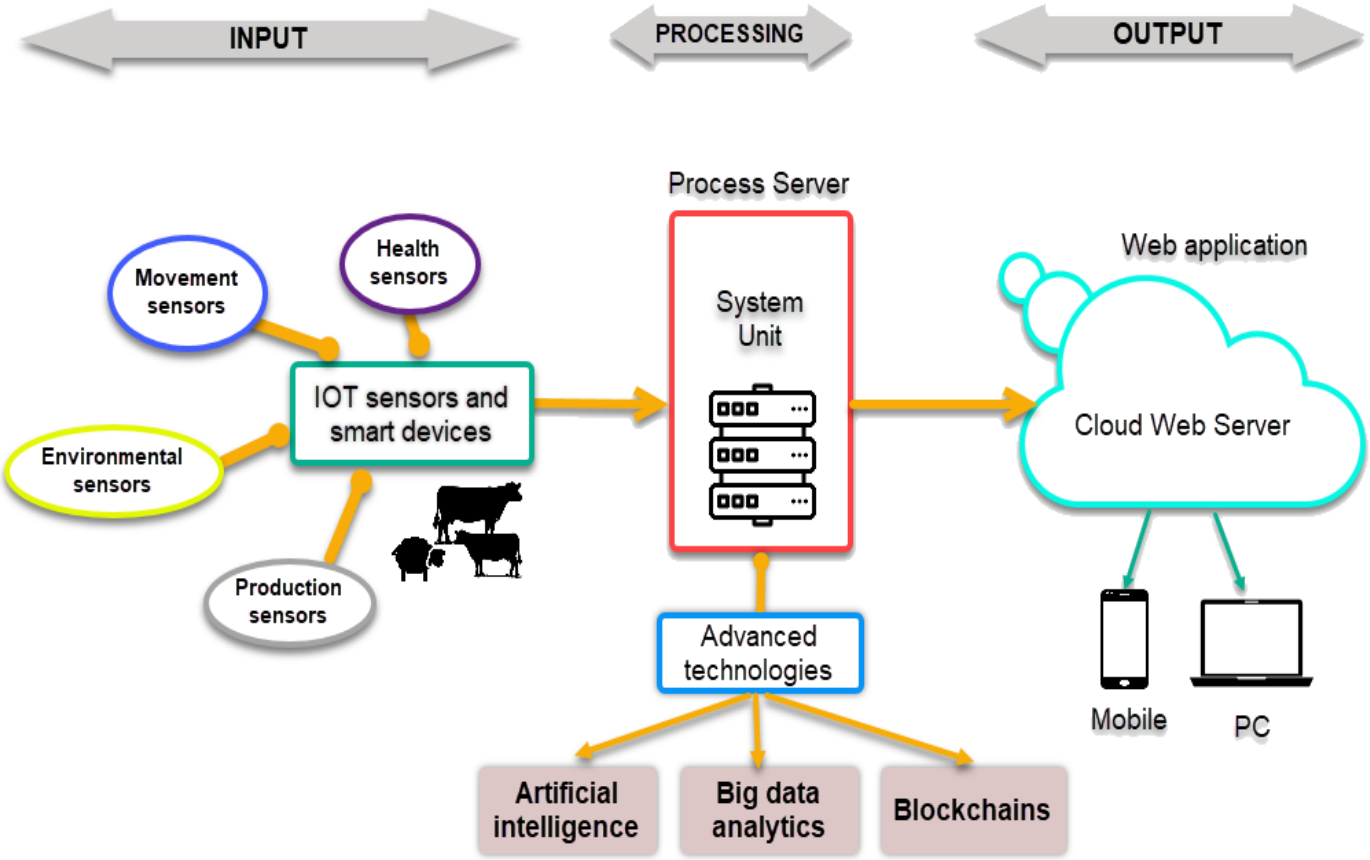


Figure 2: Connected and smart live monitoring system architecture

The figure 2 consists of a complete block diagram of the proposed system. The system comprises smart sensors and devices to track and monitor the health, movement, environment and production of the cattle. All the sensor modules data are processed in the system unit then stored in a cloud web server, allowing farmers and barn managers to remotely accessing all the data and analysis using a web application available for mobile and computer devices. Also, veterinary physicians can access the data to give the diagnosis.

This proposed system will help the different actors to keep track of all the details of their livestock. This can be used in barns as well as in veterinary hospitals. The following section presents more details on the technologies in use in the system.

In the following subsections, we highlight the latest advanced technologies in use for the smart livestock monitoring systems.

A. IOT sensors and smart devices

Smart sensors and devices based on the Internet of Things can monitor in real time the status and health of animals and record illness or disease-related data, allowing farmers or livestock manager to take preventive actions at an early stage.

These sensors and devices can be used for monitoring disease, stress, environment, feed intake, rumination, humidity and poultry, among others.

- Health sensors could include among others temperature, heartbeat, oxygen (O2) and hydrogen sulfide (H2S) sensors.
- Movement sensors may include motion, pulse, gesture and respiration sensors among others.
- Environmental sensors include cameras and ultrasonic sound smart sensors.
- Production sensors may include rumination, PH and humidity sensors among others.

B. System processing units

The system units have two use cases, the first is as a local gateway that receives data from the smart sensors and devices because mot of the sensors are not able to connect to the internet to share data. While the second use case is the control and processing of data before making it available in the cloud web servers [16].

The system processing unit use advanced technologies like Artificial Intelligence (AI) and especially Deep learning (DL) to extract relevant information from the data and Big data analytics to make decisions, in addition to blockchains to facilitate the development and deployment of the smart farming.

C. Cloud web severs

Huge Cloud servers are coupled to perform important tasks such as the storing of health data, analyzing it, and making it available for users throw a web application, and, in some cases, sending alerts to the persons concerned. Data analytics resources are also deployed on the internet or in the cloud, allowing for on-demand resources to be accessed via network infrastructure [17].

D. Communication technologies

Communication technologies consist of multiple IoT protocols to collect and analyze livestock data. The most commonly used protocols in this scenario are internet-related technologies such as WIFI, LoraWan, Zigbee, etc [18].

In IoT-enabled livestock monitoring, these communication protocols work as a nerve to transfer data. When LTE, GSM, and CDMA technologies are unavailable, Zigbee is used as the principal facilitator for communication [19].

IV. CHALLENGES AND SECURITY ISSUES

In the following subsections, we take a close look at the security issues and challenges that faces livestock monitoring and tracking systems that uses the aforementioned advanced technologies.

1) Challenges

The use of IoT sensors for livestock monitoring and tracking poses several challenges related to data management, taking into account the large amount of data that the sensor send during the continuous monitoring. Also, there is a hardware and software challenges for collecting, storing and processing the data, with the possibility of a false data injection.

Also, in some cases, we may have problems with establishing secure communication and sensor deployment issues. In addition to implementation challenges related to the high infrastructure cost, the lack of awareness and the regulatory concerns [20].

2) Security issues

Smart livestock farming systems require authentication, security, and confidentiality [21]. As a result, it's critical to protect IoT networks from outside threats. Also, securing the hardware are one of the most relevant security challenges. Thus, due to the scattered nature of IoT devices and the fact that they are deployed in a variety of environments, a single security protocol is insufficient [22]. Security attacks are generally related to data leakage that can expose sensitive information. Thus, it is necessary to propose a security model to protect the monitoring system. The most common security threats in this regard include:

- Denial of services (DoS) attacks:

DoS attacks are common in farms because of the vast number of linked devices. The normal functionality of livestock ecosystems is disrupted by these attacks.

- Cloud computing attacks:

In cloud environments, if one virtual machine (VM) is vulnerable, all auto scaled VMs are likely to be vulnerable as well. In this case, malware that has infected one VM can quickly spread to all other VMs.

- Side-channel attacks:

In this type of attack, the attackers exploit hardware and software glitches like voltage variance, cache hit, cache miss timing patterns, acoustic and sound channels.

- Misinformation attacks:

Data integrity may be jeopardized by the attackers. In this type of attack, they spread fake information about the animal's health status. It takes a lot of effort, money, and time to prove that an erroneous data report has been produced.

- Radio frequency (RF) jamming attack:

In order to improve the efficiency of tracking the livestock, the use of global navigation satellite systems (GNSS) is very common. Attackers might disrupt the GNSS over a broader area, preventing tracking systems from working properly.

- Malware injection attacks:

Malware injection attacks, in which an attacker injects malware into connected devices, are one of the most common threats that affect smart livestock farming systems [23].

Malware is a malicious software that can harm smart farm monitoring systems, spread without the consent of the livestock managers and making the monitoring systems target for attackers.

V. Conclusion & Future works

There is a common determination on the part of all actors to overcome the obstacles facing agriculture today. In this context, this research comes in order to offer farmers, livestock farm managers, and veterinarians a system based on the latest advanced technologies in order to help them track and monitor all the details of their livestock. It's a primary architecture of the smart livestock monitoring system, and as future work, we plan to develop this system and test it in a real livestock farm and deploy the different devices and technologies including the smart sensors, the system processing unit, the cloud web server and the other related devices and technologies.

References

[1] Ozdogan, B., Gacar, A., & Aktas, H. (2017). Digital agriculture practices in the context of agriculture 4.0. Journal of Economics Finance and Accounting, 4(2), 186-193.

[2] Trendov, M., Varas, S., & Zeng, M. (2019). Digital technologies in agriculture and rural areas: status report. Digital technologies in agriculture and rural areas: status report.

[3] Tang, S., Zhu, Q., Zhou, X., Liu, S., & Wu, M. (2002, June). A conception of digital agriculture. In IEEE international geoscience and remote sensing symposium (Vol. 5, pp. 3026-3028). IEEE.

[4] Andrade-Sanchez, P., & Heun, J. T. (2010). Understanding technical terms and acronyms used in Precision Agriculture.

[5] Yane, D. (2010, October). Research and analysis about system of digital agriculture based on a network platform. In International Conference on Computer and computing technologies in agriculture (pp. 274-282). Springer, Berlin, Heidelberg.

[6] Jabir, B., & Falih, N. (2020, April). Digital agriculture in Morocco, opportunities and challenges. In 2020 IEEE 6th International Conference on Optimization and Applications (ICOA) (pp. 1-5). IEEE.

[7] Rabhi, L., Falih, N., Afraites, L., & Bouikhalene, B. (2021). Digital agriculture based on big data analytics: a focus on predictive irrigation for smart farming in Morocco. Indonesian Journal of Electrical Engineering and Computer Science, 24(1), 581.

[8] Falih, B. J. N., & Jabir, B. Big data analytics opportunities and challenges for the smart enterprise. (2021). International Journal on Technical and Physical Problems of Engineering (IJTPE), (47), 20-26.

[9] Manikandan, S. G., & Ravi, S. (2014, October). Big data analysis using Apache Hadoop. In 2014 International Conference on IT Convergence and Security (ICITCS) (pp. 1-4). IEEE.

[10] Himesh, S., Prakasa Rao, E. V. S., Gouda, K. C., Ramesh, K. V., Rakesh, V., & Mohapatra, G. N. (2018). Digital revolution and Big Data: a new revolution in agriculture. CAB Rev, 13(21), 1-7.

[11] Padhy, R. P. (2013). Big data processing with Hadoop-MapReduce in cloud systems. International Journal of Cloud Computing and Services Science, 2(1), 16.

[12] Coelho, L. P., & Richert, W. (2015). Building machine learning systems with Python. Packt Publishing Ltd.

[13] Bowles, M. (2019). Machine Learning with Spark and Python: Essential Techniques for Predictive Analytics. John Wiley & Sons.

[14] Lakhwani, K., Gianey, H., Agarwal, N., & Gupta, S. (2019). Development of IoT for smart agriculture a review. In Emerging trends in expert applications and security (pp. 425-432). Springer, Singapore.

[15] Symeonaki, E. G., Arvanitis, K. G., & Piromalis, D. D. (2017, September). Cloud computing for IoT applications in climate-smart agriculture: A review on the trends and challenges toward sustainability. In International Conference on Information and Communication Technologies in Agriculture, Food & Environment (pp. 147-167). Springer, Cham.

[16] Delgado, J. A., Short Jr, N. M., Roberts, D. P., & Vandenberg, B. (2019). Big data analysis for sustainable agriculture on a geospatial cloud framework. Frontiers in Sustainable Food Systems, 3, 54.

[17] Patrício, D. I., & Rieder, R. (2018). Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. Computers and electronics in agriculture, 153, 69-81.

[18] Vigneswari, T. (2021). Smart IOT Cloud Based Livestock Monitoring System: A Survey. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(10), 3308-3315.

[19] Chaudhry, A. A., Mumtaz, R., Zaidi, S. M. H., Tahir, M. A., & School, S. H. M. (2020, December). Internet of Things (IoT) and machine learning (ML) enabled livestock monitoring. In 2020 IEEE 17th International Conference on Smart Communities: Improving Quality of Life Using ICT, IoT and AI (HONET) (pp. 151-155). IEEE.

[20] Farooq, M. S., Sohail, O. O., Abid, A., & Rasheed, S. (2022). A Survey on the Role of IoT in Agriculture for the Implementation of Smart Livestock Environment. IEEE Access.

[21] Miorandi, D., Sicari, S., De Pellegrini, F., & Chlamtac, I. (2012). Internet of things: Vision, applications and research challenges. Ad hoc networks, 10(7), 1497-1516.

[22] Li, L. (2012, May). Study on security architecture in the Internet of Things. In Proceedings of 2012 international conference on measurement, information and control (Vol. 1, pp. 374-377). IEEE.

[23] Gruschka, N., & Jensen, M. (2010, July). Attack surfaces: A taxonomy for attacks on cloud services. In 2010 IEEE 3rd international conference on cloud computing (pp. 276-279). IEEE.

The distribution of heuristic methods applied to complex systems

Elhoucine Ouassam, Nabil Hmina and Belaid Bouikhalene
LIMATI laboratory, Sultan Moulay Slimane university
Beni mellal, morocco

H. Hachimi
LGS Laboratory, ENSA of Kenitra, Ibn Tofail University
Beni mellal, morocco

***Abstract*— This work aims to study and present an overview of the complex system and heuristic techniques applied to an optimization problem, the methods presented here are the Local search methods, Population-based heuristics, and Constructive heuristics. Thereafter, we will talk about the distributed heuristic methods in which we mention Distributed CHS, Contract net protocol, and artificial potential field (APF). In addition, we will give an example of a complex system.**

Keywords*— *heuristic methods; complex systems; Optimization problems; Smart-Grid; distributed heuristics;

I. INTRODUCTION (*HEADING 1*)

Optimization problems are those where mathematical techniques are generated to find optimal solutions within a finite set of possible strategies. The set of solutions is generally defined by a set of restrictions, and the set is too large for an exhaustive search. A well-known example is the knapsack-problem, where the value of the goods carried in the knapsack has to be maximized, while the weight of the goods that can be is limited.

Combinatorial optimization (CO) is an essential field in the domain of CO and theoretical artificial intelligence that aims to use combinatorial techniques to solve discrete optimization. A discrete optimization problem seeks to determine the best possible solution from a finite set of possibilities. Computing optimal solutions are intractable for many optimization problems of industrial and scientific importance. In practice, we are usually convinced with good solutions, which are generated by heuristic or metaheuristic algorithms[2].

In this context appeared the Heuristics represent a family of approximate optimization techniques that gained a lot of popularity in the past two decades. They are among the most promising and successful techniques.

Heuristics provide 'acceptable' solutions in a reasonable time for solving hard and complex problems in science and engineering. This explains the significant growth of interest in the metaheuristic domain. Unlike exact optimization algorithms, heuristics do not guarantee the optimality of the obtained solutions. Heuristics are techniques that seek optimal or near-optimal solutions in a population-based and stochastic-based manner and are inspired by biological or human intelligence phenomena.

In this article, we will give a general view of complex systems, and the set of heuristics used to solve them, and then we will give an example of a complex system is the smart grid.

II. COMPLEX SYSTEM

A. *Definition*

First, confirm that you have the correct template for your paper size. This template has been tailored for output on the A4 paper size. If you are using US letter-sized paper, please close this file and download the file "MSW_USltr_format".

The complex system may have many components that collaborate to create a functioning whole[3]. Thereby the function creates itself, i,e it 'comes about' by the dynamical interaction of the components without an intervening regulatory body. On speaks of « self-organization » or also of « emergence ». Important is that the word « complex » is not confused with the word « complicated ». Let's consider, for example, an « architectural building complex » that serves a multitude of functionalities. It is not complicated but shows its true functional complexity when we use it.

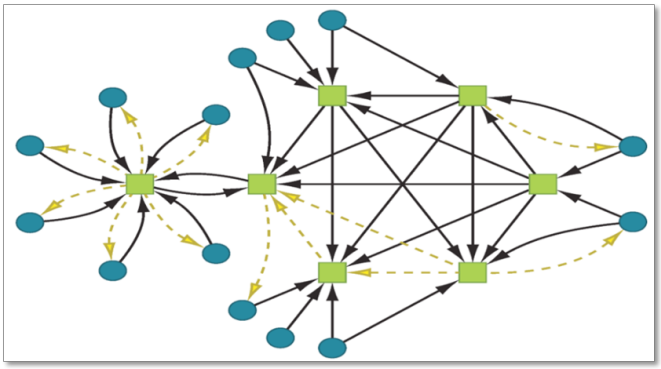


Fig1. Dynamics of complex system schematic diagram of banks (■) and firms (●) network, with connectors being either loan (◥) or deposits (◥). Banks are connected to firms and each other, whereas firms only interact with banks [4].

B. Features of a complex system

a. Non-linear feature

Complexity is the product of several different systems. This is straightforward. The more ports are the more complex it will be. Secondly, complexity is a product of the degree of connectivity between the elements. The more interconnected and interdependent they are, the more complex our system will be. Within simple systems. There are a few connections between elements, and it is relatively easy to understand the direct relations of cause and effect, that is to say, we can draw a direct line between a single cause and a single effect. Thus, we call these simple organizations, linear systems. But when we turn up the connectivity within the system and especially when there are a high number of elements, these cause and effect relations become more complex as there may be multiple causes for any given effect or vice versa, as opposed to our simple linear system. We call these more complex organizations non-linear systems, and non-linearity is a key property of complex systems. ▪ Self-organization

This gives rise to another important feature of complex systems which is self-organization. When elements have the autonomy to adapt locally, they can self-organize to form global patterns. The process through which this takes place is called emergence. Thus. As opposed to simple linear systems where order typically comes. From some form top-down, centralized coordination patterns for order within complex systems emerge from the bottom up. Self-organization is another term across all types of complex systems.

b. Diversity

Complexity is also a product of the degree of diversity between elements, when all the elements within our system are very similar or homogenous, then it is much simpler to model, design, or be managed, as opposed to dealing with a heterogeneous organization presented with numerous parts, each with its own set of characteristics.

c. Emergence

Complex systems show emergent behavior. Out of the interactions between the individual elements in the system, behaviors emerge at the level of the system as a whole. This so-called higher-order behavior cannot simply be derived by aggregating behavior at the level of the elements. The whole is more than the sum of its parts. This higher-order was not intended by the elements. It is a spontaneous order.

d. Limited predictability

The behaviors of complex systems cannot be predicted well. A small change in initial conditions or history can lead to very different dynamics over time. The existence of non-linear behavior also adds to unpredictability.

e. Feedback

The action of going back automatically, after a disturbance for example. A closed system contains a circular process in which the system's output is returned or "fed back" to the system as input.

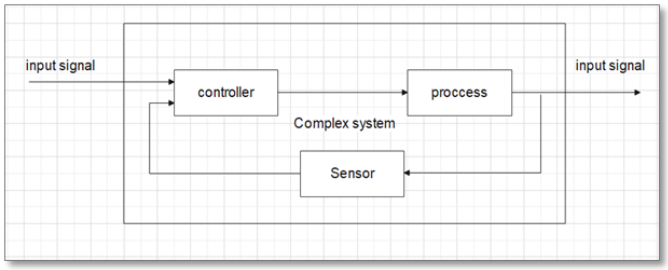


Fig2. Feedback control loop

C. Example of the complex system: the Smart-Grid system

a. Definition of smart-grid

The "Smart Grid" designates an electrical energy distribution system that automatically adapts, in autonomy, production to demand.

To achieve this just-in-time balance, the Smart Grid uses a network of sensors, and devices for the transmission and computer analysis of data in real-time (as well as Big Data), which integrate and influence modes of production and consumption to achieve optimal results in terms of energy efficiency and security.

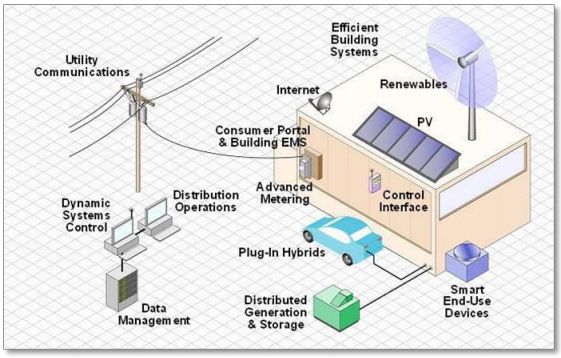


Fig3. Scheme of Smart Grid and its components.

b. Goals of smart grid

In [5] T. Vijayapriya and D. Kothari, group the objectives of the smart grid by :

- Better facility for the connection and operation of generators of all sizes and technologies.
- Allow consumers to play a part in optimizing the operation of the system.
- Provide consumers with greater information and options for how they use their supply.

- Significantly reduce the environmental impact of the whole electricity supply system.
- Maintain or even improve the existing high levels of system reliability, and quality of supply.
- Maintain or improve the existing services efficiently.

c. Smart-grid domains and their interconnections

Smart-grid communication networks support interactions between entities (users, systems, and applications) within the many organization and locations of the utility. These entities can be grouped into broad « domains ».
Traditionally, electricity flows from bulk generation sources to the transmission system of transmission lines and transmission substations systems from where it is delivered through the distribution substations to customers (consumers of energy) over feeders. With a smart grid, electricity is also generated from DG(distributed generation) sources connected to the distribution system. DG was included as part of the distribution and custom domain.

III. HEURISTIC METHODS

A heuristic method is a mental strategy that permits scientists to solve problems and make a decision quickly and efficiently. These experimental strategies decrease decision-making time and allow people to function without constantly stopping to think about their next course of action. Heuristics are helpful in many situations, but they can also lead to cognitive biases. In the following, we will present the main categories of heuristic methods.

A. Local search methods

Local search is the most successful general approach for finding high-quality solutions to hard combinatorial optimization problems in a reasonable time. It is based on the iterative exploration of neighborhoods for solutions to improve the current solution through local changes. The type of local changes that may be applied to a solution is defined by neighborhood structure.

The choice of an adequate neighborhood structure is essential for the efficiency of the local search and often has to be done in a problem-specific way. It defines the solutions that can be reached in one step of a local search algorithm. Typically, a neighborhood structure is not defined by explicitly enumerating the set of possible neighbors, but implicitly by defining the possible local changes that may be applied to a solution. The neighborhood strategy can also be conceived as a graph, called the neighborhood graph. The nodes of the graph are the solutions and two solutions are connected if they are neighbors. Thus, the execution of a local search algorithm corresponds to walking on the neighborhood graph.

finally, the solution found by a local search algorithm will not be globally optimal, it may be guaranteed to be optimal concerning local changes.

The most basic locale search algorithm is an iterative improvement. It typically starts with an initial solution generated randomly or by some constructive heuristic and tries to reputedly improve the current solution by moving to better neighboring solutions. If in the neighborhood of the current solution a better solution s' is found, it replaces the current solution and the search is continued from s'; if no better solution is found, the algorithm terminates at a local minimum. Figure 4 is given an outline of the iterative improvement algorithm.

```
Procedure FirstImprovement (s∈S)
   W=0
  While(N(s)/w ≠ ⌀) do
      s' =GenerateSolution(s)
      if f(s') <f(s)
          return s'
      else
          w=w∪s'
   end
end FirstImprovement
```

Fig4: Algorithm skeleton of iterative Improvement.

B. Population-based heuristics

P-heuristic utilizes a lot of common strategies. They could be viewed as an iterative improvement in a population of solutions. First, this population is initialized. secondly, a new population is generated. in the end, this novel population is integrated into the current one with some selection methods. The search process is stopped when a given condition is satisfied (stopping criterion). Algorithms such as evolutionary algorithms, scatter search, estimation of distribution algorithms, particle swarm optimization, bee colony, and artificial immune systems belong to this class of metaheuristics.

C. Constructive heuristics

A constructive heuristic (CH) is a novel kind of heuristic method, in that strategy initiates with an empty solution and systematically extends the last solution until an optimal solution is obtained. It differs from local search heuristics which initiate with a complete solution and then try to improve the current solution further via local moves.
CH algorithms make the choice that is best actually. in other words, it makes a locally optimal choice to that this choice will lead to the globally optimal solution. In such an algorithm of this type, we went through the following steps:

- Choice the optimal structure of the situation.
- construct a cyclic solution.

- Prove that at any stage of the recursion, one of the optimal selections is the greedy choice. Therefore, it is always correct to make a greedy choice.
- hailaitgh that all but one of the sub-problems induced by having made the constructive selection are null.
- Construct a cyclic method that executes the greedy strategy.
- Convert the cyclic method to a cyclic algorithm.

IV. DISTRIBUTION OF HEURISTIC METHODS

A. Distributed CHS(constrained Heuristic Search)(DCHS)

A distributed CHS problem is a problem where the variables are distributed among a set of agents. Each agent is responsible for a set of variables and their instantiation. Constraints and satisfiability conditions exist among variables under the jurisdiction of different agents. The instantiation of variables must satisfy that a distributed CHS problem is solved if an assignment is found of values to all variables of all agents such that all constraints and satisfiability specifications are simultaneously satisfied.

Distributed DCHS is a process carried out by a group of agents each of which has 1) limited knowledge of the environment, 2)limited knowledge of the constraints and requirements of other agents, 3) limited number and amount of resources that are required to produce a system solution. Global system solutions are arrived at by interleaving local computations and information exchange among the agents. There is no single agent with a global system view. In such an environment, DCHS is an incremental process. Agents make local decisions about assignments of values to particular variables at particular times during problem-solving and a complete solution is formed by incrementally merging partial solutions.

B. Contract net protocol

A contract net is a collection of interconnected processor nodes whose interactions are governed by a problem-solving protocol based on the contract metaphor. Each processor node in the net operates asynchronously and with relative autonomy. Instances of the execution of individual tasks are dealt with as contracts. A node that generates a task advertises the existence of that task to the other nodes in the net as a task announcement, then acts as the manager of that task for its duration. In the absence of any information about the specific capabilities of the other nodes on the net, the manager is forced to issue a general broadcast to all nodes. If, however, the manager possesses some knowledge about which of the other nodes in the net are likely candidates, then it can issue a lilted broadcast. Finally, if the manager knows exactly which of the other nodes is appropriate, then it can issue a point-to-point announce may be madcemente by various managers.

The other nodes in the net have been listening to the task announcement, and have been evaluating their level of interest in each task concerning the specialized hardware and software resources. When a task is found to be of sufficient interest, a node may submit a bid. Each bid indicates the capabilities of the bidder that are relevant to the execution of the announced task. A manager may receive several such bids in response to a single task announcement; based on the information in these bids, it selects one (or several) node(s) for the execution of the task. The selection is communicated to the successful bidder(s) through an award message. These selected nodes assume responsibility for the execution of the task, and each is called a contractor for the task.

A contract is thus an agreement between a node that generates a task (the manager) and the node that executes the task (the contractor). Note that establishing a contract is a process of mutual selection. Available processor nodes evaluate task announcements nodes by several managers until they find one of interest; the managers then evaluate the bids received from potential contractors and select one they determine to be most appropriate. Both parties to the agreement have evaluated the information supplied by the other and a mutual decision has been made.

C. Artificial potential field (AFP)

In 1986 khatib was introduced The AFP. The potential function is definable over free C-space as the sum of attractive potential pulling a robot toward the goal configuration, and a repulsive potential pushing a robot away from obstacles. An AFP is an important classic method for autonomous mobile robots. Many researchers are studying it continually all over the world. An AFP has often represented a good quality method to complete a fast and reactive response to a dynamic environment. However, this method has been widely demonstrated as suffering from unavoidable drawbacks which make it very likely that a robot will become trapped in a local minimum and oscillations. The paper [6] describes a hybrid method that integrates firstly knowledge of an environment with local perceptions to implement the assigned objectives correctly. The results indicate that this approach assumed that the robot can never be trapped in deadlocks even when operating within a partially unknown dynamic environment. Despite its good properties, the navigation system described in this paper includes a typical shortcoming: the system relies on local perceptions and navigation strategies. Another improved APF has been proposed [7] using quantum particle swarm optimization to modify the parameters of the APF to adapt to a different environment and dynamic obstacles. To address the local minima problem in the traditional APF, a method including robot regression and potential field filling has been proposed (Qi et al.,2008; Shi et al., 2010). Similar methods have been proposed in other papers (Zhang et al., 2006a; Yu et al.,2011). Before calculating the resultant force that is put on an object in the potential field, they build links among closed obstacles to optimize the planned solution. Improved APFs of other kinds have been investigated (He et al., 2011). They

determine the relative distance between a robot and the objective into a repulsive force function and change the repulsion direction to respond that the global minimum is at the objective position. Donnart and Mayer(1996) researched the learning reactive and planning rules in mobile robot path planning. The main thrust of some reports (Sheng et al.,2010; Yang et al., 2011) is that application of a virtual local target to a guide robot escapes the local minimum.

V. CONCLUSION

With the successful development of heuristic approaches, it has become difficult to choose the most suitable ones for the application at hand. So, it is very important to be able to compare different optimization algorithms, to be able to choose the most efficient for a given problem.

To have a common base, but also to have a representative set of many of the problems that can be found in reality, test protocols generics have been introduced. These test sets, called benchmarks, are used to assess the performance and convergence capabilities of optimization algorithms. They group different objective functions and test protocols to compare the results obtained by the scientists.

REFERENCES

[1] « Exact methods for the knapsack problem and its generalizations - ScienceDirect ». https://www.sciencedirect.com/science/article/abs/pii/0377221787901652 (consulté le 8 mai 2022).

[2] A. H. Gandomi, X.-S. Yang, S. Talatahari, et A. H. Alavi, Metaheuristic Applications in Structures and Infrastructures. Newnes, 2013.

[3] J. Ladyman, J. Lambert, et K. Wiesner, « What is a complex system? », Euro Jnl Phil Sci, vol. 3, no 1, p. 33-67, janv. 2013, doi: 10.1007/s13194-012-0056-8.

[4] S. Poledna, A. Hinteregger, et S. Thurner, « Identifying Systemically Important Companies by Using the Credit Network of an EntireNation », Entropy, vol. 20, no 10, Art. no 10, oct. 2018, DOI: 10.3390/e20100792.

[5] T. Vijayapriya et D. P. Kothari, « Smart Grid: An Overview », Smart Grid and Renewable Energy, vol. 2, no 4, Art. no 4, nov. 2011, DOI: 10.4236/sgre.2011.24035.

[6] A. Sgorbissa et R. Zaccaria, « Planning and Obstacle Avoidance in Mobile Robotics », Robotics and Autonomous Systems, vol. 60, p. 628-638, avr. 2012, DOI: 10.1016/j.robot.2011.12.009.

[7] F. Zhang, D. Pant, et B. E. Logan, « Long-term performance of activated carbon air cathodes with different diffusion layer porosities in microbial fuel cells », Biosensors and Bioelectronics, vol. 30, no 1, p. 49-55, déc. 2011, DOI: 10.1016/j.bios.2011.08.025.

Développement d'un robot autonome de nettoyage de panneaux solaires

H. FACOITI, L. LOMPO, S. SAFI et A. BOUMEZZOUGH
Laboratoire d'Innovation en Mathématiques et Applications et Technologies de l'Information
Université Sultan Moulay Slimane
Béni Mellal, Maroc
hassan.facoiti@gmail.com

***Résumé*—Dans cet article, nous nous concentrons sur le développement d'un prototype de robot autonome de nettoyage de panneaux solaires. Basé sur la carte Arduino et grâce à des capteurs, ce système électronique peut résoudre le problème lié à la poussière qui réduit les performances des centrales photovoltaïques à travers le monde. Sur la base d'une étude théorique et en comparant les différentes solutions qui pourraient résoudre le problème du projet, nous avons réussi à réaliser un produit qui répond aux besoins avec la possibilité d'une optimisation supplémentaire.**

***Index Terms*—Robot autonome, Panneau solaire, Centrale photovoltaïque, Énergie renouvelable.**

I. Introduction

De nos jours l'énergie solaire prend de plus en plus d'ampleur car étant une énergie propre et non épuisable. Beaucoup de pays africain, notamment le Maroc favorise la mise en place des structures à énergie solaire car cette énergie respecte la nature et l'environnement [1]. Mais malheureusement, malgré les efforts consentis par les gouvernements dans le but de mettre cette énergie en exergue ; l'énergie solaire rencontre toujours certaines difficultés. On peut citer entre autres le manque de moyen nécessaire pour la mise en place d'un système complet car étant cher, les changements climatiques qui influent beaucoup sur la production et aussi les saletés qui décroit énormément la production des centrales solaires. Au regard de l'importance qu'accorde les gouvernements à cette énergie appelée énergie du futur et aussi des différentes difficultés que rencontre cette énergie, nous avons développé un robot nettoyeur capable de nettoyer les panneaux solaires de façon autonome.

Ce robot sera à mesure de faire un nettoyage à sec ou à eau selon la commande qui lui sera envoyée par l'utilisateur et dispose de deux modes de fonctionnement(le fonctionnement automatique et le fonctionnement manuel). Étant la partie principale du projet, cet article a été scindé en deux parties dans le but de mieux structurer le travail.

Dans la première partie nous ferons une étude théorique de conception de façon judicieuse en tenant compte de tous les aspects. Et dans la deuxième partie nous présenterons les différentes phases de conceptions ainsi que les différents tests réalisés.

II. Étude théorique de réalisation du robot

A. Mode de fonctionnement

Le robot permet de faire le nettoyage des panneaux photovoltaïques. Pour réaliser cette fonction, l'utilisateur positionne le robot sur la plateforme de transport qui le transporte jusqu'aux chaînes des panneaux photovoltaïques. Une fois arrivé il le positionne via une commande manuelle sur les panneaux solaires. Après positionnement il choisit le mode nettoyage (à sec ou à l'eau), la vitesse de déplacement (variable) et mode fonctionnement (automatique ou manuelle) et lance ainsi le robot. Le robot une fois actionné commence le nettoyage selon le mode de fonctionnement choisi.

1) Mode manuel:

Dans cette partie le robot est contrôlé par l'utilisateur à distance. Celui−ci grâce à un dispositif peut contrôler tous les déplacements du robot tout en ayant une vue globale grâce à la caméra placée sur le robot. Il peut ainsi téléguider le robot afin d'assurer un nettoyage parfait en toute sécurité. La figure "Fig. 1" illustrant l'algorithme manuel élaboré pour le fonctionnement du robot.

2) Mode automatique:

Le robot assure le nettoyage de façon autonome tout en évitant les différents obstacles qu'il croise (Obstacle de fin de chaînes, de changement de niveau). Il se déplace de direction en direction afin d'assurer un nettoyage optimal et efficace. Il est important de noter qu'il offre toute fois une vue à l'utilisateur qui peut à tout instant décider de prendre la commande en cas de besoin "Fig. 2".

B. Étude technique des composants robot

1) Rôles des différents composants:

• Prise de tension de chargement : alimente le robot de nettoyage pour le charger dans le cas où il est déchargé.
• Régulateur de tension de chargement 16.8V : régle la tension de chargement à la valeur 16.8V qui est la tension de chargement d'un pack de batteries de type 4S (quatre piles lithium-ion en série).
• Interrupteur On/Off : allume ou éteindre le contrôleur radiofréquence tout en assurant l'option de chargement même si ce dernier est mis sous tension.
• Carte BMS : régle la charge, la charge des piles Lithium-ion et les protéger contre les courts circuits et la décharge

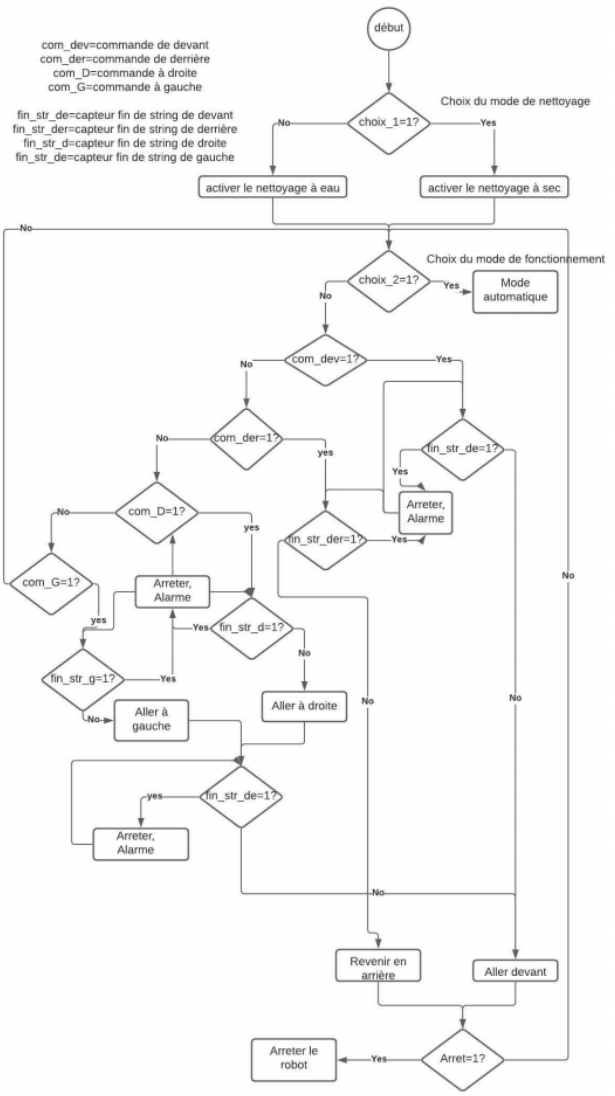


FIGURE 1. Algorithme manuel

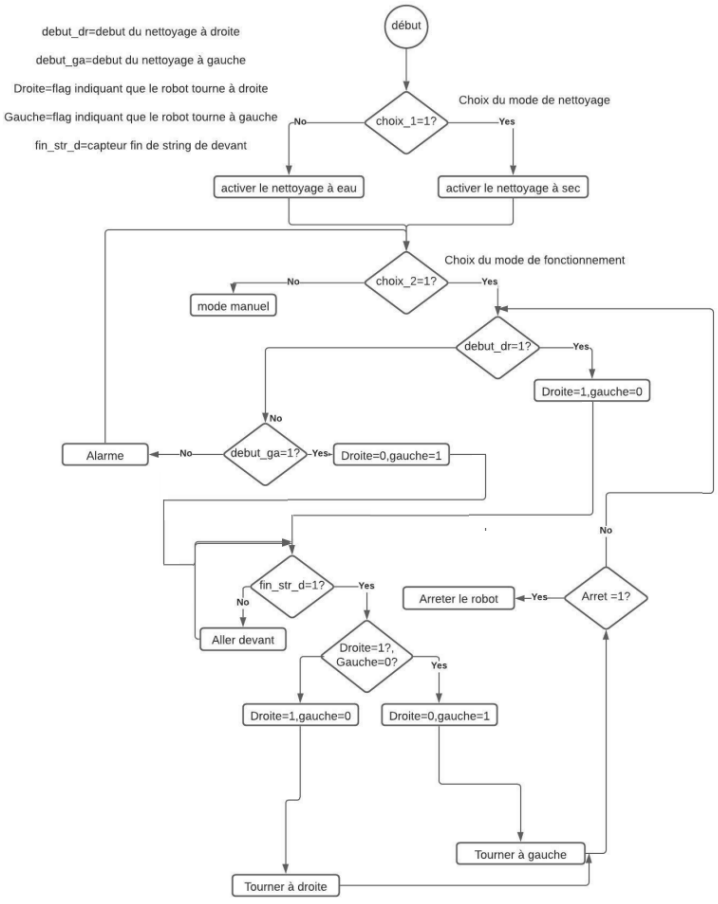


FIGURE 2. Algorithme automatique

complète.

• Régulateur de tension d'alimentation d'antenne d'émission : régle la tension d'alimentation d'antenne d'émission à la valeur nominale de 3.3V (ne peut pas utiliser la sortie 3.3V de l'Arduino parce qu'il ne peut pas donner le courant suffisant pour alimenter cette antenne en cas d'émission) [2].

• Antenne d'émission : échange les informations et les ordres avec notre robot de nettoyage.

• Condensateur de lissage : lisse la tension d'alimentation d'antenne pour stabiliser son fonctionnement et éviter les perturbations dû à l'instabilité de l'alimentation.

• Régulateur de tension de chargement 9V-12V : régle la tension de d'alimentation pour faire fonctionner les composants qui fonctionnent avec 9V-12V.

• Motor Driver : contrôle les moteurs, pour varier leur vitesse et leur sens de rotation [3].

• Moteur du robot : déplace le robot.

• Moteur de la brosse : tourne la brosse de nettoyage.

• Ventilateur de refroidissement : refroidit les composants électroniques de commandes et les batteries à cause de la température ambiante élevée dans les centrales photovoltaïques.

• Pompe : monte la pression de l'eau utilisé dans le nettoyage avec eau.

• Buzzer : informe la personne qui contrôle le robot que ce dernier est arrivé en fin de chaînes.

• Lampe : utilisée durant le nettoyage dans la nuit.

• Capteur : détecter que le robot est arrivé en fin de chaînes.

• Carte de contrôle : le cerveau principal du robot de nettoyage qui exécute les instructions venants du contrôleur radiofréquence en cas du contrôle manuel et les instructions du programme interne dans le cas du fonctionnement automatique [4].

2) Étude et choix des composants:

• Choix moteurs :

On a effectué le dimensionnement des moteurs qui vont déplacer le robot suivant la masse totale du robot avec ses deux brosses de nettoyage. La pente maximale étant de 40°, généralement au Maroc l'inclinaison maximale des panneaux photovoltaïques est de 35°. Le temps de fonctionnement

théorique de ce robot est de 2 heures s'il fonctionne dans ses conditions nominales.

Afin de rouler sur une surface horizontale, les moteurs d'un robot à roues doivent produire suffisamment de couple pour surmonter les imperfections de la surface ou des roues, ainsi que la friction dans le moteur lui-même. Théoriquement, un robot (petit ou grand) ne nécessite pas de couple important pour se déplacer sur une surface purement horizontale. Évidemment, il y aura plus de friction et de résistance dans un grand robot que dans un petit robot. Pour un robot qui roule sur un plan incliné à une vitesse constante (pas d'accélération ou de décélération), il doit produire suffisamment de couple pour contrer l'effet de la gravité. Sur une surface inclinée (à un angle θ) un seul élément de son poids (parallèle à la surface) provoque le déplacement vers le bas "Fig. 3(A)". L'autre composante, est équilibrée par la force normale à la surface exercée sur les roues.

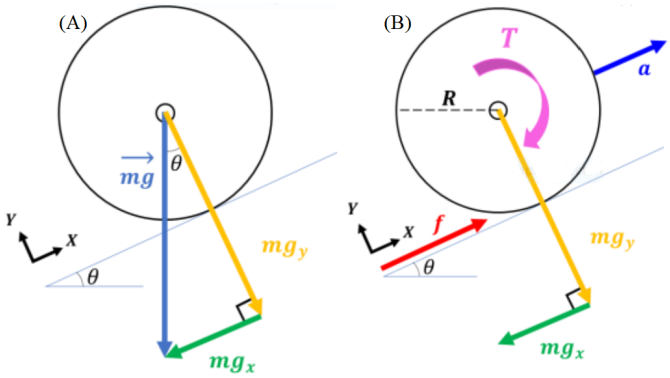


FIGURE 3. (A) : Force appliquée sur le moteur en sens incliné. (B) : Effet de la force de friction et de l'accélération sur la stabilité du robot dans un plan incliné

Afin que le robot ne glisse pas vers le bas de la pente, il faut que la friction (f) entre la roue et la surface soit suffisante "Fig. 3(B)". Le couple requis (T) est : Pour sélectionner le moteur approprié, nous devons considérer le pire des cas où le robot n'est pas seulement sur un plan incliné, mais l'accélération jusqu'à combien ?

3) Dimensionnement du moteur d'entraînement:
Pour faire le choix final des moteurs on a effectué leurs dimensionnements suivant la masse totale du robot avec ses deux brosses de nettoyage et de tous les différents paramètres. la figure "Fig. 4" montrant la procédure effectué.

4) Études de la consommation électrique des composants:
Nous avons calculé la consommation en énergie et la puissance totale des composants les plus énergivores dans le robot dans le but de bien dimensionner la partie stockage qui va assurer son autonomie pendant les heures de fonctionnement "Tab. 5".

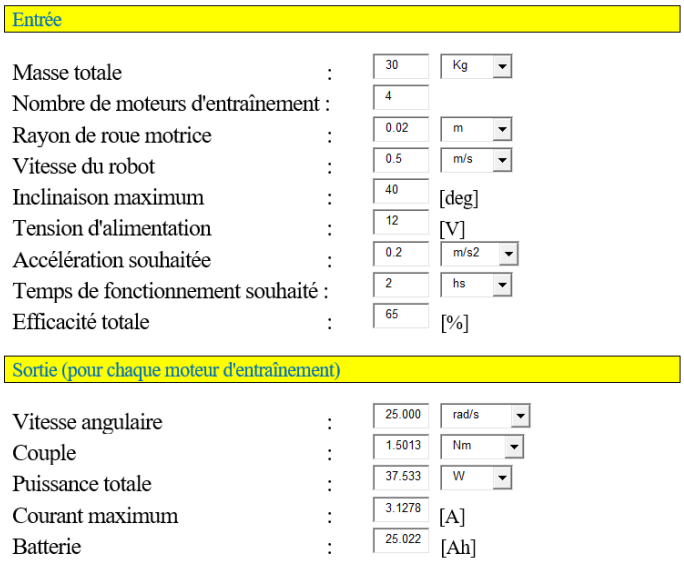


FIGURE 4. Procédure de calcul

Nombre	Appareils	Puissance (W)	Puissance totale (W)	Temps de fonctionnement	Energie journalière (Wh)
4	Moteur DC (robot)	9V*2.5A=22.5	90	2h	180
2	Moteur DC (brosse)	9V*2.5A=22.5	45	2h	90
4	Moteur Driver	12V*0.08A=0.96	3.84	2h	7.68
1	Arduino Méga	5V*0.07A=0.35	0.35	2h	0.7
1	Module NRF24L01	3.3V*0.115A=0.4	0.4	2h	0.8
4	Capteur infrarouge	5V*0.02A=0.1	0.4	2h	0.4
2	Lampe	12V*0.84A=10	20	2h	40
			P_{totale}=160w	2h	E_{totale}=320Wh/j

FIGURE 5. Tableau de consommation électrique

III. RÉALISATION DU ROBOT NETTOYEUR

A. Outils de travail

1) Logiciel de programmation ARDUINO:
Arduino est une plate-forme open source basée sur du matériel et des logiciels faciles à utiliser. Les cartes Arduino sont capables de lire les entrées lumière sur un capteur, d'activer un moteur, d'allumer une LED, de publier quelque chose en ligne. Vous pouvez dire à votre carte quoi faire en envoyant un ensemble d'instructions au microcontrôleur de la carte [5]. Pour ce faire, vous utilisez le langage de programmation Arduino (basé sur le câblage) et le logiciel Arduino (IDE).

2) Plate-forme MIT APPinventor:
MIT APPinventor "Fig. 6" est un environnement de programmation visuel et intuitif qui permet à tout le monde de créer des applications entièrement fonctionnelles pour les smart phones et tablettes Android et iOS. Ceux qui découvrent le "MIT APPinventor" peuvent avoir une première application simple et opérationnelle en moins de 30 minutes. De plus, l'outil est basé sur des blocs facilite la création d'applications complexes et à fort impact en beaucoup moins de temps que les environnements de programmation traditionnels. Le projet

"MIT APPinventor" vise à démocratiser le développement de logiciels en permettant à tous, en particulier aux jeunes, de passer de la consommation de technologies à la création de technologies.

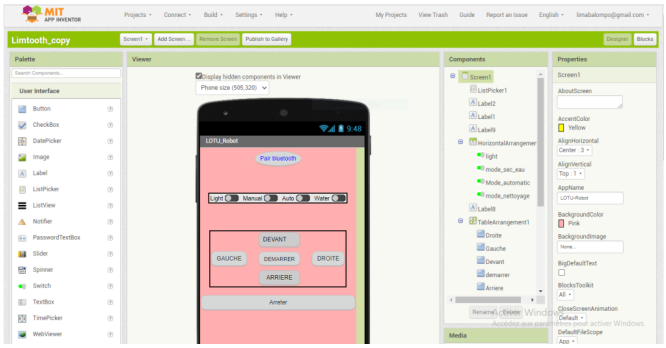


FIGURE 6. Interface de programmation MIT APPinventor

B. *modèle mécanique*

La conception mécanique du robot de nettoyage doit respecter certaines exigences mécaniques, soit du robot lui-même tel que la rigidité et la masse réduite ou bien d'autres exigences liées au panneaux photovoltaïques sur lesquels il va être placé pour le nettoyage. Le premier critère est la répartition de la masse qui ne doit pas dépasser la valeur 5400 Pa pour les panneaux photovoltaïques d'une manière constante. Cette constante est la division de la masse sur la surface du panneau qu'il ne faut pas dépasser pour garder afin de ne pas détériorer le module. Pour cette raison on a choisi un châssis "Fig. 7" avec une masse de 3.6 Kg sans prendre en compte la masse des composants mécaniques et électroniques qui va être ajoutée après.

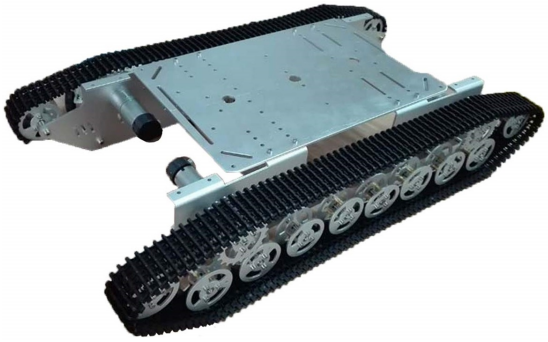


FIGURE 7. châssis du modèle

C. *Carte de programmation*

Pour la réalisation de ce projet nous avons utilisé la carte de programmation Arduino Mega qui est une carte de programmation très puissante et robuste pour la conception et le pilotage des montages électroniques, robotiques, mécaniques. L'Arduino Mega présente la configuration suivante : microprocesseur ATMega2560, 16 entrées analogiques, mémoire flash de 256 kB, mémoire SRAM de 8 kB, mémoire EEPROM de 4 kB, 54 broches d'E/S dont 14 PWM, 3 ports série "Fig. 8".

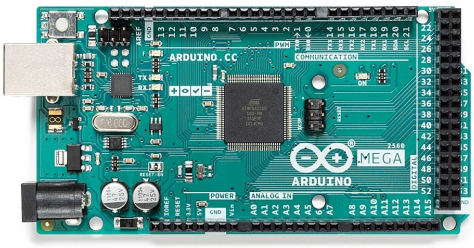


FIGURE 8. Arduino Mega

D. *Application de contrôle du robot sous MIT APPiventor*

Pour le contrôle du robot nous avons développé une application mobile sous MIT APPinventor. Cette application permet à un module bluetooth monté sur le robot de s'y connecté. Une fois le module connecté on peut contrôler les mouvements du robot à l'aide n'importe quel smart phone ayant installé l'application "Fig. 9".

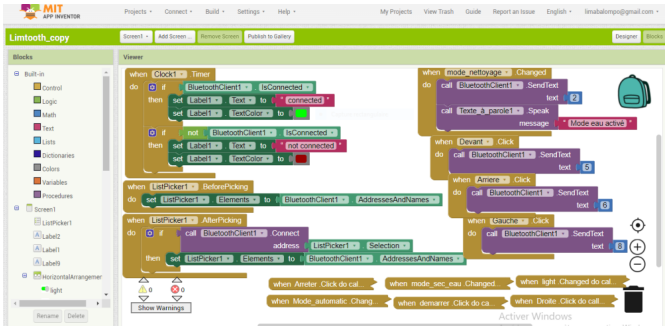


FIGURE 9. Programme sous MIT Appinventor

E. *Modèle prototype Tests de fonctionnement du robot*

Dans le but de tester le fonctionnement du robot et des différents programmes nous avons réalisé un robot de prototypage "Fig. 10". Ce robot ne possède pas les mêmes caractéristiques (électriques, contraintes physiques) que le robot réel que nous avons dimensionné mais il permet d'exécuter les différents programmes qui seront dans le robot réel afin de nous assurer du bon fonctionnement de ceux ci. Son contrôle à distance se fait via une application mobile Androïde développé sous MIT APPinventor. Ce prototype nous a permis aussi de de vérifier la détection les extrémités des Strings des panneaux photovoltaïques grâce à ces quatre capteurs d'obstacles montés sur lui. Afin que le robot réponde aux mieux les exigences fixés, les capteurs ont étés placés sur les chaque coté de celui pour éviter les obstacles de par et d'autre.

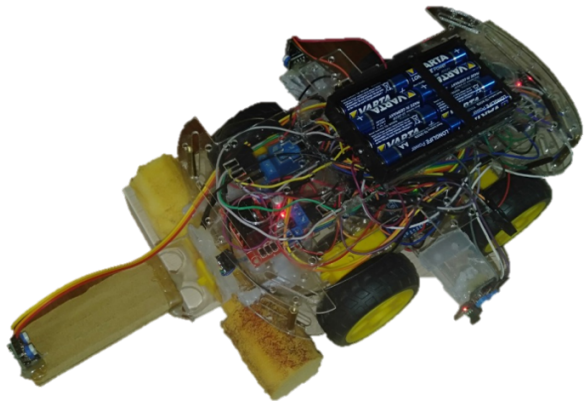


FIGURE 10. Modèle de prototype Tests

F. Interface de connexion de l'application

L'application dénommée LOTURobot possède une interface utilisateur qui permet à un utilisateur de contrôler et d'interagir avec le robot. Cette application est composée de "Fig. 11" :

- Pair Bluetooth : permet des rechercher les modules bluetooth activés, notre cas le module bluetooth a pour nom du '98 :D3 :91 :FD :A2 :C4 : HC-05'.
- LIGHT : permet d'activer et de désactiver la lumière pour les nettoyages nocturnes.
- Manuel : permet d'activer et de désactiver le mode manuel.
- Auto : permet d'activer et de désactiver le mode automatique.
- Water : permet d'activer et de désactiver le mode de nettoyage à eau.
- DEMARRER : permet de démarrer la broche de nettoyage.
- DROITE, GAUCHE, DEVANT, ARRIÈRE : permettent respectivement d'aller à droite, gauche, devant et en arrière.
- Arrêter : permet d'arrêter le robot.

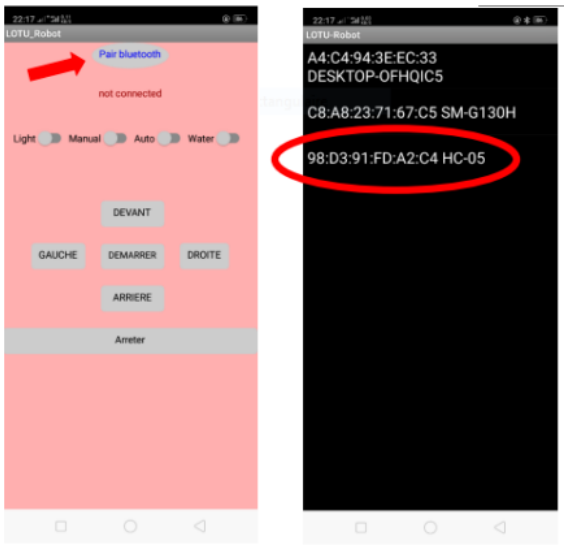


FIGURE 11. Interface utilisateur

G. Tests de fonctionnement

1) Test manuel:

Pour tester le fonctionnement manuel du robot on l'a mis sur un panneau avec quelques bouts de papiers qui ont servis d'ordures. Ainsi grâce à sa technologie de capteurs le robot circule sous l'ordre d'un utilisateur tout en éliminant les ordures placés sur celle-ci et en détectant ses bordures. Une fois qu'une bordure est détectée le robot s'arrête automatiquement et refuse toute action de dépassement jusqu'à ce que l'utilisateur change de commande.

2) Test manuel:

Le test automatique a été réalisé sous les mêmes conditions que le test manuel. Positionné sur le panneau, le robot circule de façon autonome sans aucune intervention externe. A chaque détection de fin de chaînes, le robot recule et change de chemin et continue dans ce chemin jusqu'à une nouvelle détection de fin de string.

IV. RÉSULTATS

A. Résultats positifs

Après différents tests effectués, on observe que les capteurs détectent parfaitement les extrémités des panneaux photovoltaïques. Le robot exécute les instructions mises dans le programme sans aucuns problèmes, et peut nettoyer une grande surface dans une durée réduite. Lorsque il arrive en fin des strings il déclenche l'alarme pour avertir la personne qui le contrôle et se stop. Le robot peut changer le mode de nettoyage à sec vers le nettoyage avec d'eau sans problèmes.

B. Résultats négatifs

Les capteurs infrarouges ne peuvent pas être utilisés dans le robot réel sur les panneaux photovoltaïques car ces derniers ne détectent pas les surfaces de couleur noir. On peut les remplacer par des capteurs à ultrasons car ils sont plus performants et le délai de calcul de distance est très rapide [5].

V. CONCLUSION

Ce travail nous ouvre la porte sur l'avenir des énergies renouvelables, notamment les installations photovoltaïques qui aurons besoins d'un nettoyage constant pour assurer la stabilité et la continuité de leur niveau de production durant leurs années d'exploitation. Il ouvre également les portes à la recherche.

Pour améliorer le robot doit se concentrer pour résoudre le problème de détection des extrémités des chaînes photovoltaïques, nous pouvons changer les capteurs utilisés par d'autres types (capteurs TOR). Changez le type d'antenne utilisée pour la télécommande. Optimisez le programme pour augmenter la vitesse et le temps de réponse du robot. Et, ajoutez d'autres options pour ce robot : analyse du profil de nettoyage pour optimiser la procédure de nettoyage, nettoyage programmé et analyse thermique de l'installation photovoltaïque avec une caméra intégrée pour détecter les points chauds.

RÉFÉRENCES

[1] Anne LABOURET et Michel VILLOZ. "Energie solaire photovoltaïque". Malakoff, France : Dunod, 2006.

[2] M.Carrobles Maeso, F.Rodriguez Garcia, A.Martin Hernandez. "Automates et Robotique III". Manuel de mécanique industrielle, Edition 2004.

[3] Simon Landrault et Hippolyte Weisslinger (olyte). "PREMIERS PAS EN INFORMATIQUE EMBARQUÉE" . livre, Édition du 19 février 2016.

[4] Mohamed ZAYED. "Véhicules Intelligents : Étude et développement d'un capteur intelligent de vision pour l'attelage virtuel", Thèse de Doctorat, Université des Sciences et Technologies de Lille, 12 juillet 2005.

[5] Hassan FACOITI, Said SAFI, Ahmed BOUMEZZOUGH. "Analyse, caracterisation et conception d'un systeme radar anticollision embarqué dans les véhicules. sciencesconf.org :icsat-2020 :326086, pp. 13-18, International Colloquium on Signal, Automatic control and Telecommunications, Caen, France, 10-12 juin (2020).

[6] Charles J.Jacobus, Douglas Haanpaa. "All weather autonomously driven vehicles". U.S. Patent No. 9,989,967. 5 Jun (2018).

[7] HORAUD Radu, MONGA Olivier. "Vision par ordinateur : outils fondamentaux". BOOK-Second Edition (1995).

[8] Félix INGRAND. "Architectures logicielles pour la robotique autonome". JNRR, 2003, vol. 3, p. 8-10.

[9] Thomas MAMBRINI. "Caractérisation de panneaux solaires photovoltaïques en conditions réelles d'implantation et en fonction des différentes technologies". 2014. Thèse de doctorat. Paris 11.

[10] H. FACOITI, S. SAFI, A. BOUMEZZOUGH."Comparative study between computer vision methods for the estimation and detection of the roadway," IEEE Xplore, 2021, DOI : 10.1109ICOA51614.2021.9442633

Energy Efficiency of a Public Lighting Network: from Diagnosis to Improvement of Electrical Performances

B.Boukhris[1], L.Elmahni[2], F.Douslimane[2]

[1]LASIME, ENSA, Ibn Zohr university, Agadir, Morocco

[2]RMEL, Faculty of sciences, Ibn Zohr university, Agadir, Morocco

I Introduction

Public lighting (PL) is a major expense in the operating budget of local authorities. In this context we seek to upgrade their public lighting network through the following objectives: controlling the existing public lighting system; modernize the public lighting network, improve the quality of service provided to citizens; rationalize energy consumption, reduce budget expenditure related to public lighting; acquire the necessary studies and diagnoses of its public lighting network and finally optimize the mode of operation and maintenance of the public lighting network.

This research includes the technical inventory of all the components of the public lighting network, the technical evaluation (photometric and electrical) of these elements, Indeed, the old ferromagnetic type power supply equipment and control systems, can be advantageously replaced by electronic ballasts which offer additional functionalities, such as automatic voltage regulation and dimming. Autonomous or centralized control systems make it possible to control ignition times and modulate illuminance values according to the needs defined by the public authority. Moreover, all these improvement solutions are the objective of an energy and technical diagnostic project of the public lighting network, in order to reduce energy consumption, improve quality and safety as well as optimize investment and operating costs, so we are talking about improving energy efficiency.

II Description of the Project

For carrying out a diagnostic project for an existing public lighting network; we must inventory and collect global data of a public lighting network subject of study [1].

2-1 Technical data of the PE network

Types	Main Features
Types of networks:	% Underground network % Twisted aerial network on facade % Twisted aerial network on concrete/wooden pole % Soft Aerial – bare copper % Overhead-underground network
Type of cables:	Copper: supplies a number of light points Aluminum: supplies a number of light points
Order types:	Astronomical clock mechanical clock
support Types:	% Of supports are candelabra

	% Polls % In front % (in number) non-existent supports % Without skirt
support material	% Hell % Concrete posts
Height of light points:	Number of 10m light points Number of 12m light points
Types of lamps:	High pressure sodium Low pressure sodium Fluo-compact Incandescent LEDs Metal halide Halogen mercury vapor
Power of lamps:	70W: 250W: 150W: 400W: 1000W: 2000W:
Types of ballast:	Electronic ballasts

2.2 Expenses relating to public lighting:

We must at least have the different operating budgets in (MAD) and investment (MAD) for at least 3 years which will be concerned by the study in MAD including tax. Similarly, the electricity consumption of the various items billed for the years concerned by the study in KWh and MAD including tax.

III Inventory and technical diagnosis of the public lighting network

The technical inventory is a census operation of all the elements of a PE network with an assessment of their states and their geo-referencing using a geographical information system [2]. At the end of this operation, a database is created. The exploitation of this database is multiple. It makes it possible to know the state of each element of the PE network, serves as a basis for establishing preventive maintenance and constitutes a reference for the planning of any renovation or modernization project of the PE network.

The aggregate data of all cabinets and light points is filed in the database in two ways:

- A file containing all the data in a tree structure: each cabinet followed by the corresponding light sources.
- A file containing several sheets in the number of items. Each sheet contains the data of a substation and its corresponding light sources.

This database will constitute the public lighting layer of the geographical information system (GIS) of the site in question.

IV Analysis of measurements in a conventional public lighting station

Before starting the study and sizing of the proposed solution for public lighting using solar PV, we are going to present the current state of one of the public lighting stations of a study site. In fact, the study concerns the state of one of the conventional power supply stations for public lighting. First, some electrical quantities characterizing this station such as voltage, current, power factor and fundamental frequency [3]. Second, the parameters identifying the quality of the post-conventional electrical network, such as the various THD harmonic distortion rates (Total Harmonic Distortion), the spectrum of voltage and current harmonics and their compatibility levels and finally the electrical consumption involved in this conventional substation.

In order to perform the measurements presented below, we used a "CIR-e3" type network analyzer [4], the aim being to check compliance or not with standards EN50160 and CEI61000-2-2concerning the HVn and HIn harmonic levels. Using this measuring device, we recorded the evolution of the various electrical parameters and those characterizing the quality of the electrical network for a period of 24 hours: Phase voltage U,Current of the phases I

Power factor, Network frequency:

Parameters identifying the quality of the electrical network

According to the measurement results we find that the total harmonic distortion rate THDf with respect to the voltage fundamental VTHD does not exceed 3%, therefore is less than 5%, which is acceptable according to standard EN50160 [3].

Therefore, in the case of this network it is non-polluting in voltage, which is not the case in current. The THDf of the ITHD current is around 13.8% during the night, which exceeds the 10% imposed by the standard. To minimize this ITHD, we must implement passive anti-harmonic filters [5] [6].

Spectrum of voltage and current harmonics

In the case of public lighting, the lamps used constitute a non-linear load which then corresponds to a distorted current call, therefore rich in harmonics. [7]

V Photometric analysis of the public lighting network

The photometric analysis of the public lighting network makes it possible to conclude on the performance of the network in terms of average minimum illumination and the uniformity to be maintained for each zone, in order to guarantee the minimum lighting values required for users.

To assess the quality of the service rendered, namely the quality of the lighting, standard simulations were carried out using the DIALUX ECLAIRAGE PUBLIC software [8],

although the results and the resulting conclusions remain approximate given the previous considerations.

5.1 Photometric simulation of light points at the start of a cabinet supplied from a transformer station:

Simulation Data:

- Number of carriageways: 2- Number of sidewalks: 2- Pavement width: 7m- Sidewalk width: 3m- Average central reservation width: 0.5m- Height of the light point: 12m- Arrangement of light points: bilateral facing each other- Number of lights per post: 1- Average distance between 2 light points: 39m- Lamp power: 250W SHP Philips- Luminous flux of the luminaire: 19710 lm

Simulation Results:

- Chaussée 1 et 2 :

Classe d'éclairage choisie: ME4a (Toutes les exigences photométriques ne sont pas remplies.)

	L_{moy} [cd/m²]	U0	Ul	TI [%]	SR
Valeur effective selon calcul:	0.70	0.69	0.83	52	0.96
Valeurs de consigne selon la classe:	≥ 0.75	≥ 0.40	≥ 0.60	≤ 15	≥ 0.50
Rempli/Non rempli:	✗	✓	✓	✗	✓

- Trottoir 1 et 2 :

Classe d'éclairage choisie: CE5 (Toutes les exigences photométriques ne sont pas remplies.)

	E_{moy} [lx]	U0
Valeur effective selon calcul:	7.48	0.60
Valeurs de consigne selon la classe:	≥ 7.50	≥ 0.40
Rempli/Non rempli:	✗	✓

We can notice that the photometric performances calculated for this avenue do not meet, but remain close to, the requirements of the lighting class chosen by Dialux. [9]

5.2 Photometric simulation of the light points of a cabinet feeder powered from another transformer substation:

Simulation Data:

- Number of roads: 1- Number of sidewalks: 2- Pavement width: 13m- Sidewalk width: 3m
- Average central reservation width: 0m- Height of the light point: 12m- Arrangement of light points: unilateral- Number of lights per post: 1- Average distance between 2 light points: 41m
- Lamp power: 250W SHP Philips- Luminaire luminous flux: 19710lm

Simulation results:

- Chaussée :

Classe d'éclairage choisie: ME4a (Toutes les exigences photométriques ne sont pas remplies.)

	L_{moy} [cd/m²]	U0	UI	TI [%]	SR
Valeur effective selon calcul:	0.29	0.31	0.67	58	0.93
Valeurs de consigne selon la classe:	≥ 0.75	≥ 0.40	≥ 0.60	≤ 15	≥ 0.50
Rempli/Non rempli:	✗	✗	✓	✗	✓

- Trottoir 1 :

Classe d'éclairage choisie: CE5 (Toutes les exigences photométriques ne sont pas remplies.)

	E_{moy} [lx]	U0
Valeur effective selon calcul:	2.03	0.51
Valeurs de consigne selon la classe:	≥ 7.50	≥ 0.40
Rempli/Non rempli:	✗	✓

- Trottoir 2 :

Classe d'éclairage choisie: CE5 (Toutes les exigences photométriques ne sont pas remplies.)

	E_{moy} [lx]	U0
Valeur effective selon calcul:	5.14	0.60
Valeurs de consigne selon la classe:	≥ 7.50	≥ 0.40
Rempli/Non rempli:	✗	✓

Results interpretation:

It is an avenue lit following a unilateral arrangement by luminaires whose average distance is 41m. Theoretically, the recommendations stipulate that the height "h" of the fireplaces depends on the width "l" of the roadway, is.:

Type of location	Unilateral; Axial; Bilateral retro; quincunx	Bilateral opposite
Fire height "h"	h≥1	h≥1/2

In our case, we have a unilateral arrangement, the height must be at least equal to 13m (here h=12 m). Also, the inter-distance must be::

Type of location		Unilateral; Axial; retro bilateral; bilateral	quincunx
Spacing e	Luminaire equipped with balloon or SBH	e=3h	e≤2.7h
	Luminaire equipped with SHP tubular lamp	3.5h≤e≤4h	e≤3.2h

e_{min}=3.5*12=42m. The average location measured being 41m.

Upgrading work on many control cabinets must be scheduled. Based on the results of this diagnosis, giving priority to priority actions related to the safety of people and installations, then actions related to the quality of lighting. The budget estimate for the entire project then depends on the nature and number of improvement actions.

VI energy diagnosis

6.1 Energy consumption indicators:

Overall consumption: We give the overall consumption of all contracts in kWh and its equivalent in MAD for the years of study and the evolution of its energy consumption.

Overall specific consumption:This is consumption per inhabitant. This ratio reflects the over dimensioning of energy consumption in relation to the real needs of citizens. Lighting uninhabited areas, for example, can have a negative impact on the result.

In the absence of a national average in Morocco, and as an example the national average in France is 76 kWh/inhabitant and 6.3 euros/inhabitant. Global kWh price: The global kWh price thus calculated appears reasonable.

6.2 Analysis of photometric simulations

Photometric simulations of the light points were carried out: case of one luminaire per mast. The simulation concerned the existing so-called reference (SHP 250w sources) and the projected case called action (77w LED) producing the same level of illumination required. The detail of the simulations is

Simulation results:

➢ ***Chaussée 1 et 2 :***

1. **Résultats de simulation par une lampe SHP 276W de Philips dont le Flux lumineux est de 19710lm :**

Classe d'éclairage choisie: ME4a (Toutes les exigences photométriques ne sont pas remplies.)

	L_{moy} [cd/m²]	U0	UI	TI [%]	SR
Valeur effective selon calcul:	0.70	0.69	0.83	52	0.96
Valeurs de consigne selon la classe:	≥ 0.75	≥ 0.40	≥ 0.60	≤ 15	≥ 0.50
Rempli/Non rempli:	✗	✓	✓	✗	✓

2. **Résultats de simulation par une lampe LED 77.5W de Philips dont le Flux lumineux est de 10140lm :**

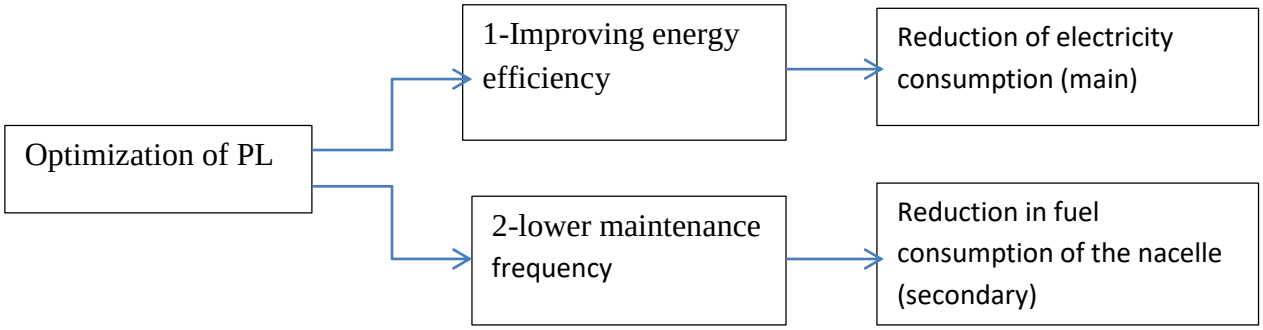
Classe d'éclairage choisie: ME4a (Toutes les exigences photométriques sont remplies.)

	L_{moy} [cd/m²]	U0	UI	TI [%]	SR
Valeur effective selon calcul:	0.85	0.84	0.78	9	0.90
Valeurs de consigne selon la classe:	≥ 0.75	≥ 0.40	≥ 0.60	≤ 15	≥ 0.50
Rempli/Non rempli:	✓	✓	✓	✓	✓

The 77w LED light source provides average illuminance higher than that of the 250w SHP light source, and allows both to meet photometric requirements (minimum illuminance required) and reduce energy consumption by approximately 1/3.

The evaluation of the gains expected from the operation of replacing all the luminaires equipped with 250w HPS by luminaires equipped with 77w LEDs is given below:

Step 1: define the purpose of the evaluation
Estimate the financial gains generated by the action
Step 2: define the action
Replacement of conventional luminaires with LED luminaires
Step 3: build the action consequence tree



Description of each consequence	
Consequence 1	The replacement of conventional luminaires with LED luminaires leads to a reduction in electricity consumption.
Consequence 2	LED luminaires require less frequent and faster maintenance: the nacelle used for maintenance will therefore consume less fuel.
Step 3: define the scope of the action	
The party concerned	Classic luminaires equipped with SHP 250w
Observation time	15 years
Step 4 Gather activity data	
Activity data	Observation time
	Total number of LED luminaires commissioned
	Duration of implementation
	% Of luminaires implemented during the 1st third of the implementation period
	% Of luminaires implemented during the 2nd third of the implementation period
	% Of luminaires implemented during the 3rd third of the implementation period
	Average power of LED luminaires (excluding dimming)
	Average power of the reference luminaires
	Equipment consumption
	Annual duration of use
	Number of gas-discharge luminaires that would have been replaced
	Number of gas-discharge luminaires that would have had to be maintained (lamp change every 4 years)
	Lifetime of an LED luminaire
	Lifetime of a classic luminaire
	Lifetime of a bulb
	Number of cleaning interventions per point each year
	Number of replacement interventions per point each year
	Number of maintenance interventions per point each year
	Time required to clean a luminaire
	Time required for the systematic maintenance of a luminaire (lamp change + cleaning)
Consequence 1 calculations: Reduction of electricity consumption	
Activity data	Total number of LED luminaires commissioned
	Duration of implementation
	% of luminaires implemented during the 1st third of the implementation period
	% of luminaires implemented during the 2nd third of the implementation period
	% of luminaires implemented during the 3rd third of the implementation period
	Average power of LED luminaires (excluding dimming)
	Average power of the reference luminaires
	Equipment consumption
	Annual duration of use
	Observation time
Data to calculate	Energy saving per luminaire replaced
	Financial gain per luminaire replaced
	Energy savings for replaced luminaires
	Financial gain for replaced luminaires
	Energy gain for luminaires replaced during the 1st period
	Financial gain for luminaires replaced during the 1st period

	Energy gain for luminaires replaced during the 2nd period
	Financial gain for luminaires replaced during the 2nd period
	Energy gain for luminaires replaced during the 3rd period
	Financial gain for luminaires replaced during the 3rd period
	Financial gain for luminaires replaced for 15 years
Stock	Replacement of luminaires from the first period: Investment and annual gain
	Replacement of the luminaires of the 2nd period: Investment and annual gain
	Replacement of 3rd period luminaires: Investment and annual gain
Consequence calculations 2: Reduction of fuel consumption by the nacelle	
Activity data	total number of LED luminaires commissioned
	Number of gas-discharge luminaires that would have had to be maintained (bulb change every 4 years)
	Number of cleaning interventions per point each year
	Time required to clean a luminaire
	Number of replacement interventions per point each year
	Time required for the systematic maintenance of 1 luminaire (change of lamp + cleaning)
	Observation time
Gain	Gain in distance traveled per day
	Saving in fuel consumption per day
	Financial gain per day
	Savings in fuel consumption per year
	Financial gain per year

6.3Installation of photovoltaic panels:

The installation of an autonomous solar street lamp with a power of 33W LED equivalent to a 150W SHP lamp.

6.4 Use of power optimizers:

The power optimizers in public lighting are devices, which can be placed at the level of the control cabinets or at the level of the luminaires (dimmable ballasts). The principle of power optimizers is based on lowering the voltage, which leads to a reduction in the power consumed and subsequently to a reduction in the luminous flux of the source. The energy savings achieved, in public lighting, are 20% to 50% [10].

6.5 Changing ferromagnetic ballasts to electronic ballasts: In a classic installation with ferromagnetic ballasts:

- 1% overvoltage results in 3% overconsumption;
- Permanent operation at a 7% overvoltage can reduce the life of lamps and equipment by 50%.

In addition, given the risks of losses and disturbances on the network, it is necessary to check the good condition of the capacitors.

It is therefore recommended to use electronic ballasts which can achieve energy savings of up to 10 to 15% and extend the life of the lamps by controlling overvoltage.

The electronic ballast offers the advantage of ensuring the automatic cut-off of the power supply When the lamp is defective. It also guarantees the stabilization and regulation of the voltage and thus contributes to better maintenance of color temperatures over time.

6.6. Use of astronomical clocks:

The use of astronomical clocks allows an energy saving of 5%

6.7. Changing lamps at the end of their lifespan:

A lamp at the end of its lifespan consumes more than its nominal power, and lights less. If the maintenance was preventive, the exact date of commissioning of the lamps is known. Based on the table of average lamp lifetimes, a systematic check is carried out and the replacement of lamps at the end of their lifetime is automatic.

6.8. Minimization of joule losses in conductors:

In the absence of precise figures from measurements on the network, the calculation of the joule losses of a network on the basis of assumptions and data will be considered close to reality, the resulting conclusions are approximate.

6.9. Phase balancing:

Among the great advantages of a balanced three-phase system is the absence of neutral current, hence zero losses in the neutral.

According to the data from the measurement campaign, many cases of imbalances have been observed. As the phase balancing operation is relatively tedious in comparison with the gains to be generated (around 3% energy saving), it is recommended to:

6.10. Installation of capacitors:

Joule losses quadruple in the absence of capacitors, in fact the current is inversely proportional to the cosα: $P = 3UI\cos\alpha$ and the losses are proportional to the square of the current [$Pj = RI^2$] for an uncompensated discharge lamp $\cos\alpha=0.5$ on average , with cosα compensation approaches 1, so the current doubles and consequently the losses

VII. Conclusion

The calculation of the overall energy efficiency indicators showed that the typical EP network does not present any major problems in terms of electricity consumption. Nevertheless, some anomalies are raised for some transformer stations. Potential energy savings are addressed, and mainly relate to the change in light source technology: the replacement of existing sources by LED light sources makes it possible both to meet photometric requirements (minimum illumination required) and to reduce energy consumption by around 1/3. Also gains in operating budget will be made since the maintenance of LED luminaires costs much less than that of conventional luminaires. The overall investment for this change being significant, the return-on-investment time is approximately 5 years, for an average lifespan of the equipment of 15 years.

Other actions are described to improve energy efficiency (the change of ferromagnetic ballasts by electronic ones...). It is the realization of these consolidated actions that makes it possible to have a significant energy saving. Nevertheless, it is necessary to make the compromise between the investment cost and the expected gains. Energy efficiency gains can be generated indirectly by bringing electricity grid elements into compliance

References

[1] Norme européenne EN 13201, "Critères photométriques Réception et diagnostic des installations", Janvier 2009 AFE-CIE France, 22 Janvier 2009 à l'Espace Hamelin, http://www.afe-eclairage.fr/docs/10080-ext.pdf.

[2] La norme NF EN 13201, "Métrologie en éclairage public : Etat des lieux et prospectives", Colloque.

[3] European Standard: CENELEC, EN50160, "Voltage characteristics of electricity supplied by public electricity networks", July 2010

[4] Circutor, http://circutor.com/en/products/measurement-and-control/portable-power-analyzers/ portable-power-analyzers-port/m85020-detail#documentation, "Portable analyzer for energy audits", M85020, Site web : http://circutor.com/docs/FT_CIR-e3_FR.pdf.

[5] Dr. S.K. Purushothaman, "Design and analysis of elimination of harmonics using wind energy conversion systems", Journal of Theoretical and Applied Information Technology, (JATIT), Vol. 61, n. 3, 2014.

[6] A.Y. Kadri, A. Hamidat, "Compensation partielle et globale du réseau électrique BT à l'aide de l'énergie solaire photovoltaïque", Revue des Energies Renouvelables, Vol. 12, n. 3, pp. 433-447,2009.

[7] Aleksandar Nikolic, Dragana Naumovic-Vukovic, Slobodan Skundric, Dragan Kovacevic, Vladeta Milenkovic, "Methods for power quality analysis according to EN 50160", 9th International Conference on Electrical Power Quality and Utilisation, EPQU'2007, Barcelona, 9-11 October 2007

[8] Software: DiaLux 4.12, Site web: www.dial.de

[9] La norme NF EN 13201, "Métrologie en éclairage public : Etat des lieux et prospectives", Colloque AFE-CIE, 22 Janvier 2009 à l'Espace Hamelin, France, http://www.afe-eclairage.fr/docs/10080-ext.pdf

[10] F. Outferdine, L. Bouhouch, M. Kourchi, & al. Feasibility of substitution of the conventional street lighting installation by the photovoltaic case study on a municipality in Agadir in Morocco. International Journal of Electrical and Computer Engineering, 2017, vol. 7, no 5, p. 2287

A review of Radio Frequency IDentification technologie to combat COVID-19

1st El hassania Rouan
University Sultan Moulay Slimane, Faculty of Sciences and Technics
Beni-Mellal, Morocco
rouanelhassania@gmail.com

2nd Said Safi
University Sultan Moulay Slimane, Faculty Poly-disciplinary
Beni-Mellal, Morocco
mail

3rd Ahmed Boumezzough
University Sultan Moulay Slimane, Faculty Poly-disciplinary
Beni-Mellal, Morocco
mail

***Abstract*—As a result of the rapid spread of the novel coronavirus disease 2019 (COVID-19), the world has experienced an unprecedented pandemic that directly affected public health. In order to prevent or slow down the outbreak of this disease, most countries impose strict curfews and emergencies that have a negative impact on human society and global economy. Protection methods, early detection/diagnosis of symptoms, isolating the infected, and tracing of contacts was the most effective methods to face the COVID-19 disease, but these methods are difficult to apply using current technologies and materials. This paper explores and reviews the importance of using Radio Frequency IDentification (RFID) technology to respond to the challenges posed by COVID-19.**

***Index Terms*—COVID-19, Healthcare, Internet Of Things, Radio Frequency IDentification, World Health Organization.**

I. Introduction

Since December 2019, a new type of coronavirus called COVID-19 has been first identified in China. Then COVID-19 spread rapidly in whole China and the world. In March 2020, COVID-19 was declared a global pandemic virus by the World Health Organization (WHO) [1], [2].

COVID-19 is an infectious disease of humans that in severe cases causes acute respiratory distress syndrome. The presence of this disease is manifested in infected people with many symptoms ranging from asymptomatic / mild symptomatic to severe illness and death. The most common symptoms are fever, cough, and difficulty of breath [2].

The main challenges of COVID-19 is rapid person-to-person transmission or respiratory droplets and contact with contaminated objects during the incubation period, which can be up to more than two weeks, or in the case of a recessive infection [3]. Clinical diagnosis of COVID-19 is based on clinical, manifestations, molecular diagnostics of the viral genome by RT-PCR, chest x-ray or CT scan, and serology blood tests [2].

The main reason for the spread coronavirus 2019 is that clinical diagnosis are not available in all private and public places with large number of people such companies, stores, universities,...

Due to lack or limited diagnostic materials, the most institutions rely on high body temperature or fever symptoms to restrict people movement and social distancing.

In order to check the body temperature or fever, thermal screening are used to identify the probably infected people such as thermal camera, infrared thermometers,...

To control the outbreak of COVID-19, a new non-pharmaceutic solution appeared which is social distancing. social distancing means the physical distance between people that must be maintained to protect against the spreading of infectious diseases. This safety distance is set at approximately two metres [3], [4].

These preventions methods aim to limit or slow down the spread of infectious viruses, but they are still not really effective for two mains reasons:

- Spends time to check the body temperature from every person,
- Social distancing might not be respected during the screening process resulting in the spread of the coronavirus, from infected people to the person performing the screening process and other people.

In order to answer to these above flaws, an alternative and Efficient technology is needed in healthcare to minimize the effect of the pandemic during peak hours when public places are crowded and it is difficult to practice social distancing in a limited space.

The Internet of Things (IoT) is attaining increasing attention and is progressively being used in healthcare to predict, stop and control severe diseases. IoT is a network of interconnected devices, objects, or people with unique identifiers that are attached with sensors and software to collect and transfer data with least human intervention.

The most frequently IoT healthcare systems used to provide solutions to the challenges posed by coronavirus disease are Radio Frequency IDentification (RFID), Wireless Sensor Network (WSN), Wireless Body Area Network (WBAN), Wireless Fidelity (WiFi), Global Positioning System (GPS)

and wearable devices.

This paper provides a survey of key IoT based RFID solutions that have had an impact on COVID-19 in healthcare. RFID can be used in healthcare to improve methods of personal/public protective, detection of infected people and contact tracing.

The remainder of this paper is organized as follows: Section II provides an overview on RFID system including its main components and applications. Section III provides a review of studies on different proposed approaches for fighting COVID-19. Section IV presents some conclusions.

II. An overview on RFID sytem

RFID is a contactless technology that uses radio waves to uniquely identify and track tagged animate or inanimate objects. RFID has many applications in various fields due to its benefits and features such as healthcare, object tracking, smart parking systems, etc. [5]

There are three main components in the RFID system, which are Reader(s), Tag(s) and RFID middleware. The RFID reader is transmitter/reciever device that is used for identify tagged objects within a given radio frequency range. RFID tag is a small device equipped with microchip and antenna that is used to transmit data to the reader.

In an RFID system, the readers emit continuously radio signals. When tags detect those specific signals, they use the energy issued from the transmitted reader signal to activate the microchip and generate their response. The obtained data is transferred to a smart part called middleware which is in charge of controlling readers, managing their exchanged data and aggregating it to the dedicated applications [5].

Typically, RFID tags can be categorized into three types according to their power source: passives, semi-passives, and actives.

An active tag is powered by an internal battery that powers the integrated circuits (such as sensors) and the transmitter. This kind of tag can broadcast continuously a radio frequency signal even if the reader has not called it up. Active tag is often used for long read-range applications up to 100 m. A passive tag has not a battery assisted, it rely only on reader's field to operate the microchip and generate its own response. Passive tag is used for short read-range applications less than 1.5 m. A semi-passive tag has an embedded power source but is only used to aliment the integrated circuits, e.g. to store periodic temperature information from an integrated temperature sensor. Moreover, the semi-passive tag acts also as passive tag for collecting the radio waves power from the RFID reader to transmit the data [5].

III. Literature Review

In this section, we have presented, analysed, and discussed survey reviews relevant to proposed approaches for facing the coronavirus disease outbreak focused on RFID-based solutions. The most RFID-inspired solutions to contain COVID-19 disease during the pandemic are detecting infected people, maintaining social safety distances and contact tracing.

A. RFID and protection

Due to the outbreak of COVID-19 pandemic, wearing Filtering Facepiece Respirators (FFRs) in crowded places can significantly reduce the risk of infection caused by this disease because the FFRs avoid frequently touching face with unwashed hands, but FFRs humidity can reduce their effectiveness [6]. The World Health Organization advises to replace the face masks when they become wet.

The authors of [7] have developed a new kind of wearable personal device which is sensorized face masks using the FFRs integrated with RFID-based sensors in order to monitor the wetness of face masks through controlling the moisture emitted during breathing.

Another solution proposed by The World Health Organization to prevent or slowdown the breakout of coronavirus disease is social distancing. But this social distancing is violating usually in crowded places.

The RFID technology can be used in social distancing to control the number of people inside a restricted area, for example, a supermarket or workplace [4].

An RFID system can be installed inside an indoor building to maintain a certain number of people that is allow a non-violating social distancing measures. Here in this technique, each person inside that building must carres a RFID tag that is characterized by an Unique IDentifier (UID). The RFID reader installed in that building scans RFID cards allotted to the registered users and receives their UID . Based on difference between the arrival and the departure of each person, the system calculate the number of people present within the building by increasing or decreasing the counter value. If the total number of people is exceeded, the system can notify the admin to take appropriate actions either to force people to queue before entering or to leave the place [3].

B. RFID and screening process

Most patients with coronavirus disease have an elevated body temperature, which is considered the most common warning sign and basis for pandemic prevention and detection. Infrared body temperature sensor or thermo-cameras monitoring system are direct tools that need direct line of sight or close distance conditions to extract the correct measurements, but these tools may lead to the serious outbreak of COVID-19. RFID sensors have already demonstrated great potential to counteract the problems caused by conventional body temperature measurement equipment by enabling remote, unsupervised and instantaneous monitoring of this vital sign [9].

RFID sensors can be a wearable smart thermometers applied directly over the skin in order to record continuously or periodically of body temperature. The epidermal RFID temperature sensors can be accessed when users are monitored by mobile or fixed remote reading system during the crossing of surveillance gate [8]. The access gate must be installed at certain distance within the reading range of RFID reader to enable a correct read of body temperature measurements. In such a system, each user must have a RIFD skin-sensor attached to his body. When the user arrives at the surveillance

point, the system automatically records his pass, captures his Unique IDentifier (UID) and medical records stored into the integrated circuitry memory. This collected data can be eventually used as indicator sign to detect suspected cases or it can also be stored in the integrated circuitry itself to be used for comparison during the next control stage [8].

C. RFID and tracing contacts

Effective tracking and tracing plays an important role in monitoring the spread of coronavirus disease by identifying people who have been close to suspected cases in order to take the necessary actions, e.g **undergoing quarantine and treatment.** Due to the huge number of cases that are increasing excessively every day, the tracking and tracing process cannot be done effectively with the simple current processes.

In [10] an IoT based Tracing and Tracking system has been developed using RFID technology and a mobile phone that acts as an reader to identify potential contacts. In this system, each person must carry an RFID Tag which contains various data related to his personal and medical information. This stored information are readable by Near Field Communication (NFC) protocol provided in mobile phones. The authors also are developed a mobile app that continuously runs in the background. This mobile app can detect if there is a tagged person in close proximity (within its reading range) and then record and collect their stored data from their mobile app. The collected information will be transferred to the edge device for further processing. If a person is identified as suspected case, all the people they have crossed with it would be identified also as possible infection cases through the mobile app. afterward, those people will be alerted through their mobile apps installed in their mobile phones.

Another solution based RFID can implemented in indoors. In [11] Each user must tagged with a unique identifer. When a user arrives to the entrance, a message will be generate and send to edge server. this message contains user identifier and login details. If one of the users are subsequently found to be infected with the COVID-19 virus, the organization must be notified by the infected user himself using the mobile application. Moreover, all users who entered the same place on the same day will be notified via SMS, email and mobile application.

IV. Conclusions

The spread of COVID-19 has changed the lives of society and disrupted economies around the world. This pandemic has opened many research challenges and opportunities that may benefit our community and help it to face Covid-19 and other infectious diseases that could emerge in the future.

In this article, a review of RFID-based solutions is examined to highlight their features and advantages in the fight against coronavirus. Most of the proposed solutions based on RFID technology are based on protecting individuals, diagnosing infected people, and tracking and tracing infected cases.

References

[1] Zhu H, Wei L, Niu P, "The novel coronavirus outbreak in Wuhan, China", Glob Health Res Policy. 2020 Mar 2;5:6. doi: 10.1186/s41256-020-00135-6. PMID: 32226823; PMCID: PMC7050114.

[2] Esakandari, H., Nabi-Afjadi, M., Fakkari-Afjadi, J. et al. "A comprehensive review of COVID-19 characteristics", Biol Proced Online 22, 19 (2020). https://doi.org/10.1186/s12575-020-00128-2

[3] Mahesh G, Dr.Chidananda Murthy M V, Dr. M. Z. Kurian HOD, "A Review on Various Social Distancing Technologies to combat COVID-19 contagious disease",International Journal of Computer Science & Communication (ISSN: 0973-7391) Special Issue, Page 1-9, December 2020.

[4] C. T. Nguyen et al., "A Comprehensive Survey of Enabling and Emerging Technologies for Social Distancing—Part I: Fundamentals and Enabling Technologies," in IEEE Access, vol. 8, pp. 153479-153507, 2020, doi: 10.1109/ACCESS.2020.3018140.

[5] Marc van Lieshout, et al., "RFID Technologies: Emerging Issues, Challenges and Policy Options" , Office for Official Publications of the European Communities, Luxembourg, 2007, 197.

[6] G. M. Bianco, N. Panunzio and G. Marrocco, "RFID Research Against COVID-19 – Sensorized Face Masks," 2021 IEEE International Conference on RFID Technology and Applications (RFID-TA), 2021, pp. 241-243, doi: 10.1109/RFID-TA53372.2021.9617305.

[7] G. M. Bianco and G. Marrocco, "Sensorized Facemask With Moisture-Sensitive RFID Antenna," in IEEE Sensors Letters, vol. 5, no. 3, pp. 1-4, March 2021, Art no. 6000604, doi: 10.1109/LSENS.2021.3059348.

[8] N. Panunzio, G. M. Bianco, C. Occhiuzzi and G. Marrocco, "RFID Sensors for the Monitoring of Body Temperature and Respiratory Function: a Pandemic Prospect," 2021 6th International Conference on Smart and Sustainable Technologies (SpliTech), 2021, pp. 1-5, doi: 10.23919/SpliTech52315.2021.9566334.

[9] S. Amendola, G. Bovesecchi, A. Palombi, P. Coppa and G. Marrocco, "Design, Calibration and Experimentation of an Epidermal RFID Sensor for Remote Temperature Monitoring," in IEEE Sensors Journal, vol. 16, no. 19, pp. 7250-7257, Oct.1, 2016, doi: 10.1109/JSEN.2016.2594582.

[10] Rajasekar, Sakthi Jaya Sundar,"An Enhanced IoT Based Tracing and Tracking Model for COVID -19 Cases," SN computer science vol. 2,1 (2021): 42. doi:10.1007/s42979-020-00400-y

[11] N. Petrovic and D. Kocic, "IoT for COVID-19 Indoor Spread Prevention: Cough Detection, Air Quality Control and Contact Tracing," 2021 IEEE 32nd International Conference on Microelectronics (MIEL), 2021, pp. 297-300, doi: 10.1109/MIEL52794.2021.9569099.

Logout in identity and access management systems

BELFAIK Yousra
LIMATI Laboratory
FPBM, USMS University
Beni Mellal, Morocco
Email: belfaik.yousra@usms.ac.ma

SADQI Yassine
LIMATI Laboratory
FPBM, USMS University
Beni Mellal, Morocco
Email: yassine.sadqi@ieee.org

SAFI Said
LIMATI Laboratory
FPBM, USMS University
Beni Mellal, Morocco
Email: safi.said@gmail.com

***Abstract*—Identity and access management(IAM) systems are generally designed to facilitate the management of users' identities and data access permissions within organizations. They are closely integrated with other systems such as those that provide the single sign-on mechanism and the verification of their credentials. The IAM systems have been widely used in Cloud, mobile, and IoT applications. However, logout implementation presents a real challenge for IAM systems developers. Implementing logout can be considered more complex to design and test in some cases than login. Even though logging out of an IAM service is important as login, it has received less attention in the research community. In this article, we describe challenges related to logout in identity and access management on web-based services. Moreover, we present the logout mechanisms of the three commonly used industry-standard identity protocols in IAM systems, namely SAML2.0, OAuth2.0, and OpenID Connect, and some implementation guidelines for reliable logout.**

1. Introduction

Digital identity plays an increasingly important role in our interconnected, digitalized world. The term "identity" can be defined as a collection of attributes associated with a specific person or entity in a particular context [1]. For example, most of us have several digital identities, associated with our workplace, our personal life, and other professionalrelated activities. Nowadays, there are thousands of websites providing different services that require users to provide their identities for authentication in order to access those services with the appropriate level of access. This explains the growing reliance on identity and access management (IAM) systems to manage and secure our identity information through a set of operations, including controlling user access, registering, updating, revoking and looking up digital identities [2]. There are many solutions used to manage identity information and provide authentication and authorization. Some of these solutions are Single SignOn (SSO), Identity Federation and Multifactor authentication [3]. Single Sign-On allows users to authenticate once and access other services using their credentials on an Identity Provider (IdP). Identity Federation solution allows users to use a single identity across different enterprises in different trust domains based on a trust factor. The multiFactor Authentication solution authenticates a user by challenging him with multiple authentication factors (e.g., password, SMS, and fingerprint). Recently, IAM systems have been more likely to rely on industrystandard protocols to provide their services, because they have been reviewed for flaws by many people, provide interoperability between applications and service providers, and save time as many programming languages have SDKs that support them [1]. In this work, we will talk about the three commonly used standard identity protocols in IAM systems, namely SAML2.0, OAuth2.0, and OpenID Connect.

Many of the IAM services consider their task complete after authenticating the user at the beginning of a service session and providing his level of access, which is not the case. However, it's very important for users to have a way to terminate their session. This is especially true for shared-device environments that are used to access sensitive applications. For example, if a user connects to the service using a public computer in a library or cafe and does not log out, the next user of the computer may gain access to the service with the previous user's privileges. Terminating a session if it is no longer needed eliminates the chance that the session can be hijacked by others. This can complement other measures of a global security strategy and is beneficial in situations where devices are stolen or confiscated.

The main contributions of this work are as follows :

- Overview of identity and access management system, its different models, and its major functions.
- Present the challenges related to logout in identity and access management on web-based services.
- Describe the logout mechanism in the three most commonly used identity standard protocols used in IAM systems, namely: SAML2.0, OAuth2.0, and OpenID Connect.
- Propose some implementation guidelines for reliable logout.

The rest of this paper is organized as follows. In Section 2, we study the previous works related to our review. Section 3 discusses the identity and access management system, its different models and major functions. In Section 4, we

discuss logout challenges in IAM systems. In section 5, we describe the logout flows on SAML2.0, OAuth2.0, and OpenID Connect protocols. Finally, section 6 provides some logout implementation guidelines.

2. Related work

Logout has not been studied widely, although identity and access management systems have been developed and researched extensively. To the best of our knowledge, we did not find studies talking about the log out in IAM global system; we have just found some research studied log out in SSO and federated identity systems.

In 2014, S. suoranta et all [11] describe challenges related to logout in federated identity management systems and analyze the unexpected behaviour in logout situations. They also give some implementation guidelines for reliable logout and describe a polling-based solution for creating a system-wide logout mechanism that only requires minor changes, does not burden the identity provider, and gives users the option to log out of only one service. Upon their solution's evaluation, they have found that users like to have the ability to choose between the single logout and the local logout options, but they did not understand the difference between them. In addition, they have found that a majority of the users do not log out of the services at all, they just close the browser window.

In 2018, M. Ghasemisharif et all [19] investigate th, security implications of SSO and offer an in-depth analysis of account hijacking on the modern Web. Their analysis shows that the current implementation of SSO exposes users to numerous dangerous and stealthy attacks, some of which extend to services not connected to the original IdP. They also introduce novel attacks that leverage SSO for maintaining long-term control of victims' accounts. Even worse, they have found that the majority of SSO services don't provide the functionality that would allow users to terminate their active sessions and remediate the account takeover. Guided by their findings, they designed single sign-off, an access revocation extension to OpenID Connect that enables users to efficiently recover from an IdP account hijack.

In our previous work [20], we have discussed the huge security risk that SSO systems present in the case of a hijacked IdP account. If the centralized IdP account has been compromised, the attacker can gain access and control of all the accounts associated with the victim's IdP account. We have also performed a comparative analysis of four OIDC mechanisms that can be used as authentication revocation solutions for SSO systems: OIDC Session Management, OIDC front-channel logout, OIDC back-channel logout and Single Sign-Off.

3. Identity and access management systems

Identity and access management (IAM) is a framework of policies and technologies that allow organizations to initiate, capture, record and manage user identities and ensure that only authorized individuals can access the associated resources in an organization [6]. The IAM systems consist of two independent aspects: identity management and access management. Identity management describes the process of authentication, authorization and user privileges across system boundaries, whereas access management is more focused on the access control to verify whether users are granted privileges to access particular services or resources. The decision result is evaluated based on policy rules, user roles and other elements that are predefined by the administrators [5].An IAM system has three main entities: (i) Identity Provider (IdP), which is the entity responsible for generating identities, maintaining user information and authenticating users; (ii) Service Provider (SP), which is a public s available service offers resources and services to registered users; and (iii) the user or device, the entity that uses a service and needs to be authenticated.

3.1. IAM models

IAM systems follow models categorized as traditional, centralized, federated and user-centric [7]:

- **Traditional model:** In the traditional model, the SP operates as both SP and IdP. In this model, there is no identity sharing among SPs. Thus, the user needs to create accounts for each SP he wants to access, with different identifiers and credentials.
- **Centralized model:** In this model, user authentication and account management are handled by a central trusted Identity provider IdP. All the SPs trust this IdP and log users through it. The main advantage of this centralized system is that SP developers don't need to worry about the user management task and users don't need to maintain different user accounts. Single Sign-On (SSO) is a strong application of this centralized architecture. SSO enables users to authenticate once and to use the authentication with all the SPs. However, the IdP is a single point of failure. Also, as the IdP has control over the user's identity information, it may do whatever it wants with such information.
- **Federated model:** This model allows users to use a single identity across different IdPs and SPs in different trust domains based on a trust factor. SPs accept the authentication token issued by an IdP, due to trust relationships established among IdPs and SPs in the federation. This makes access easy, as users are not required to remember a different set of credentials for every application they use. However, the users still have to provide their credentials to each application separately although the credentials used are the same.
- **User-centric model:** in the above models, there is a lack of user control over identity information stored on the IdP, because it has full control over such information and can disclose it to SPs. The user-centric model solves this problem. This model

aims to give users more control over transactions that involve their identity data. For example, in some IAM systems, disclosure of identity attributes requires user consent.

3.2. IAM major functions

The IAM is widely accepted as the first defence line of today's information and communications technology (ICT) infrastructures and services because of the different functions it provides, including identification, authentication, authorization, auditing and account management.

- **Identification:** is the process of assigning a unique identifier (Id) to a human being, a computer, or a network component. Even inanimate components, such as network interface cards, firewalls, and printers, need Ids. In computer systems, this identity is used to determine if a subject can access a particular resource [8].
- **Authentication:** is the process of determining whether the user credentials are authentic. The users are required to prove their right and identity to access information in the system, basically through a username and password or biometric identities. The system verifies user identity by matching the provided credentials with a specific abstract user object that is stored in the system. Once the two objects match, the access is authenticated.
- **Authorization:** is the process of allowing users to perform an action with the resource that they are granted. For example, when a user logs in to his email and provides his credentials, he can add and delete an email and even make certain changes that theys are authorized to make. But, he cannot modify the mail server because he does not have the permission to do so. Furthermore, the users can be authenticated using a certain identity but they can be authorized to access a specific resource under a different identity.
- **Auditing:** is the process of recording security events related to the accounting and traceability process. It can provide historic information about when and how a user accessed the assets, and whether there were any attempts to violate authentication policies. The historic information of user status is stored in the log files for further analysis.
- **Account management:** is the process of identifying and controlling the state of users when logged into the system. It encompasses a set of administrative functions such as identity creation, propagation, and maintenance of user identity and privileges. This enables the administrator to have the better granularity to control the user authentication and manage the lifespan of user accounts through user lifecycle management, thus ranging from the initial stage of authentication to the final stage when the user logs out of the system. The flexibility provided by account management allows administrators to implement IAM efficiently with a closer match to the security policy [5].

4. Logout challenges in IAM

Logg out means clearing the sessions that are created when the user is logged in. In a traditional web service, the user session is usually terminated either when the user explicitly presses the logout button or when the service timeout is reached [11]. Logout can be complex to implement in environments with single sign-on, because there may be multiple sessions to worry about. However, we can find three different scenarios of user authentication sessions in IAM systems (see figure 1): (1) The user has at minimum an application session, (2) The application and the identity provider have active sessions for the user in the case of delegated authentication, (3) in the case of an application uses an authentication broker to facilitate handling different IdPs and protocols, the authentication broker will also have an active session for the user.

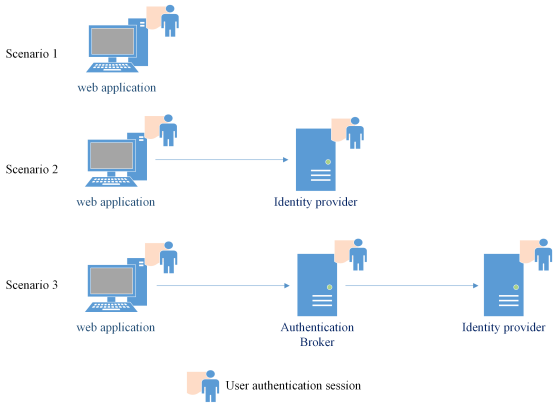


Figure 1. Scenarios of user's authentication session.

In the case of single sign-on, the logout is more complicated because there might be even more sessions to consider. If a user accesses multiple applications via SSO, there could be an additional session in each of those applications. The user might want to log out either from a single service or from the whole SSO session. The mean problem here is users could not predict what will happen when they press the logout button. Some of the services closed only their session and the SP session, logging the user only out of the local service and not of the IdP. On the other hand, we have some services also terminated the session on the IdP using the single logout feature. In single logout (SLO), users can log out of both the IdP session and all applications in the current IdP login session [10]. This logout could be initiated by the SP or the IdP. A user needs to have the ability to decide what should happen when any of the user's sessions are terminated, or at least know which one of his sessions are terminated and which are not (i,e the application session, the authentication broker session, and/or the identity provider session).

In practice, designing logout must consider where sessions exist and which should be terminated when the logout is initiated either directly by the user, or by an IdP or SP due to a session timeout, administrator command, etc. Indeed, there is no defined rule to determine which logout feature service should invoke; rather, it appears to be an implementation decision independent of each service developer. In enterprise environments, the corporate security policy may dictate that a logout in an application must trigger the termination of the identity provider session and possibly all the user's open sessions in the relaying applications. In an environment where the user logs in with a social provider, it may be less justified or impossible for an application's log out to terminate the user's session with the social identity provider. Therefore, the features supported by an identity provider or authentication broker for logout are another factor that will influence logout design [1].

User experience is another important factor to take into consideration in designing logout. Terminating all application and SSO sessions for a user with one logout provides a convenient way to terminate all access at once. This may be desirable in an enterprise environment because if users have to log out of each application individually, they may forget one. However, if the impact of such a logout is not clear to a user, this may cause him to lose his work in other applications that rely on the same SSO session, depending on how the logout is implemented. For example, using scenarios 2 and 3 in figure1, when an application sends a logout request to the authentication broker or the IdP, these last could simply terminate their own session for the user and let the user continue working in other applications until his session times out. Upon timeout, the application would check if the user's session is valid at the SSO server or not. If not, the user would need to log in again to continue to access this application.

5. Logout in IAM protocols

In this section, we describe the logout flows in the three most commonly used standard identity protocols in dentity and access management systems, namely SAML2.0, OAuth2.0, and OpenID Connect.

5.1. SAML2.0 logout

OASIS, the standardization organization behind Security Assertion Markup Language (SAML) standard, defines it as an XML-based framework for describing and exchanging security information between online business partners. This security information is expressed in the form of portable SAML assertions that applications working across security domain boundaries can trust. The SAML standard defines precise syntax and rules for requesting, creating, communicating, and using these SAML assertions [12]. With SAML 2.0, an SP application can terminate a user's session at an IdP and all the relying parties (SP applications) with an active session for the user. Upon receiving a logout request message from the SP, the IdP terminates the SSO session it holds for the user, identified by a subject identifier in the request and possibly a session identifier for the session. The IdP may also update or remove its session cookie in the user's browser. Next, the IdP responds to the SP application with a logout response message. Moreover, upon the termination of the IdP session or receipt of a logout request message from an SP, the IdP can send a logout request message to each of the other relying parties (i.e, SPs) with the same user session. The SPs are supposed to terminate their session and respond with a logout response message to the IdP. If the single logout (SLO) was initiated by one SP, the IdP returns a logout response message to the relying party that initiated the logout [1]. Figure 2 illustrates the SAML2.0 SP-initiated Single Logout.

The SAML2.0 specification defines two different logout mechanisms: the frontchannel logout and the backchannel logout. While the frontchannel logout messages are sent via a web browser, the back-channel logout messages require direct connectivity between the identity provider and the relying parties. Many SAML 2.0 identity providers reside behind corporate firewalls, and using a front-channel implementation avoids issues with firewalls. Even though back-channel logout mechanism may be more reliable if the logout messages need to be sent to multiple parties, it's still challenging to implement this mechanism, especially for components behind corporate firewalls.

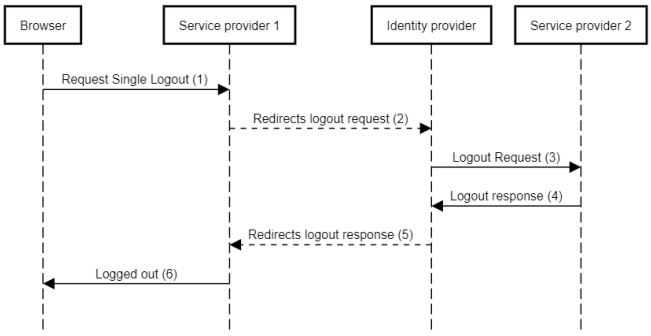


Figure 2. SAML2.0 SP-initiated Single Logout.

5.2. OAuth2.0 logout

OAuth2.0 is an authorization standard protocol, created in 2012 by the OAuth IETF working group. This protocol allows users to grant limited access to their resources on one application, to another, without having their credentials exposed [14]. The OAuth 2.0 does not contain a logout endpoint because it is designed just for users authorization and not authentication. However, OAuth2.0 indicates that authorization servers should provide a mechanism for clients (i.e, SPs) to revoke access tokens, and OAuth2.0 defines the Token Revocation extension as a standard for this.

This revocation mechanism allows an application to invalidate its tokens if the user logs out, changes identity, or uninstalls the respective application. Notifying the authorization server that the token is no longer needed allows

the authorization server to clean up data associated with that token (e.g., session data) and the underlying authorization grant. This prevents abuse of abandoned tokens and facilitates a better user experience since the invalidated authorization grants will no longer appear in the list of authorization grants the authorization server might present to the user [13]. The OAuth2.0 token revocation flow is shown in figure 3.

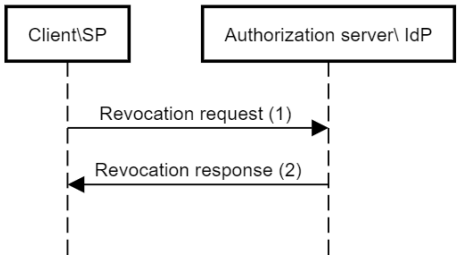


Figure 3. OAuth2.0 Token revocation.

(1) The client (i.e, SP) sends a revocation request including the token that wants to get revoked and the client authentication credentials.

(2) The authorisation server (i.e, IdP supports OAuth2.0) verifies the request, invalidates the token, and responds with HTTP status code 200 if the token has been revoked successfully or if the token is already invalid.

5.3. OpenID Connect logout

OpenID Connect (OIDC) is an authentication and authorization standard protocol based on building a simple identity layer on top of OAuth2.0, created in 2014 by the OpenID Foundation [15]. It allows applications, called relying parties RPs, to authenticate a user by obtaining his basic profile information (name, family name, given name, email, picture, etc.) from an OpenID Provider (i.e, IdP supports OIDC). OIDC defines three mechanisms for logout namely, OIDC session management, OIDC frontchannel logout and OIDC backchannel logout.

The OIDC Session management is a logout mechanism that defines how to monitor the user's login status at the IdP on a continuous basis so that the SP can automatically log out a user who has logged out of the IdP [16]. This mechanism uses two invisible iframes, RP iframe and OP iframe. The OP iframe is loaded from the IdP and has access to the browser state from the IdP. Therefore, the RP iframe is loaded from the RP itself. The RP iframe polls the OP iframe with a postMessage containing the client_id and the actual session_state parameters. Next, the OP iframe accesses the IdP browser state cookie and calculates the new session state parameter and compares it with the one sent by the RP, and then responds with a postMessage containing the status as changed, unchanged or error. If the status is changed (the user's session at the IdP has changed), the RP application can redirect the user to the IdP with a new authentication request using prompt=none, and if this request receives an error response, it indicates the user session at the IdP is no longer valid.

The OIDC frontchannel logout mechanism provides a solution for IdPs to send logout requests to RPs to notify them when the IdP session has terminated. The frontchannel logout relies on redirecting logout messages between RPs and an IdP via the user's browser. The RPs supporting this logout mechanism must register a logout URI called frontchannel logout URI, which must be accessible by the user's browser. Upon receiving a logout request, the IdP provides a response page with the logout URI for each RP that has an active session with the IdP. Then, The browser sends a logout request to each RP containing the frontchannel logout URI. Upon receiving a logout request, the RP clears the state associated with the connected session, removing all associated cookies and data stored locally within the user's browser [17].

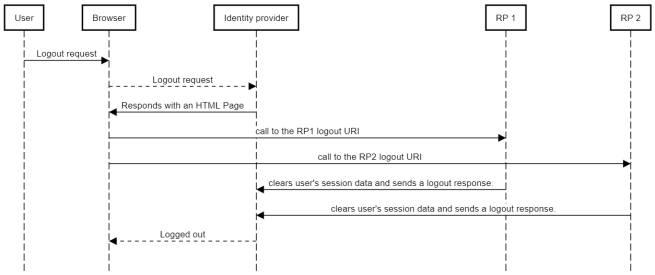


Figure 4. OIDC frontchannel logout.

The OIDC backchannel logout mechanism provides a solution for an IdP to send logout requests to RPs via backchannel communication between servers rather than via frontchannel browser redirects. This may provide a more reliable logout option than frontchannel log out when there are many relying parties. The RPs supporting this solution should register a backchannel logout URI with the IdP as part of their client registration. On the other hand, the IdPs need to monitor the group of the RPs to which a user has logged in via their IdP session. When the IdP session for the user is terminated, the IdP first generates a logout token, which is a JWT, and then sends an HTTP-POST request to the registered backchannel logout URI for each RP including the generated logout token. Upon receiving a valid logout token, each RP clears its session for the user and returns a status response to the IdP [18].

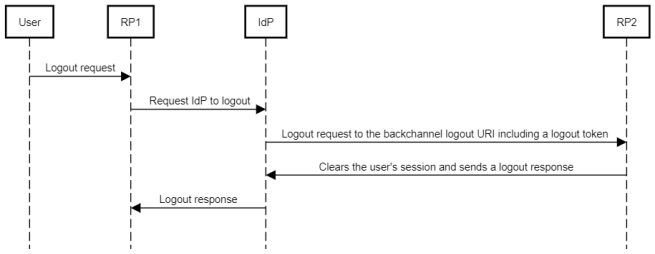


Figure 5. OIDC backchannel logout.

6. Logout implementation guidelines

In this section, we discuss some possible mitigations in logout implementations to make users able to securely log out of the services that are being used and also end the IdP session whenever they want.

Cookie management in web browsers can be considered the most important key for a reliable log out, especially for solutions based on browser redirections. In the current model of session management, there are some rules on how cookies should be handled to mitigate problems caused by attacks. These rules prevent the IdP and SP from seeing each other's cookies. Thus, for making single logout possible in this case, an IdP could create a collection of cookies connected to the SSO session, where SPs could add their own cookies to the collection. Then, the IdP could remove cookies and set delete rights on all servers in the federation. However, changes in cookie management may cause unexpected security weaknesses if not done with special care.

Moreover, it's very important to have a standard for the process of removing session cookies and managing their timeout periods. This standard should redefine the process of local logout and single logout, the difference between the two, and the order of the session removal in each scenario. Users need to know that single logout removes all the session information from all the services. After closing the web browser, the sessions are on the client-side end. Contrariwise, the sessions on the server-side remain until timeout, but they cannot be used without knowing the session cookies. The problem here is users also have other web browser windows and tabs open and do not want to close them all but just log out of one of the services. Thus, browsers need to be improved to allow the ending of sessions without closing the whole browser.

7. Conclusion

Although the Identity and Access Management (IAM) system has been widely used in web, mobile and IoT applications, the logout feature and how user sessions are terminated in this system are still not clear. Even worse, this issue has not gained the attention it deserves in the scientific community. In this paper, we have presented the logout challenges in IAM web-based services. In addition, we have described the logout mechanism in SAML2.0, OAuth2.0, and OpenID Connect identity protocols, and proposed some implementation guidelines for reliable logout. We hope this work highlights the big problem of logout in IAM systems and draws more attention to this issue. In future work, we will propose a reliable logout mechanism facing the challenges described above.

References

[1] Yvonne W. and Abhishek H. (2019).Solving Identity Management in Modern Applications. https://link.springer.com/book 10.1007/978-1-4842-5095-2.

[2] Zhu, X. , and Badr, Y. (2018). Identity Management Systems for the Internet of Things: A Survey Towards Blockchain Solutions. Sensors, 18(12), 4215. doi:10.3390/s18124215.

[3] B. Piyathilake. Identity and Access Management Components. Medium, 6 mars 2022. https://bojithapiyathilake.medium.com/identity-and-access-management-components-c501d786db4c.

[4] Pöhn, D. , and Hommel, W. (2020). An overview of limitations and approaches in identity management. Proceedings of the 15th International Conference on Availability, Reliability and Security. doi:10.1145/3407023.3407026.

[5] Zhang, Z. , and Meddahi, A. (2017). Identity and Access Management in NFV. Security in Network Functions Virtualization, 135–155. doi:10.1016/b978-1-78548-257-1.50004-2.

[6] Y. Liu, D. He, M. S. Obaidat, N. Kumar, M. K. Khan, et K.-K. Raymond Choo, Blockchain-based identity management systems: A review, J. Netw. Comput. Appl., vol. 166, p. 102731, sept. 2020, doi: 10.1016/j.jnca.2020.102731.

[7] Cordeiro Domenech, M. , Comunello, E. , and Silva Wangham, M. (2014). Identity management in e-Health: A case study of web of things application using OpenID connect. 2014 IEEE 16th International Conference on e-Health Networking, Applications and Services (Healthcom). doi:10.1109/healthcom.2014.7001844.

[8] R. Sandhu, J. Hadley, S. Lovaas, et N. Takacs, Identification and Authentication, in Computer Security Handbook, John Wiley and Sons, Ltd, 2012, p. 28.1-28.21. doi: 10.1002/9781118851678.ch28.

[9] Karunanithi, D. , and Kiruthika, B. (2011). Single sign-on and single log out in identity. International Conference on Nanoscience, Engineering and Technology (ICONSET 2011). doi:10.1109/iconset.2011.6168044.

[10] Configure Single Logout in app integrations — Okta. https://help.okta.com/en/prod/Content/Topics/Apps/Apps_Single_Logout.htm.

[11] S. Suoranta, K. Manzoor, A. Tontti, J. Ruuskanen, et T. Aura, Logout in single sign-on systems: Problems and solutions, J. Inf. Secur. Appl., vol. 19, no 1, p. 61-77, févr. 2014, doi: 10.1016/j.jisa.2014.03.005.

[12] sstc-saml-tech-overview-2.0.pdf. https://www.oasis-open.org/committees/download.php/27819/sstc-saml-tech-overview-2.0-cd-02.pdf

[13] T. Lodderstedt, S. Dronia, et M. Scurtescu, OAuth 2.0 Token Revocation, Internet Engineering Task Force, Request for Comments RFC 7009, août 2013. doi: 10.17487/RFC7009.

[14] Hardt ¡dick.hardt@gmail.com¿ D. The OAuth 2.0 Authorization Framework. Disponible sur : https ://tools.ietf.org/html/rfc6749.

[15] Final: OpenID Connect Core 1.0 incorporating errata set 1. https://openid.net/specs/openid-connect-core-1 0.html

[16] Draft: OpenID Connect Session Management 1.0 - draft 31. https://openid.net/specs/openid-connect-session-1_0.html

[17] M. Jones. OpenID Connect Front-Channel Logout 1.0 - draft 05, April 18, 2022. https://openid.net/specs/openid-connect-frontchannel-1 0.html.

[18] Jones, M. B., and Bradley, J. OpenID Connect Back-Channel Logout 1.0 – draft 07, April 18, 2022. https://openid.net/specs/openid-connect-backchannel-1_0.html.

[19] Ghasemisharif M, Ramesh A, Checkoway S, Kanich C, Polakis J. O Single Sign-Off, Where Art Thou ? An Empirical Analysis of Single Sign-On Account Hijacking and Session Management on the Web .USENIX. https://www.usenix.org/conference/usenixsecurity18/presentation/ghasemisharif.

[20] Y. Belfaik, A. Lmouhandiz, Y. Sadqi, et S. Said, Single Sign-On Revocation Access, 2022, p. 535 544. doi: 10.1007/978-3-030-91738-8_49.

Low overhead routing in wireless sensor networks

1st Maryem LACHGAR
Department of Mathematics and Informatics
Polydisciplinary faculty, University Sultan Moulay Slimane
Beni-Mellal, Morocco
maryem.lachgarfpb@usms.ac.ma

2nd Anouar DARIF
Department of Mathematics and Informatics
Polydisciplinary faculty, University Sultan Moulay Slimane
Beni-Mellal, Morocco
anouar.darif@gmail.com

3rd Hicham MOUNCIF
Department of Mathematics and Informatics
Polydisciplinary faculty, University Sultan Moulay Slimane
Beni-Mellal, Morocco
h.mouncif@usms.ma

***Abstract*—Minimizing energy consumption means minimizing the amount of data transmitted through the network.**
Link breakage is one of the main challenges facing routing protocols in wireless sensor networks. It suspends the traffic flow or stops it completely sometimes. Therefore, the routing protocol design must take into account the link break to find an alternative route or handle the problem locally by providing a local link repair that replaces only the broken part of the route with minimal routing overhead to save energy.
In this work, the Lightweight Routing Protocol (LRP) with local repair and different mechanisms to reduce routing overhead have been introduced.

***Index Terms*—Network Lifetime, Wireless Sensor Network, ROUTING, LRP, local repair**

I. Introduction

Routing in multi-hop wireless sensor networks [4] is crucial for discover, establish and maintain routes along which data packets can be routed. this necessity arises from the fact that networks are usually too large to be comprehended in their entirety by the human mind, that the quality of wireless links is highly variable, nodes may go down, new nodes may join the network, and nodes may change their location, and the calculations to find the best route among those available is a fastidious task. So, the routing protocol needs to guarantee that if the topology is modified, the routing structure is properly adapted without creating loops. Moreover, it has to introduce minimal overhead and take into account the contraints of the available energy.
LRP, the Lightweight Routing Protocol combines the collection tree or DODAG topology of RPL [5] to extract traffic out of the network through the sink and the reactive route establishment of LOADng [6] to distribute traffic inside the network. this combination aims at generating as little traffic as possible while remaining simple.
in this paper we start with the description of the protocol LRP (Section II) .Then we define the different mechanisms to reduce the routing overhead (Section III) .(Section IV) present results that we obtain and a discussion of performances in each scenario and finally a conclusion is given in (Section V).

II. Wireless Sensor Network

Wireless sensor networks (WSNs) are application-oriented and mainly used in the following fields, such as battleground surveillance, border security, and disaster management [8]. The use of these sensors on a large scale has led to the creation of a multitude of new applications and services capable of improving our quality of life. But WSNs have constrained resources in terms of processing capability, memory, power, whereas sink is a resource-rich device having a continuous power supply and communication capability. Therefore, the sink can communicate with other computers using additional networks, such as a local area network, wireless local area network, wireless personal area network and Internet [9]. Fig 2 shows the architecture of a typical WSN.

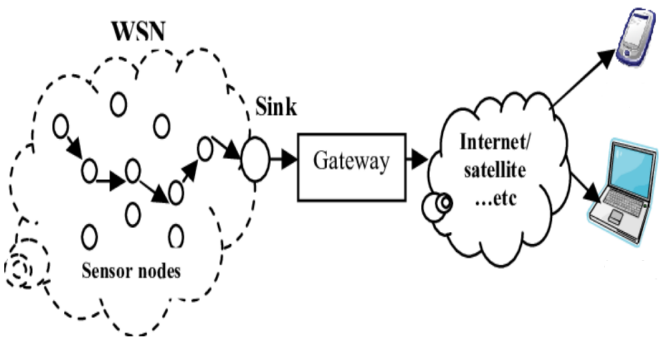


Fig. 1. Architecture of a wireless sensor network.

because of their unique characteristics, the transposition of the solutions and protocols existing in traditional networks to this type of environment is impossible. Standardization efforts have led in recent years to the design of many protocols adapted to the constrained environment of the motes used such as the LRP routing protocol which is important for improving results in a wide range of sensor-based applications. In the following sections, we will see the mechanism used by LRP to minimize the overload of control messages.

III. LRP - Protocol Overview

A. Collection tree

To easily extract data generated by nodes outside the sensor network lrp organizes the nodes into a collection tree.This type of traffic is routed the same way regardless of the source of the packet and uses a single route called the default route with the set of default routes installed in the routing tables of the individual nodes forming a tree rooted on the sink. The positions in the collection tree are computed using the same distributed Bellman-Ford algorithm [1].
After initiating the construction of the tree by the sink, the DIO propagates in the network. When receiving a DIO, a node compares its current position with the one it would have by selecting the sender of the message as successor. If the proposed route allows it to approach the sink, it has found a better position in the tree.

fig2 shows the process of the Bellman-Ford algorithm for the construction of the collection tree.

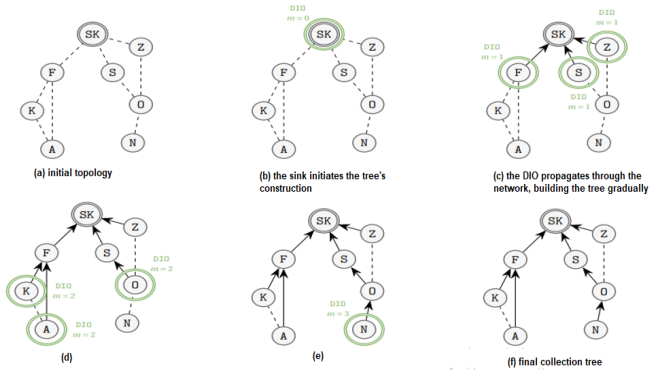


Fig. 2. step process of the distributed Bellman-Ford algorithm for the construction of the collection tree.

B. Host routes

LRP also allows the nodes to establish host routes to distribute traffic into the network. The initial establishment of the host routes is done in a proactive way, and the search for a host in the network is done in a reactive way. To prevent loops from occurring, an implicit data path validation mechanism in the LRP detects routing loops the moment they are used [2].

C. Local repair

To ensure there are no loops into the network, LRP imposes that a node must never accept a route that would put it further away from the sink than its current position. However, there are some particular situations that are problematic, such as when a node tries to find a new successor among its neighbors, but none of its neighbors is close enough to the sink to offer a satisfactory path, so it cannot connect to the sink at all. At this point, the tree needs to be repaired.

Local repair impacts only a small number of nodes by proceeding locally. It is activated whenever a node loses all successors and thus has to turn to using a node in its own sub-tree as successor [3], It can be triggered by any of the nodes in the network when the need arises. This is done in two parts: searching for a path from the detached node to the sink using a BRK (Break) message and rebuilding part of the tree from the sink without a routing loop using the UPD (Update) message.

– Find an alternative route to the sink The BRK (BRoKen) message is transmitted up through the network as a multicast within the node's subtree, and as a unicast along the collection tree outside using the default routes. This step has two purposes: to notify the sink of the need for a local repair, and to find a path between the sink and the detached node as show in fig3

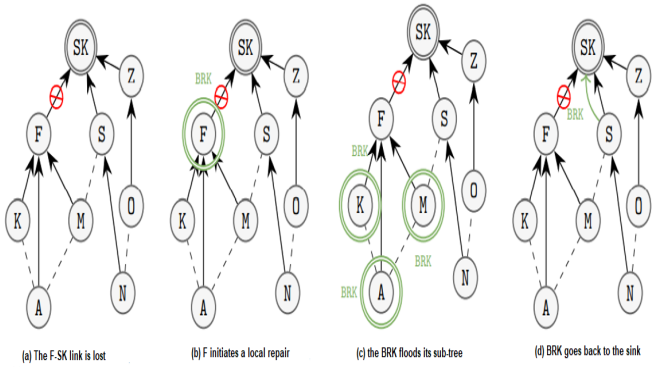


Fig. 3. Local repair "find an alternative route to the sink".

– Updating the collection tree without creating a loop When the sink receives the BRK message, it replies with an UPD (Update) message. The path taken by the Break msg to reach the sink is stored by the nodes in the first step, as it passes the sink. The UPD message is transmitted from node to node on the same path in the opposite direction. When a node receives an UPD message, it switches to using the UPD sender as successor. And so the path is rebuilt and F finds another successor.

IV. Reduction of Routing Overhead

A. Reduce BRK flooding

When a detached node initiates a local repair, a BRK message is broadcast throughout the detached sub-tree. In general, flooding the entire sub-tree is not necessary, in many cases the repair routes could be found through a not so distant predecessor. In order to reduce the number of messages broadcast, LRP proposes the use of the an expanding ring search scheme(ERS) [7]. the BRK message specifies a ring-size field, describing the maximum distance that the BRK is allowed to travel in the sub-tree. it is decremented at each hop and, when it reaches zero, it inhibits the retransmission of the BRK packet. If the nearest nodes cannot be used to repair the

collection tree, there is nothing to prevent the detached node from setting the ring size to a larger value.
The BRK range limitation applies only inside the subtree. When the BRK is outside, the route to the sink has been found and the message is unicast. The ring-size field is then ignored.

B. Spontaneous Response to a DIO

When B the neighbor of a node A at a distance n from the sink broadcasts a DIO describing a distance to the sink which is larger than A. the latter replies spontaneously by sending back a unicast DIO message describing its own distance, thus inciting the neighbor B to change its successor.
Using this mechanism, N can easily explore its neighborhood and detect potential successors by publishing its own distance to the sink in an DIO. It will only receive responses from neighbors that it can use as successors.

C. Host routes repair

In a local repair several nodes change successor, The following problem is the set of nodes in its subtree, whose host routes are also no more usable.
After having detected that host routes are broken , the sink will reactively look for their destinations by flooding several RREQ into the whole network. This my create a significant broadcast traffic that we want to avoid. Proactively rebuilding the host route by transmitting an RREP on behalf of the nodes in its subtree is a naive solution because it cannot modify the sequence number it received previously with the RREP. It is therefore imperative that the host generate a new sequence number itself, thus rendering the old one obsolete. This is the purpose of the confined RREQ as presented in the figure below.
each node receives confined RREQ transmits it further, however, the CONFINED flag it carries indicates that it should only be considered when a node receives it from its successor. F signals the situation to its entire sub-tree by broadcasting in confined RREQ, subsequently Spontaneous RREP go up the collection tree and finally All host routes are reconstructed.

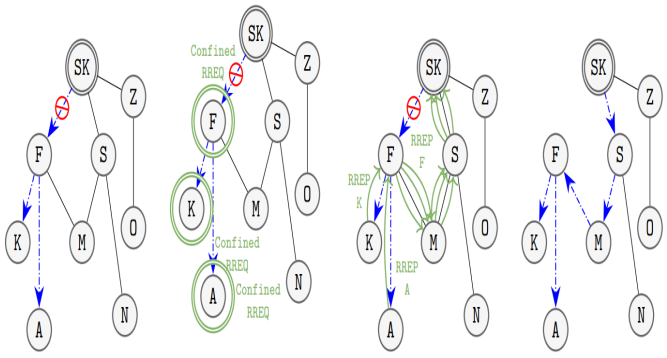


Fig. 4. Local repair "find an alternative route to the sink".

V. Results and Discussion

The objective of this experimentation is to compare the two versions o-LRP and LRP in a real situation, where several functionalities are not activated in o-LRP, in order to verify the efficiency of these mechanisms defined in sections IV-A, IV-B and IV-C.

TABLE I
Node parameters during the experiment

Random delay for packet transmission following a broadcast	0.5 s
Frequency of neighborhood probing when disconnected	5 min
Blacklist timeout	10 min
Metric type	Hop count

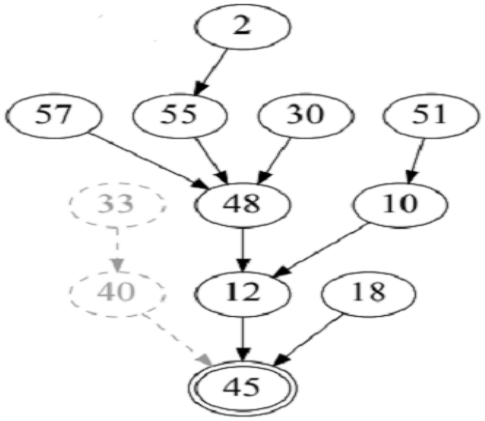


Fig. 5. Collection tree of the experimentation.

TABLE II
Number of messages sent during the whole experiments

		o-LRP	LRP
Broadcast	DIO	37	4
	BRK	7	5
	RREQ	11	3
Unicast	DIO	n/a	1
	BRK	20	4
	UPD	12	5
	HELLO	n/a	2
	RERR	3	-
	RREP	25	17

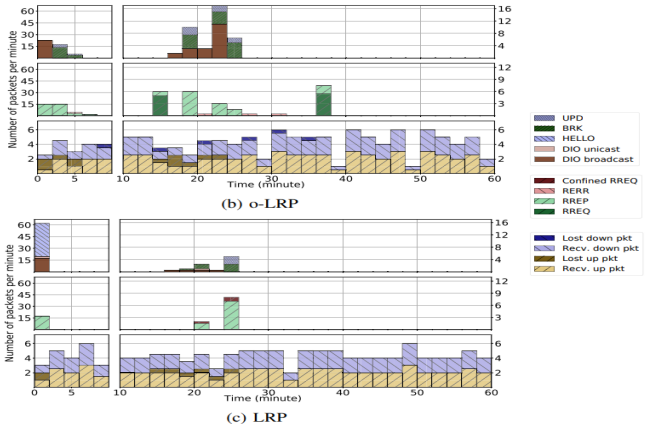


Fig. 6. 60 min experiments with 11 client nodes and 1 sink.

o-LRP network: node 40, having no way to reach the sink, applies the local repair algorithm, resulting in the appearance of DIO, then BRK and UPD messages, shortly after 15 min. This experiment does not go exceptionally well. Indeed, a network saturation phenomenon occurs, due to another node that had failed to proactively build its host route. Since in o-LRP (the original version) the detached node must be removed and rebuilt reactively, this results in a flood of RREQ and RREP messages.

LRP network : local repair performs in the same way as o-LRP, but at this point host routes that access the repaired subtree are not necessarily updated. So it also proactively repairs all host routes with the confined RREQ mechanism.

In both experiments, connectivity is successfully restored in a few minutes. However, as shown in table II o-LRP uses more RREQ, RREP, and RERR messages compared to LRP, which leads to higher control traffic and energy consumption

VI. Conclusion

With the implementation of these mechanisms presented above, we can consider that LRP is one of the lightest and most efficient routing protocols. Nodes send few control messages, the General approach of LRP is to initialize the collection tree proactively, and then completely stop transmitting messages. If the links disappear, they are detected during the transmission of data packets and LRP reacts immediately to repair the tree and proactively refresh these host routes to avoid their reactive rebuilding, and thus a network flood.

The perspectives : Link quality estimation is a distinct issue from routing, however, it is not a negligible issue.A good routing protocol must have a link quality estimator that is both reliable so that good routing decisions can be made about that link because the accuracy of routing decisions is biased in advance if the link quality is incorrectly estimated, and fast to construct that estimate in a very small number of packets.

The next step is to look for and propose a methode that allows LRP to be able to estimate precisely the link quality when few exchanges have been made on it.

REFERENCES

[1] R. BELLMAN. "On a Routing Problem". In : Quarterly of applied mathematics (1958). DOI : 10.1090/qam/102435.

[2] H. J. Audeoud, M. Krol, Martin Heusse, Andrzej Duda. Low overhead loop-free routing in wireless sensor networks. 11th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), Oct 2015, Abu Dhabi, United Arab Emirates. pp.443-451, ff10.1109/WiMOB.2015.7347996ff. ffhal-01286291f.

[3] udeoud, H.-J., Heusse, M. (2017). Experimental Comparison of Routing Protocols for Wireless Sensors Networks: Routing Overhead and Asymmetric Links. 2017 29th International Teletraffic Congress (ITC 29).

[4] H. El-Sayed, Hamdy and Al Bayatti, Hilal (2021) "Comparisons of Some Multi-Hop Routing Protocols in Wireless Sensor Networks," Information Sciences Letters: Vol. 10 : Iss. 3 , Article 14.

[5] Royaee, Z., Mirvaziri, H., Khatibi Bardsiri, A. (2020). Designing a context-aware model for RPL load balancing of low power and lossy networks in the internet of things. Journal of Ambient Intelligence and Humanized Computing. doi:10.1007/s12652-020-02382-4.

[6] Thasan Leenas,Shriparen Sriskandarajah."Comparison of Proactive, Reactive, and Hybrid Routing Protocols in Mobile Ad Hoc Networks".Conference: 2021 10th International Conference on Information and Automation for Sustainability (ICIAfS).

[7] Duy Ngoc Pham, Van Duc Nguyen, Van Tien Pham, Ngoc Tuan Nguyen, Xuan Bac Do, Trung Dung Nguyen, … Kaiser, T. (2010). An expending ring search algorithm for Mobile Adhoc networks. The 2010 International Conference on Advanced Technologies for Communications. doi:10.1109/atc.2010.5672725.

[8] G. Preeti, S. Tripathi, and S. Singh, "Energy efficient hotspot problem mitigation techniques using multiple mobile sink in heterogeneous wireless sensor network," Int. J. Commun. Syst. (2020).

[9] Mehto, A., Tapaswi, S., Pattanaik, K. K. (2019). A review on rendezvous based data acquisition methods in wireless sensor networks with mobile sink. Wireless Networks. doi:10.1007/s11276-019-02022-6

The contribution of RBF neural networks to the command and control of 5MW wind turbine Horizontal axis class Based on differential speed regulation

By using a Bond Graph Approach

Mohssine KARIMI *

Industrial Engineering Laboratory
Sultan My Slimane

University Beni Mellal, Morocco
karimi.mohssine@gmail.com

Mustapha Zekraoui

Industrial Engineering Laboratory
Sultan My Slimane

University Beni Mellal, Morocco

Zakaria Khaouch

Industrial Engineering Laboratory
Sultan My Slimane

University Beni Mellal, Morocco

Mouad Kadiri

Industrial Engineering Laboratory
Sultan My Slimane

University Beni Mellal, Morocco

Youssef Khoya

Industrial Engineering Laboratory
Sultan My Slimane

University Beni Mellal, Morocco

Abstract

The study of the dynamic behaviour of wind turbine is one of the most attractive fields for researchers in the area of renewable energies. Aiming at the problems arising from the wind turbine with the frequency converter, this paper proposes a new design of wind turbine based on differential speed regulation. The first input shaft of the differential gear transmission is driving by a rotor of the wind turbine, the second input shaft of the DGT had an control signal delivered by using an servo machine, which controlled by an RBF neural networks; the output shaft of DGT is connected to the grid via the asynchronous generator. In order to control the speed of servo machine to the optimal value, the proportional and integral (PI) controller is used. The neural networks and evolutionary algorithms are tools that provide a suitable ground to determine the optimal PI gains. The robustness and effectiveness of the proposed design and controller are verified. The simulation results are conducted to confirm the validity of the proposed design. Hands-on experience is carried out by means of the 20-sim software package (a demo version is available for free on the Internet).

Keywords:

- **Bond Graph, Wind Turbine Modeling, DGT, gearbox, power split, PSO, RBF, Neural Networks, Mechatronic control.**

I- *Introduction*

Several wind turbine uses a special generator with a frequency converter, which consumes a higher power to control the current delivered by the generator to the electrical network. It's worth mentioning that the majority of wind energy conversion systems are based on permanent magnet synchronous generator (SG) and the doubly-fed induction generator, this variety of electrical machines can causes some problems like high reactive power dissipation and a big harmonic content moreover can a low ability of low voltage ride through, and low reliability of frequency converter [1].

The objective of this study is to elaborate a novel control system which helps the wind energy conversion system WECS to deliver an electrical power more stable, and reduces the energy dissipation. The variety of wind turbine operating areas possess many challenges to maintain our objective specially at modelling level, subsequently determine the best ways to the command and control system.

Wind turbine is a complex mechatronic system, which contains many different subsystems [2]:

- Aerodynamic-subsystem
- Mechanical-subsystem
- Electrical-subsystem
- Pitching-subsystem
- Control-subsystem

This variety of technical area makes the mechatronics approach one of the best way to design a wind turbine as an integrated system. There are many publication deals with this topic [3-4], in [3] a bond graph methodology is employed to obtain a per-unitized dynamic model for wind turbine two-mass drive train, other contribution, in [4] a bond graph doubly-fed wind turbine generator Control is proposed, only the power curve is considered. In addition, [5] detailing a mechatronic modelling of a 750 KW fixed speed win energy conversion system using bond graph approach, all these contributions highlight the importance of having a mechatronic model of the wind turbine, in order to conduct dynamic study and analyzes of the integrated system.

In the same reference frame, in order to overcome the shortages in the existing wind turbine. This paper proposes a novel wind turbine with power split, and develop a new mechatronic control based on RBF neural networks. The structure of the wind turbine based on a standard planetary gearbox, which has a sun gear, a ring gear and a carrier is used. The main shaft of the wind turbine, servo machine, and an induction machine are connected to the three terminals of the Differential Gear Train (DGT). In order to ensure that turbine follows the pre-defined power curve, the proposed mechatronic control employs radial basic function neural networks RBF-PI and evolutionary algorithms PSO to provide an optimal PI gain. All system components are modelled in detail by using the Bond Graph Approach (BGA), the implementation of the global model has been carried out by using 20-sim software (a demo version is available for free on the internet). All the parameters of the wind turbine were taken from NACA 4415 profile [6-7] .Reporting the verification of results. Conclusions of the conducted research investigation are drawn.

II- *Modelling*

1- Wind turbine model:

The wind turbine studies in this paper is a class of 5MW (VSWT) variable-speed horizontal axis with power splitting transmission. The schematic structure of this turbine is represent in fig 1.The major components of this wind turbine model are the aerodynamics subsystem, the pitch actuator, the drive train, the servomotor and the generator. The aerodynamic subsystem converts the wind energy captured by the blades into an aerodynamic torque ***Ta***, and drives the main shaft which modelled by an elastic linkage (***Kms***, ***Dms***). The mechanical energy will be divide into two-power flow by the planetary transmission. One of them is convert to the electrical energy through the main generator, which is driven by the sun gear; the other is connected to the servomotor by the ring gear, the shaft servomotor's is modelled by an elastic linkage ***(Ksm, Dsm).*** The deferential gear transmission increase the rotor speed to value more suitable, typically from 20-50 rpm to 1000-1500 rpm.

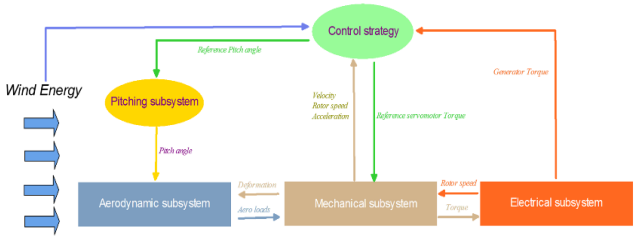


Fig 1: Schematic structure of the wind turbine with power splitting transmission.

2- *Aerodynamic subsystem:*

The kinetic energy of the wind turbine captured by the turbine blades is transformed into mechanical energy, which is transformed to a motor torque Ta rotating the rotor at angular speed ***wr*** , the aerodynamic torque depends nonlinearly on the wind speed V, the tip-speed ratio λ ,and the blade pitch angle β, such that we have [8]:

$$T_a = \frac{1}{2}\rho\pi R^2 V^2 \frac{C_p(\lambda,\beta)}{\lambda} \qquad \{E.1\}$$

With:

$$\lambda = \frac{R.w_r}{V} \qquad \{E.2\}$$

Were ***R*** the rotor radius, ***ρ*** is the density of air, ***Cp*** denote the aerodynamic power coefficient. The power coefficient curve ***Cp(β,λ)*** is shown in Fig 2, the plot is made with different pitch angle and tip-speed ratio, the maximum value ***Cp_max*** is taken when the pitch angle β is set on 2,2o and tip-speed ration set on 8,7 .To calculate the aerodynamic power coefficient we use the following expression proposed by HEIER[9]:

$$C_p(\lambda,\beta) = 0{,}5176\left(\frac{116}{\lambda_i} - 0{,}4\beta - 5\right)e^{-\frac{21}{\lambda_i}} + 0{,}0068\lambda$$

With:

$$\lambda_i = \frac{1}{\frac{1}{\lambda + 0.08\beta} - \frac{0.035}{\beta^2+1}}$$

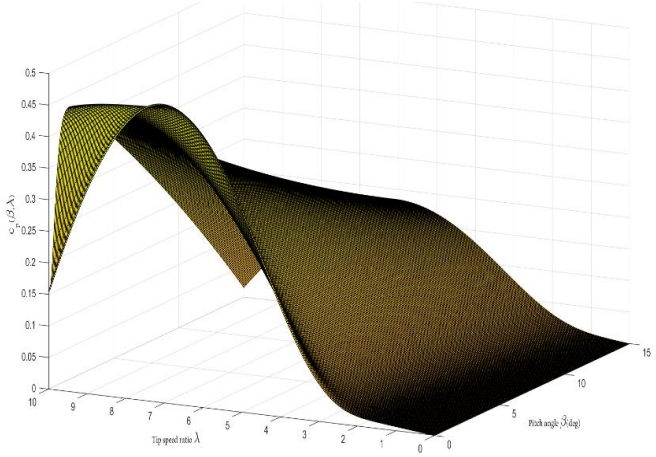


Fig 2: Wind Turbine power coefficient curve.

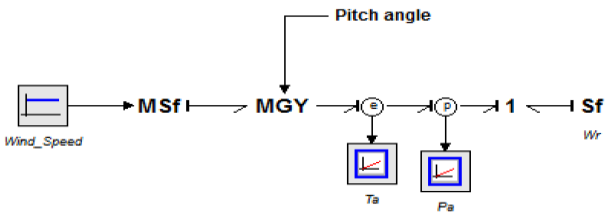


Fig 3: Bond graph Model of the Aerodynamic system.

3- *Mechanical subsystem*

The mechanical subsystem of our wind turbine contains two main parts ; the gearbox and gear set of the servo machine, in Fig 4 which shown the Mechanical subsystems model, the first input of the gearbox receives the aerodynamic power ***Pa*** delivered by the aerodynamics subsystems, the second input is meshed with servo machine by the third gear set.

The transient load and rotational speed component of the planetary gear transmission are shown in Fig 5, basing on the analysis of component which shown in fig 4, we gets the following question:

$$J_C\dot{\omega}_C = T_C - 3F_3R_C \qquad \{E.5\}$$
$$J_R\dot{\omega}_R = T_R - 3F_2R_R \qquad \{E.6\}$$
$$J_S\dot{\omega}_S = T_S - 3F_1R_S \qquad \{E.7\}$$
$$J_P\dot{\omega}_P = (F_2 - F1)\times R_P \qquad \{E.8\}$$
$$m_P\dot{\omega}_C R_C = F_3 - F_2 - F_1 \qquad \{E.9\}$$

The relationship between the angular speeds w_p, w_c and w_R can be obtained by the following equation:

$$w_P = \frac{w_s R_s - w_C(R_S + R_P)}{R_P} \qquad \{E.10\}$$

With: R_p, R_c and R_s are radius of planet gear, carrier and sun gear respectively. Fig 5 shows the bond graph model of the differential gear transmission.

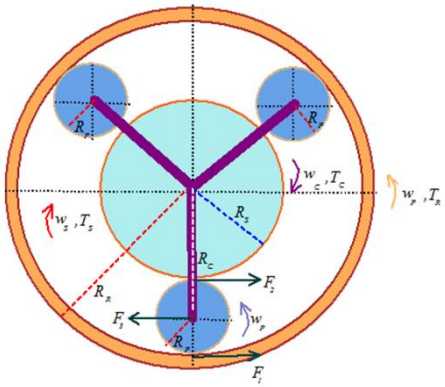


Fig 4: Velocity and internal load analysis of planet gear.

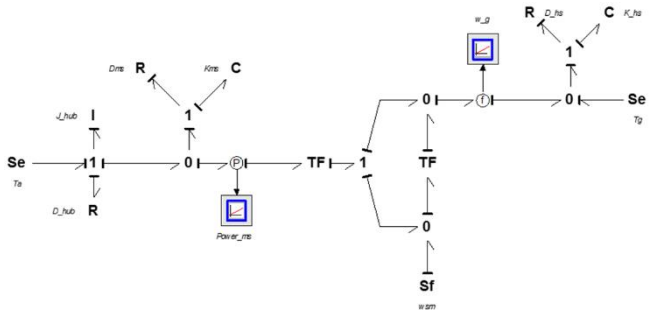


Fig 5: Bond graph model of the DGT.

4- Servo machine model

The bond graph model of the servo-machine system is depicted in Fig 6. The first part of the model contains the electrical subsystem is represented by the coil resistance Rind, which is represented by an R-element $R_{\_ind}$ and the motor inductance $L_{\_ind}$ represented by an I-element. The second is the mechanical subsystem is represented by I-element ($J_{\_SM}$: rotor inertia of servomotor) and R-element ($D_{\_SM}$: the friction at the rotor). Fig 6 shows the Bond Graph sub-model.

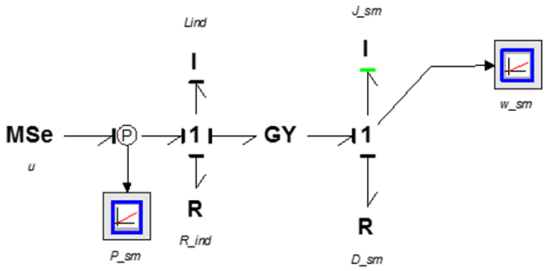


Fig 6: Bond graph model of the Servo machine.

III- PROPOSED CONTROL

In the wind energy conversion systems, one of the operational problems is the variability and discontinuity of the wind [10]. In most cases, wind speed can fluctuate rapidly. Hence, quality of produced energy becomes an important problem. Several control techniques have been designed to improve the quality of power generated from wind turbines [11].

In this paper this paper, we propose a Mechatronic Control Model of a Novel Variable Speed Wind Turbine Concept with Power Splitting Drive Train. In order to maintain the rotational speed of the asynchronous generator at the rated speed, Fig 16 describes the structure of the proposed controller. To minimize the difference between the actual speed and rated speed of generator this method proposed use an optimal proportional and integral PI controller. To achieve generator speed ωg at the optimal speed $\omega_{g\text{-}op}$, the proposed PI controller commands the servomotor to provide a suitable speed $\omega_{SM\text{-}ref}$, the referential speed of servomotor and dynamic error are described by the following equation:

$$\omega_{SM-ref}(t) = k_p e(t) + k_i \int_0^t e(\tau) d\tau_R \qquad \{E.11\}$$

$$e(t) = \omega_{SM-op} - \omega_{SM-ref}(t) \qquad \{E.12\}$$

Where; kp and ki are respectively proportional and integral gains of the PI controller.

To adjust the PI gains, we use the RBF neural networks, which gets the wind speed as input and gives the PI gains as output. In order to propose an optimal PI collective Servomotor controller, the RBF neural networks must be trained with optimal PI gains, the RBF parameters's is mentioned in table 2 (Appendix). Fig 7 describes the PSO-PI method [12] using to obtain an optimal training dataset for different wind speed. The objective function of PSO algorithm used in the model bond graph is an integral absolute error IAE, which calculated by using the following equation:

$$e(t) = \omega_{SM-op} - \omega_{SM-ref}(t) \qquad \{E.13\}$$

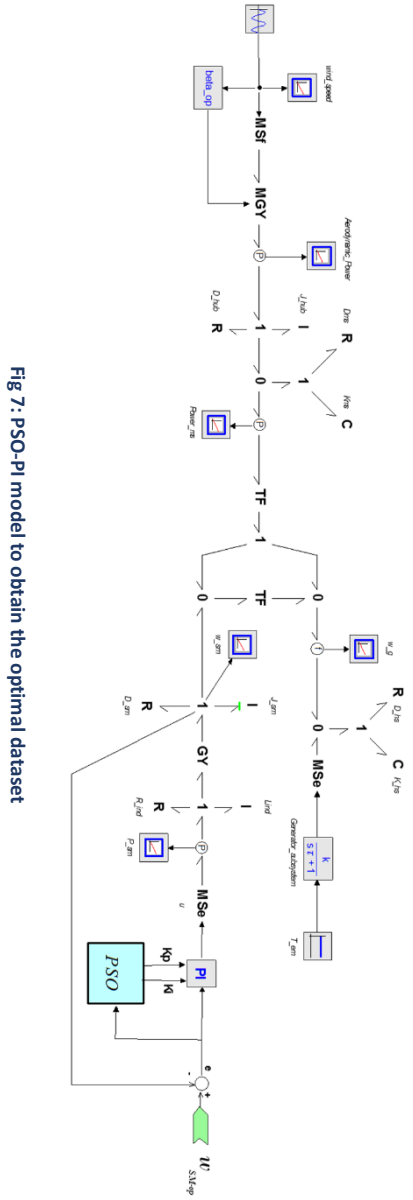


Fig 7: PSO-PI model to obtain the optimal dataset

IV- SIMULATION RESULT

In order to evaluate the performances of our Mechatronic model, and the proposed collective speed controller CSC, and the interactions between the different subsystems of the wind turbine, we apply a real data for 5MW wind turbine to our model [13].

Table 1 and Table 3 (Appendix) illustrates respectively all wind turbine's parameters values and PSO's parameters. To obtain an optimal dataset for RBF-training, Table 3 exposes the structure of RBF neural networks, which contains one neuron as input represents

the wind speed and two neurons as output; it provides PI gains, in addition ten neurons from the hidden layer. To train the RBF neural networks we use a dataset formed by 500 constants obtained from PI-PSO, ten of them are shown in Table 4.

Fig 9 shows a variable wind profile applied to our model shown in Fig 8, Fig 10 exposes the behaviors model for a real wind flew. We observe that when the wind speed is less than speed rate, the generator speed ωg follow the optimal speed which equal 116.34 rad/s, with error under 1e-3. The power coefficient C_p achieved at a maximum value which equal C_P=0,42. In the top of Fig 9, the hub torque oscillates between two values 99,556 KN*m and 127,334 KN*m. For the third curve, which shows the high shaft torque's variation illustrates that the torque is maintained proximately at 43,45 KN*m, we observe also that by using the proposed controller the power generator curves stay stable approximately on 5MW.

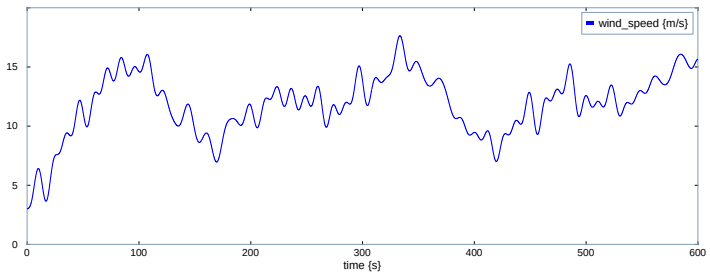


Fig 8: Wind Speed profile.

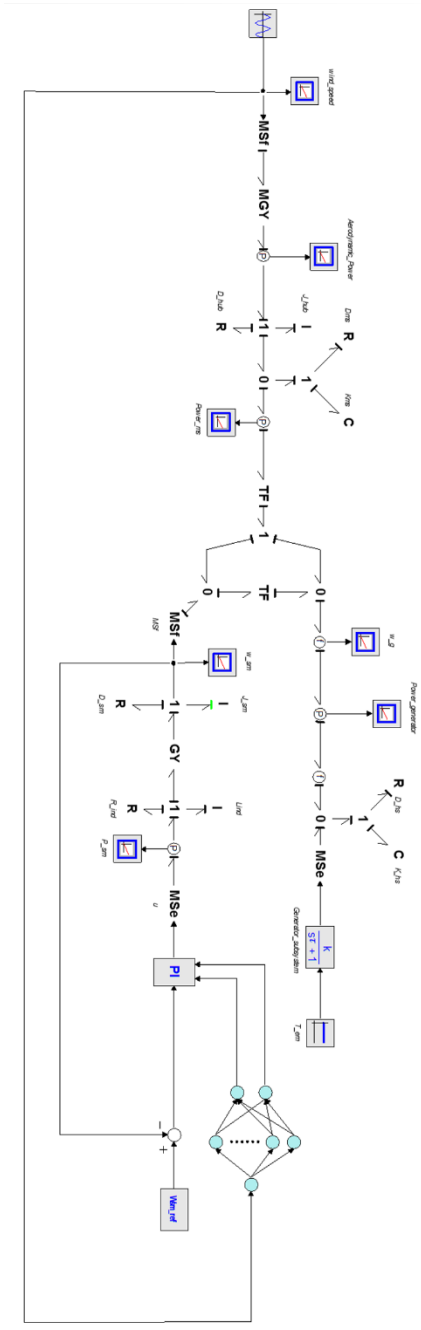


Fig 8: complete wind turbine model with proposed controller.

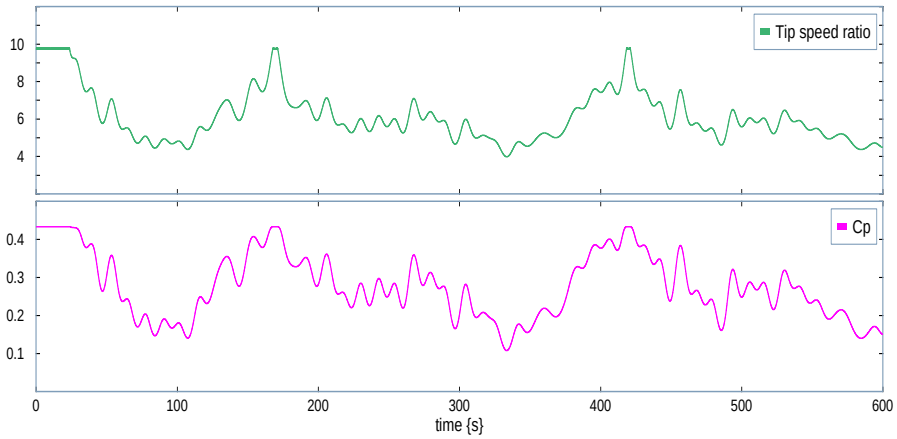


Fig 10-a: Tip Speed ratio and Power coefficient.

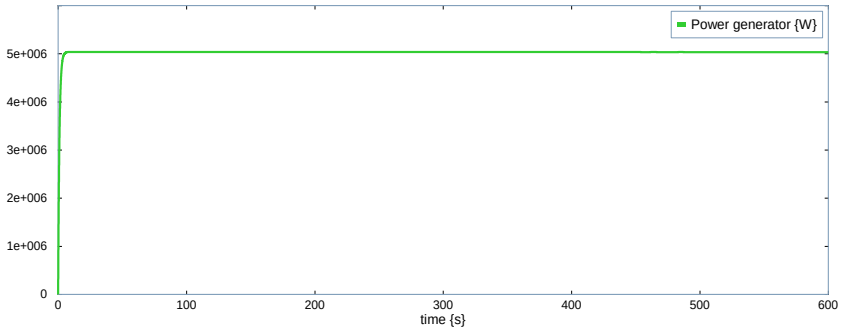


Fig 10-b: Power generator.

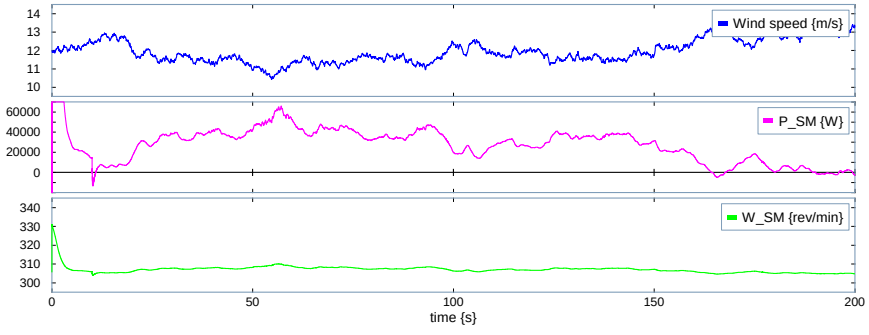


Fig 10-c : The Power and angular velocity curves of the Servomotor.

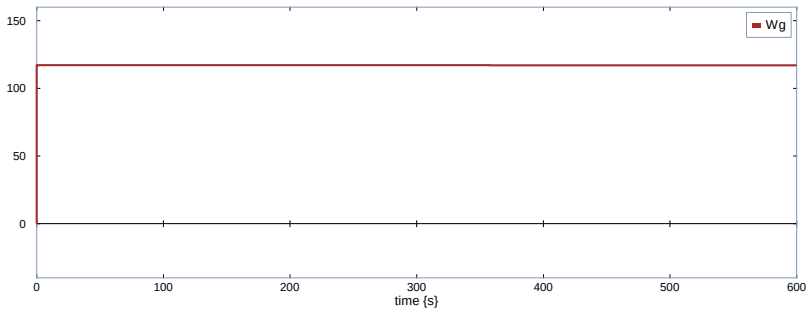


Fig 10-d: Angular velocity of generator {rad/s}.

V- CONCLUSION

In order to control the angular velocity of generator, and to reduce the losses and maintenance caused by the electronic converters, this paper proposes a novel strategy based on a RBF-PI controller of 5MW wind turbine class.

The RBF-PI controller is set up on the system including the servo machine for an optimal production point. It is worth mentioning that the global system is modeled by a mechatronic approach using Bond Graph representation. The simulation results of the proposed controller shows that the RBF-PI model has more performance, and helping the integrated model developed to provide an electrical power, more stable in every operation regions of the wind turbine.

Thanks to the simulation results, the proposed power split system can be used in areas where there is non-uniform wind profile in order to have a well-optimized production.

VI- APPENDIX

Table 2:

A portion of optimal dataset gated from PSO method.

Wind speed (m/s)	K_p	K_i	IAE {s}
7.1	1.23	10.3	4.56
8.5	3.44	12.33	2.341
9	-0.23	7.22	2.05
10.75	1.223	9.43	2.515
11	-0.334	10.4	1.25
11.5	-0.443	12.34	3.144
13	2.34	14.43	2.34
14.25	5.443	10.44	2.07
17	4.334	15.01	4.45
20	7.334	16.786	2.44

Table 1:

Parameter of wind turbine.

Parameter	Values
Power capacity	5 MW
Cut-in wind speed	2 m/s
Cut- out wind speed	25 m/s
Rated wind speed	12 m/s
Rotor radius	55 m
Rated generator speed	116.34 rad/s
Maximum power coefficient	0.434
Hub and main shaft	
J_{hub} , D_{hub} , K_{ms} , D_{ms}	5000 kg m^2, 1000 N/m, 3.67e^7 N/m, 200N/m.
High speed shaft	
K_{hs}, D_{hs}	10^7 N/m, 54.51 N.s/m.
Electrical Networks	
V_n ,f	460 V, 50 Hz,.
Servo motor	
J_{SM}, L_{ind}, R_{ind}, D_{SM}, K_{SM}	3.43kg.m^2, 4e^{-3} H, 3.6 Ω, 2.65e3 N.s/m, 0.4.

Table 3

Parameters of trained RBF neural networks.

Parameters	Values
Input neuron	1
Hidden neurons	20
Output neurons	2
Learning rate	0.4
Spread	0.74
Training time	10,5 s

VII- REFERENCE

[1] R Dev Shukla, R Kumar Tripathi. A novel voltage and frequency controller for standalone DFIG based Wind Energy Conversion System. Renewable and Sustainable Energy Reviews 37 (2014) 69–89.

[2] P baghri, Q sun. Nonlinear control of a class non-affine variable-speed variable-Pitch wind turbine with radial basic function neural networks. ISA Transaction 2022.

[3] Poultangari I, ShahnaziR, Sheikhan M. RBF neural network based PI-pitch controller for a class of 5MW wind turbines using particle swarm optimization algorithm.ISATrans2012; 51:641–8.

[4] J Mcmorland, A coraddu. A review of operations and maintenance modelling with consideration for novel wind turbine concepts. Renewable and Sustainable Energy. September 2022.

[5] Z khaouch, M zekraoui, M mabrouki. Mechatronic modelling of a 750kw fixed speed wind energy conversion system using the bond graph approach. ISA Transaction 2016.

[6] Y Liu, K Jiang. Numerical Study of the effect of surface grooves on the aerodynamic performance of a NACA 4415 airfoil for small wind turbine. Journal of wind engineering and industrial aerodynamic. November 2020,104263.

[7] Hoffmann MJ, Reuss Ramsay R, Gregorek GM. Effects of grit roughness and pitch oscillations on the NACA 4415 airfoil. The Ohio State University Columbus, Ohio; July 1996.

[8] S Zhao, Z Lui. Research on the rotor speed and aerodynamic characteristics of a dynamic yawing wind turbine with a short-time uniform wind direction variation. Energy,15 June 2022.

[9] Y Wang, X Sun, Y Jiang, L Wang. A new method for prediction of power coefficient and wake length of a horizontal axis wind turbine based on energy analysis. Energy Conversion and management, 15 January 2022,115121.

[10] F Sun, T Jin. A hybrid approach to multi-step, Short-term wind speed forecasting using correlated features. Renewable Energy, March 2022, Page 742-754.

[11] N Alessandro Vargas, L Acho. Optimal control of variable-speed wind turbine modelled as Markov jump systems. Journal of the Franklin Instistute.13 May 2022.

[12] R Sushmitha, S Chitra Devi, A Mangula, R Niraimathi. A novel methodology to capture the maximum power to capture the maximum power from variable speed wind turbines using fuzzy proportional integral controllers modified PSO-GSA algorithm. Materials Today: Proceedings. 13 February 2021.

[13] J.Y He, Y.C He, Q.S Li, L Zhang, H.L. Yang. Observation study of wind characteristic, wind speed and wind turbulence profile during super Typhoon Mangkhut. Journal of wind Engineering and Industrial Aerodynamics. November 2020, 104362.

Coverage Sets Scheduling for lifetime maximization of wireless sensor networks: Genetic algorithm approach

1st Ibtissam LARHLIMI
Department of Mathematics and Informatics
Polydisciplinary faculty, University Sultan Moulay Slimane
Beni-Mellal, Morocco
ibtissam.larhlimifpb@usms.ac.ma

2nd Hicham OUCHITACHEN
Department of Mathematics and Informatics
Polydisciplinary faculty, University Sultan Moulay Slimane
Beni-Mellal, Morocco
h.ouchitachen@gmail.com

3rd Hicham MOUNCIF
Department of Mathematics and Informatics
Polydisciplinary faculty, University Sultan Moulay Slimane
Beni-Mellal, Morocco
h.mouncif@usms.ma

***Abstract*—Wireless sensor networks (WSNs) are an evolving research area with a wide range of applications. In most utility applications, the energy resources of sensor nodes are limited. Therefore, it is important to manage the energy consumption effectively to increase the network lifetime. The challenge is to efficiently program sensor availability and activity and maximize network lifetime [20] for coverage. It has been shown that this problem is NP-hard [1], [11] and is called the maximum coverage set scheduling problem (MCSS) [15]. The objective of this paper is firstly, to propose a genetic algorithm for extending the network lifetime that guarantees finding a feasible scheduling for the collection of cover sets to maximize the network lifetime. Secondly, to compare it with the work reported by Chuanwen Luo, Yi Hong, Deying Li, Yongcai Wang, Wenping Chen, and Qian Hu (Greedy algorithm and MCSSA approximation algorithm). The simulation results show that the proposed algorithm also improves the network lifetime, which indicates the effectiveness of this algorithm.**

***Index Terms*—Network Lifetime, Wireless Sensor Network, Coverage Sets Scheduling, Genetic Algorithm, Greedy algorithm, Approximation Algorithm**

I. Introduction

Nowadays, innovation is progressing at a rapid pace. At first, sensors were simply used to measure and monitor a physical value at the level. With the advent of the internet and research into wireless technologies, these sensors were equipped with connectivity and gave rise to wireless sensor networks.
IoT is considered the next step in the evolution of the internet. It can connect and communicate via the Internet with almost any heterogeneous object in the real world to facilitate information sharing. With the help of wireless sensor network technology, the IoT can collect, analyze and deploy a huge amount of data that will be converted into meaningful information and knowledge. So wireless sensor networks (WSNs) play a critical role in supporting the IoT infrastructure.
The energy optimization approach we consider is based on inactivity/activity. In short, this paper proposes a method to form subsets of sensor nodes that will be successively activated for defining durations while the other nodes will be on standby, maximizing the duration. This problem, known as the Maximum Lifetime Coverage Problem (MLCP), is NP-hard and has been widely studied in the literature without taking into account the energy consumed by the sensors that are in standby mode [17], [18], [19].
The coverage problem can be split into two steps: The first step is to find as many coverage sets as possible for the sensors in the network, and the second step is to schedule the coverage sets obtained in the first step in a way that maximizes the lifetime of the network. Therefore, the maximum coverage set scheduling (MCSS) problem is important in all coverage problems. For a coverage set C = $(C_1, \cdots, C_m)$, in which each coverage set $C_j \in C$ consists of a set of sensors with sufficient battery power and can fully cover all targets (or the entire coverage area), the MCSS problem is to find the optimal scheduling strategy for the coverage set C to maximize the network lifetime [15].
In this work, we well study the MCSS problem to maximize the lifetime of the network (Section II), and then (Section IV) we well talk about the algorithms found in the literature to solve this problem: Greedy Algorithm and Probabilistic Algorithm, then (Section V) we well propose a Genetic Algorithm for this problem, and finally (Section VI), we make a comparison of the results obtained.

II. Related Works

The coverage in the WSN is often considered a very important performance and quality of service measure [22]. According to [1], it is, therefore, necessary to ensure that an event occurring at any point in the area monitored by the sensors is detected by at least one particular sensor. The

problem of coverage has attracted the attention of researchers and many approaches have been proposed, e.g. strategies for using nodes to solve the coverage problem, node scheduling schemes to extend the lifetime of the WSN while maintaining coverage, etc.
There are three typical types of coverage fig:1: Target(point) coverage, Area coverage, Barrier coverage [2], [3].

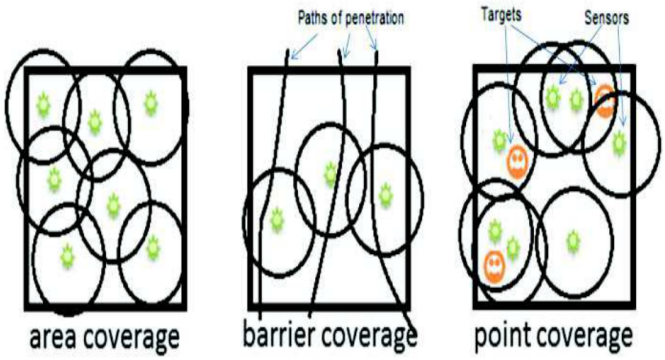


Fig. 1. Types of coverage [23].

A. Target coverage

Target coverage applications are the category of WSN applications where multiple discrete points of interest, called targets, need to be continuously monitored by the sensor nodes [17].
Since sensor nodes have a limited energy capacity, it is important to manage their energy consumption in a way that maximizes the lifetime of the network. Scheduling the appropriate subsets of active or dormant sensor nodes to maximize the lifetime of the network is called the maximum lifetime coverage problem (MLCP).
In [6], Lu and al. investigated the area coverage planning problem (MLCS) on a Euclidean plane considering target coverage and data acquisition and proposed a constant factor polynomial-time approximation algorithm for this problem.
In [2], the authors studied the MLCP problem considering the energy consumption of the sensors in standby mode. To solve the MLCP problem, they proposed a new iterative method based on the sensor level and the blacklist concept, which is a binary version of the relationship between a sensor and critical targets.

B. Area coverage

The most frequently used coverage in WSNs is the area coverage. The latter is always addressed for contiguous space. Therefore, the transformation of the area coverage problem into a discrete space problem is usually difficult, which complicates the solution of the area coverage problem. The coverage can be full or partial, depending on the requirements of the application.
Tao and al. presented a centralized area coverage enhancement algorithm to minimize the detection area overlap between directional sensors in a directional sensor network [7]. Tezcan and Wang proposed an algorithm for determining the coverage of a sensor and finding the sensor orientation to maximize the free coverage area for multimedia networks with directional sensors [8]. In [9], Gao and al. studied the problem of covering complex events to optimize the quality of coverage under the constraint that the total budget is not exceeded.
By considering the costs of different heterogeneous sensor deployments and the total budget for a monitored region. Wu and al. developed a new model called "full-view coverage" that requires selecting the minimum number of sensors to provide coverage for a given region of interest, and proposed a solution for each of two different deployment strategies [10]. In [21] they interested in solved this problem by used dynamic planning algorithm to increase the coverages area and covered the hole positions, the dynamic planning algorithm based on tracking the route to find the optimal path with the smallest cost.

C. Barrier cover

The goal of barrier coverage is to prevent undetected intrusions through the conceptual barrier formed by the sensors. Barrier coverage can be classified as weak or strong coverage [14]. In the case of weak coverage, the sensors must detect an object crossing the barrier, while in the case of strong coverage, we must detect any kind of suspicious behavior.
Kumar and al. first introduced the concept of barrier coverage. They then defined the k-barrier coverage notation and proposed algorithms to determine whether k-barrier coverage covers or not a belt area in sensor networks [9]. Dobrev and al. studied the algorithmic complexity of several natural generalizations of the barrier coverage problem with sensors of varying sizes.
In mobile sensors, li and Shen studied the problem of barrier coverage with these mobile sensors in a two-dimensional environment to form a specified line segment barrier while minimizing the maximum sensor movement [12].
Finally, Wang and al. have considered the barrier coverage problem for sensor position errors and presented the problem of formulating a barrier coverage with fault tolerance to establish a barrier with a guarantee [13].

III. Modelling and problem formulation

We use the following definition of WSN lifetime according to [16]: if that t_s is the period when the WSN starts its operation, and t_f is the period when the network is no longer covered, then the network lifetime is $TimeLife = t_f - t_s$.
To better understand, the coverage problem can be divided into two main phases:

- First: Find as many coverage sets as possible from the network sensors.
- Second: Take the obtained coverage sets and plan them to maximize the lifetime of the network.

Therefore, the maximum coverage set scheduling (MCSS) problem is important in all coverage problems. For a coverage set C = $(C_1, \cdots, C_m)$, in which each coverage set $C_j \in C$ consists of a set of sensors with sufficient battery power and

can fully cover all targets (or the entire coverage area), the MCSS problem is to find the optimal scheduling strategy for the coverage set C to maximize the network lifetime [15].

A. Mathematical Model

To formulate the MCSS problem as an integer linear programming (ILP) problem [15], the following terms are used: i: index of sensors, where $1 \leq i \leq n$ and j: index of the cover sets, where $1 \leq j \leq m$.
For each sensor s_i has b_i active time slots, the definition of the binary variable $a_{i,j}$ is as follows:

$$a_{i,j} = \begin{cases} 1, & \text{if } s_i \in C_j \\ 0, & \text{otherwise} \end{cases}$$

The definition of T_j is the active time of the cover set C_j, where $1 \leq j \leq m$. $(T_1, T_2, \cdots, T_m)$ represents the scheduling strategy for the collection of cover sets C. The MCSS problem can be formulated as the following integer linear programming (ILP):

$$max \sum_{i=1}^{m} T_j \tag{1}$$

Subject to:

$$\sum_{j=1}^{m} (a_{i,j} T_j) \leq b_i, \forall s_i \in S \tag{2}$$

Function 1 aims to maximize the total active time for the collection of cover sets C. Constraint 2 ensures that for each sensor $s_i \in S$, the total active time intervals in the scheduling strategy are less than or equal to the original active time intervals.

IV. Some algorithms proposed in the literature

We focused on the algorithms proposed by the researchers in [15] to solve the MCSS problem and investigated an algorithm called Greedy-MCSS. Based on this algorithm, we then studied an approximation algorithm, the MCSS algorithm (MCSSA).

A. Greedy algorithm

A Greedy Algorithm is an algorithm that follows the problem heuristic of making the locally optimal choice at each step, a greedy heuristic can produce locally optimal solutions that approximate a globally optimal solution in a reasonable amount of time. This algorithm is based on the propriety level method [15] of cover sets in C.

Demonstration: The algorithm iteratively selects a cover set C_j with the lowest priority level in C. Then C_j is executed E_j times, and $C_j$is added E_j times in CG sequentially, with $E_j = \min\{b_i | s_i \in C_j\}$ is the lifetime of C_j, and CG represents the scheduling strategy obtained by the algorithm [15].
Therefore: $TGlife = TGlife + E_j$
With: $TGlife$ is the working time of CG and its lifetime. The initial value of $TGlife$ is set to 0.
On the other hand: the active time slots of all sensors of C_j are updated, i.e. for each sensor $s_k \in C_j$, $b_k = b_k - E_j$.
Then, for each set of covers $C_j \in C$, the lifetime of C_j is updated, and the C_j will be deleted from C, if its lifetime is equal to 0.

B. Probabilistic Algorithm (MCSSA)

This is an algorithm based on the Greedy Algorithm, which allows finding a CG scheduling strategy, and the network lifetime, by performing the treatment we will see later. On a set of sensors $S = \{s_1, s_2, \cdots s_n\}$, where each sensor s_i has two states: inactive and active, and the set of coverages C, where each set of coverages C_j is a subset of S that can cover all targets [15].

V. Proposed approach: Genetic Algorithm

In this section, we propose an efficient algorithm for the MCSS problem, its objective is to improve the network lifetime by finding a scheduling strategy for the cover sets in C so that only one cover set is active at each time slot and the total active time of the cover sets in C is maximized. We propose a genetic algorithm for MCSS problem as follows:

Coding: The chromosome is represented as a vector of integer values of $_j$ between zero and m "Priority level".
The number of genes, which is the length of the vector, is equal to the number of covers $C_j \in C$, and each gene represents the set of covers C_j to be programmed as in table I.

TABLE I
Genetic algorithm chromosome creation.

Coverage set C_1	**Coverage set** C_2 | **...** | **Coverage set** C_m |

Initialization: The initial population of a limited number of chromosomes is randomly generated. A gene represents the cover setC_j. Therefore, the chromosome represents the scheduling strategy of the collection of cover sets C, the lifetime can be calculated by the sum T_j of the genes in the chromosome.

Fitness: The fitness must ensure that not all sensors in the candidate solution have exceeded the energy constraint. For a sensor $s_i \in S$, the sum of energy consumption across all covers in a schedule must not exceed the initial energy b_i(Section: III-A - Equation:2).

$$\sum_{j=1}^{m} (a_{i,j} T_j) \leq b_i, \forall s_i \in S$$

The fitness function will exclude the candidate schedule from the population that does not meet this constraint from future GA processes.

Selection: From the members of the adapted population, select the two best candidate programs with the maximum lifetime as parents ((Section: III-A - Equation:1)).

$$max \sum_{i=1}^{m} T_j$$

Crossover: An appropriate crossover is applied to generate children for the next generation and update the population. Double point crossovers are used. The crossover points are selected randomly, as is normally the case in GA.

Mutation: The appropriate mutation is applied to improve the children and avoid the local optimal solution. One or more genes are randomly selected and updated. We consider that enhancement increases the value of the genes, and thus the chromosome. The mutation genes are selected at random.

Fitness of children: The fitness function ((Section: III-A - Equation:2)) is reapplied for the "new chromosome" children to ensure that no sensor exceeds its initial energy. Parents are updated if the new and improved children have a better gene pool or a better lifespan.

VI. Simulation and Results

To analyse the performance of the proposed algorithm, we first compare it with the Greedy-MCSS algorithm presented in Section IV-A then with the MCSSA algorithm presented in Section IV-B. We simulate a network with N sensors randomly placed in a 1000 m × 1000 m area and detecting T targets. We apply the proposed algorithm to these coverage sets to measure the lifetime of the network. Each test result is an average of 100 runs. The MATLAB R2020 software was used for all simulations. In our simulation, we consider the following parameters (Table:II):

TABLE II
Parameters of simulations.

Length of chromosome	The scheduling strategy of the collection of cover sets C
Iteration	100
R (Sensing range of each sensor node)	200
T (Target)	10 to 20
M (Number of coverages sets)	100

As shown in Figure 2, we used T=10 targets and N=300 sensors, we assigned active time slots of each sensor between 10 and 30 at random, and varied the number of coverage sets from 10 to 60. We can notice that the network lifetime obtained by the proposed algorithm increases proportionally to the number of coverage sets. The results show that the lifetimes calculated by our algorithm are better than those obtained by the Greedy-MCSS and MCSSA algorithms. For example, in a network with 40 cover sets, we find the network lifetime found by our algorithm is 389s, and the lifetime calculated successively by the Greedy MCSS and MCSSA algorithm is equal to 170s and 184s.

In the second experiment, we used T=20 randomly distributed targets, M=100, and varied the number of sensors from 200 to 1000, with an increment of 200, for each sensor we set the time slot to 10. In Figure 3, we show our results in comparison with the Greedy MCSS and MCSSA algorithms [15], as a function of the number of sensors. We observe that the lifetime of the network increases in proportion to the number of sensors.Indeed, the fitness and selection part of the genetic algorithm selects the two best proposed solutions with the maximum lifetime as parents, so the possibility of combining the best parents to produce offspring with good genomes increases proportionally to the number of probes. The results show that the lifetimes calculated by the proposed algorithm are better than those obtained by the Greedy MCSS and MCSSA algorithms.

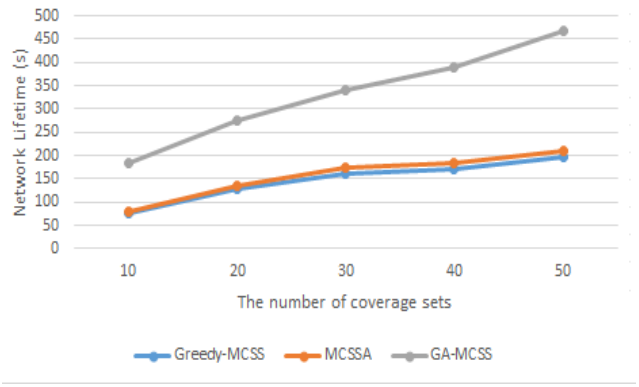


Fig. 2. varying the number of coverage sets from 10 to 60.

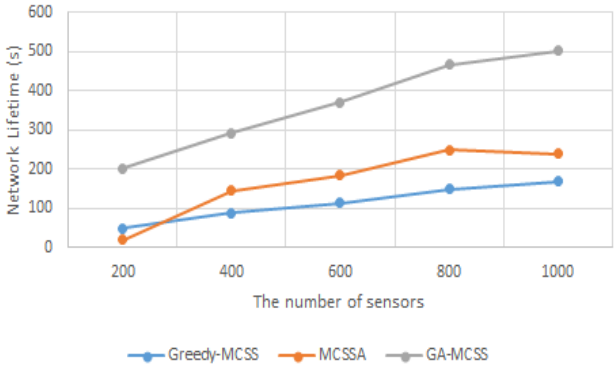


Fig. 3. Varying the number of sensors from 200 to 1000.

Figure 4 shows the lifetime calculated by our algorithm as a function of the active time slots of the sensors, in comparison with the two algorithms Greedy-MCSS and MCSSA [15]. For this, we set T=20, N=100, M=100 and change the active time slots of the sensors from [0, 10] to [50, 60], where [0, 10] means that the active time slots of each sensor are assigned from 1 to 10, randomly. The results show that the lifetimes calculated by our algorithm are better than those obtained with the two algorithms Greedy-MCSS and MCSSA thanks to the parts of the proposed algorithm that ultimately produce offspring with good genomes each time we select and combine the best choices as parents, thus obtaining parents with maximum lifetimes.

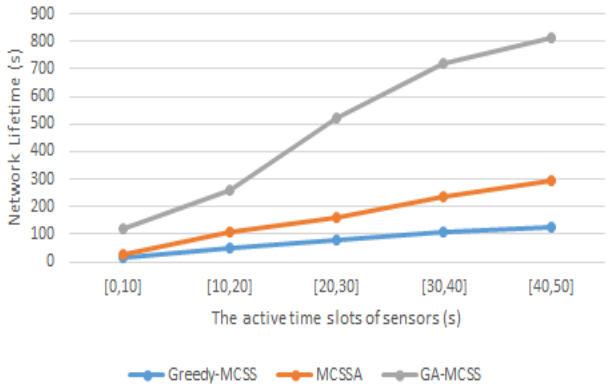


Fig. 4. The active time slots of sensors (s).

VII. Conclusion

In this paper, we study the MCSS problem, which involves finding the best optimal coverage and scheduling strategy to maximize the lifetime of the WSN. The MCSS problem has been formulated as an integer linear programming problem since it has been proved to be NP-hard. GA-MCSS is proposed as an algorithm to solve this problem and increase the lifetime of the network.

To demonstrate the effectiveness of the proposed algorithm, we compare it with the Greedy-MCSS algorithm and the MCSSA algorithm [15]. The simulation results show that the proposed Algorithm provides a longer network lifetime than the proposed algorithm in [15]. In fact, the network lifetime of the proposed algorithm increases significantly when the number of sensor nodes, coverage radius and activation interval are increased.

In the future, this study can be improved to analyze the impact of network and algorithm parameters (crossover, mutation) on the network lifetime and performance.

References

[1] Khoufi, I., Minet, P., Laouiti, A., Mahfoudh, S.: Survey of Deployment Algorithms in Wireless Sensor Networks: Coverage and Connectivity Issues and Challenges. In International Journal of Autonomous and Adaptive Communications Systems, vol. 10, no 4, p. 341 390. (2017)

[2] Wang, Y., Wu, S., Chen, Z., Gao, X., Chen, G.: Coverage problem with uncertain properties in wireless sensor networks: A survey. In Computer Networks, 123, 200–232. (2017). doi:10.1016/j.comnet.2017.05.008

[3] J. Amutha, S. Sharma, J. Nagar: WSN strategies based on sensors, deployment, sensing models, coverage and energy efficiency: review, approaches and open issues Wirel. Pers. Commun., 111 (2) (2020), pp. 1089-1115

[4] Cardei, M., Thai, M.T., Li, Y., Wu, W.: Energy-efficient target coverage in wireless sensor networks. In IEEE International Conference on Computer Communications (INFOCOM), vol. 3, pp. 1976-1984. (2005)

[5] Berman, P., Calinescu, G., Shah, C., Zelikovsky, A.: Efficient energy management in sensor networks. In Ad Hoc and Sensor Networks,Wireless Networks and Mobile Computing, vol. 2, pp. 71-90. (2005)

[6] Lu, Z., Li, W.W., Pan, M.: Maximum lifetime scheduling for target coverage and data collection in wireless sensor networks. In IEEE Transactions on vehicular technology, vol. 64, no. 2, pp. 714-727.(2015)

[7] Tao, D., Ma, H., Liu, L.: Coverage-enhancing algorithm for directional sensor networks. In Proceeding of the 2nd International Conferencen on Mobile Ad-Hoc and Sensor Networks, pp. 256-267. (2006)

[8] Tezcan, N., Wang, W.: Self-orienting wireless multimedia sensor networks for maximizing multimedia coverage. In Computer networks, vol.52, no. 13, pp. 2558-2567.(2008)

[9] Kumar, S., Lai, T.H., Arora, A.: Barrier coverage with wireless sensors. In Proceedings of the 11th annual international conference on Mobile computing and networking, pp. 284-298.(2005)

[10] Wu, P.F., Xiao, F., Sha, C., Huang, H.-P., Wang, R.-C., Xiong, N.X.: Node Scheduling Strategies for Achieving Full-View Area Coverage in Camera Sensor Networks. Sensors, 17(6), 1303.(2017). doi:10.3390/s17061303

[11] Sangwan, A., Singh, R.P.: Survey on coverage problems in wireless sensor networks Wirel. Pers. Commun., pp. 1475-1500. (2015)

[12] Li, S., Shen, H.: Minimizing the maximum sensor movement for barrier coverage in the plane. In IEEE International Conference on Computer Communications (INFOCOM), pp.244-252. (2015)

[13] Wang, Z., Chen, H., Cao, Q., Qi, H., Wang, Z., Wang, Q.: Achieving location error tolerant barrier coverage for wireless sensor networks. In Computer Networks, vol. 112, pp. 314-328, 2017

[14] Zhang, X. Fan, H., Lee, V.C.S., Li, M., Zhao, Y., Lin, C.: Minimizing the total cost of barrier coverage in a linear domain. In Journal of Combinatorial Optimization, pp. 1-24. (2018)

[15] Chuanwen, L., Yi, H., Deying, L., Yongcai, W., Wenping, C., Qian, H.: Maximizing Network Lifetime Using Coverage Sets Scheduling in Wireless Sensor Networks. In Ad Hoc Networks (2019). doi: https://doi.org/10.1016/j.adhoc.2019.102037

[16] Shi, T., Cheng, S., Cai, Z., Li, J.: Adaptive connected dominating set discovering algorithm in energy-harvest sensor networks. In IEEE International Conference on Computer Communications (INFOCOM), pp. 1-9. (2016)

[17] Thai, M.T., Wang, F., Du, D.H., Jia X.: Coverage problems in wireless sensor networks: designs and analysis. In International Journal of Sensor Networks, vol. 3, no 3, p. 191. (2008)

[18] Fan, G., Jin, S.: Coverage Problem in Wireless Sensor Network: A Survey. In Journal of Networks, vol. 5, no 9, p. 1033 1040.(2010)

[19] Cardei, M.M.T., Thai, Y., Li, Wu, W.: Energy-efficient target coverage in wireless sensor networks. In Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies, Miami, FL, USA, p. 1976–1984. (2005)

[20] Vijayalakshmi K., Anandan P.: Global levy flight of cuckoo search with particle swarm optimization for effective cluster head selection in wireless sensor network Intell. Autom. Soft Comput., 26 (2) (2020), pp. 303-311

[21] Alhaddad, Z.A.,Manimurugan S. :Maximum coverage area and energy aware path planner in WSN,nMaterials Today: Proceedings,(2021)

[22] Elif Zeynep Serper, Ayşegül Altın-Kayhan: Coverage and connectivity based lifetime maximization with topology update for WSN in smart grid applications,Computer Networks (2022)

[23] ami Mnasri, Nejah Nasri, Thierry Val: The Deployment in the Wireless Sensor Networks: Methodologies, Recent Works and Applications. International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks (2016)

Cybersecurity Virtual Labs for Online Learning: Current Trends and Recommendations for Improvements

Abdeslam Rehaimi
Laboratory LIMATI, FPBM
Sultan Moulay Slimane University
Beni Mellal, Morocco
abdeslam.rehaimi@usms.ac.ma

Yassine Sadqi
Laboratory LIMATI, FPBM
Sultan Moulay Slimane University
Beni Mellal, Morocco
yassine.sadqi@ieee.org

Yassine Maleh
Laboratory LaSTI, ENSAK
Sultan Moulay Slimane University
Beni Mellal, Morocco
yassine.maleh@ieee.org

***Abstract*—Currently, there is a significant need for modern approaches for cybersecurity education. Virtual laboratories have been introduced as a practical solution to complement theoretical education and enhance students by applying hand-on learning. In this paper we present an overview of both virtualization and containerization based portable solutions, which are widely used for developing cybersecurity labs. In addition, a comparative study of current solutions is presented. Finally, we discuss the challenges and recommendations to follow while designing virtual cybersecurity labs. The findings of our study can be useful to both scholars and practitioners interested in improving online learning.**

***Index Terms*—Cybersecurity, Virtual Laboratory, Hands-on Labs, Virtualization, Containerization**

I. Introduction

Designing a portable cybersecurity virtual laboratories is one of the required skills by educational institutions recently. Nowadays, the wide availability and use of online services has made the world increasingly interconnected. Thus, the counted number of severity of cybercrimes have increased in almost real-time[1][2]. As a result, the demand for cybersecurity specialist to stand against such attacks is also grown considerably. In 2019 ENSIA[3] has published a report in which they have founded that a set number of challenges are negatively impacting cybersecurity education. They are: Education institutions lack of cybersecurity staffing and educators; a mismatch between educational programs and job market demands because of the poor interaction with the industry; Obsolete education tools which focuses only on theoretical-based education instead of learning by practice.

Due to the favorable usability of virtual laboratories in education field, it has known a big consideration in the last decade by universities and companies as an effective tool to enhance student's learning[4]. In fact, the practical activities of cybersecurity increasingly bring opportunities for all type of learners to see and practice by doing, which is very beneficial [5]. In 2020 with the spread of COVID-19, schools and universities were all shifted to emergency remote education, as a result the need for portable virtual laboratories as a tool for students was a main difficult task faced to complement online course teaching and provide learners with pr-econfigured virtual labs to perform hands-on tasks by practicing legally on a secure and isolated environment [6]. Furthermore, Zeng et al [4] in there study found that providing lab material such as a user guide, video tutorials, or presentations plays a significant role to increase students retention than acquired learning by only practice.

Recently with the world digitalization movement in many fields, cybersecurity has become an extremely critical concern for many organizations, companies, and governments [1] that are unprepared for a sudden data breaches. Indeed, education and training is one of the key aspects to stand against such thing by understanding the latest defenses and threats to apply prevention techniques against it whenever it is faced [10] [11]. Virtualization and containerization-based virtual laboratories is a very promising tool for future specialists in this field by focusing on practical learning[9] as a complementary to theoretical-based education. The main contributions of this work are as follows:

- Presents an overview of the virtualization and containerization technologies used for designing portable virtual laboratories.
- Presesnts a comparative analysis between those two technologies.
- Gathers a combination of efficient variables to consider while designing a particular cybersecurity laboratory.

This paper is organized in sections as follows: (S1) the introduction; (S2) the related work; (S3) an overview of virtualization and containerization technologies; (S4) a comparaison between virtual machine and containers; (S5) Factors of developing an efficient hands-on cybersecurity lab; and (S6) the conclusion.

II. Related Works

There are many papers, which are discussing both virtualization and containerization technologies for designing cybersecurity labs in education. Catuogno et al. [10] proposed a methodology for measuring the effectiveness of using the

isolation properties of a containerized virtualization system against high resource consumption caused by malicious malwares, in which, the main host operating system might suffer an unwanted computing resource consuming due to suspicious services running. Karagiannis et al. [11] in their study have measured the computing resources utilization such as CPU, RAM, Disk I/O to demonstrate the benefits of using containerization technology's less overhead compared to the use of virtualization technology to deploy cybersecurity exercises.

Zeng et al [4], in there study used a web-based virtualized platform for cybersecurity education as they combine three virtual machines to construct a server and client web-based application in which students can access to the deployed lab through the internet. In their study they have also evaluated students' performance toward performing the lab as they find that students knowledge is increased notably by the factor of reading lab material than acquired knowledge while performing the lab itself. In contrary to the use of Virtual Machine Shahriar et al [5], proposed an approach of delivering only the application layer containerization of machine learning algorithms and datasets through Docker containers instead of running it in a virtual machine as they claimed that dockerized applications use host computing resources more efficiently than virtual machines. There proposed approach reduces troubles for students such as configuration coding, searching for dataset so they focus only on the content instead of configuration issues.

Significant efforts have been also done in the past to provide portable virtual laboratories bundled with multiple cybersecurity exercises. SEED Labs[12], for example, provides prebuilt virtual machine images with over 30 exercises covering a broad range of cybersecurity topics. Similarly, Labtainer [13], provide 50 cybersecurity lab exercises within a secure and consistent execution environments in which students complete labs totally on their computer as they provide docker installation and pre built virtual machine images, regardless of the Linux distribution and packages installed on the student's workstation or virtual machine hypervisor installed.

III. Portable cybersecurity tools

Setting up and configuring an effective environment for cybersecurity laboratories is highly on demand and a subject study by many researchers recently, as the number of cyberattacks grows the need for specialist also grows. Nowadays, various tools exist to develop and train successful cybersecurity professionals (e.g., Vulnerable virtual machines that can be deployed and exploited or portable docker cotainers that containerized every component by sandboxing them utilizing either OpenVZ[14] or Docker[15]). Thus, this makes it simple to publish or distribute the preconfigured portable files among learners using usb flash drive, hosting at college facilities or deployed in online cloud providers. Learners are then requested for tasks to be accomplished. most common tools widely used for educational purposes are as follows[16] [17] [18]:

- **Virtual Machines :** Virtual Machine is a concept that was developed back in the 1960's[19] that is widely used testing environment emulating the main physical host. Performing a cybersecurity lab in a secure condition to understand how an intrusion sample such as malwares interacts with the file system without harming the actual physical host, a virtual machine is utilized as an ideal solution for this scenario. Replication and isolation between the VM and the main physical host and also between it and other VMs [20] is it's main feature.
- **Containers :** Docker containers are used to automate a complete lab execution environment, thereby, reducing lab setup and configuration distractions. By using containers, lab designers can incorporate complex topologies without suffering the overhead as usually resulted from running multiple VMs [21] [22] as it delivers only the software virtualized application.

The fact that replicating multiple instances of an operating system or application on the same physical host provide a realistic emulation environment while remaining highly secure. Thus virtualization and containerization are exceptionally a handy tools in designing cybersecurity hands-on labs for educational purposes.

IV. Virtualization and containerization Technologies

According to Sahoo et al [20], three virtualization technologies are widely used: hypervisor-based provisioning for both full virtualization and paravirtualization and the container-based virtualization [10]. In this section, the three virtualisation technologies are been discussed.

A. Virtualization Technology

Virtualisation is a technique consist of provisioning resources by creating multiple virtual machines (VMs) on a single physical machine. Those VMs are manipulated using Virtual Machine Monitor (VMM) also known as a hypervisor laying between the host and guest operating system. The hypervisor is a tool that is used for virtualisation purpose as it only and mainly control the state of the VMs (e.g., power on and power off) and adjusting the amount of shared ressources given to each individual VM such as memory, computing processing speed, storage, and the networks without affecting the main host operating system. The virtualization consist of four main advantages such as virtualizing: 1) hardware virtualisation as required to run the guest operating system 2) software virtualisation wich runs an isolated software application on the guest isolated form the main host 3) storage virtualisation and lastely 4)Network virtualisation.

As shown in Fig. 1. virtualisation exist as two mechanisms like server virtualization (i.e., full virtualization) and desktop virtualization (i.e., paravirtualization) [23] each one of them with it's own feature as follows:

- **Full virtualization (Type 1) :** Type 1 (bare-metal hypervisors) with a purpose of virtualizing the undelaying

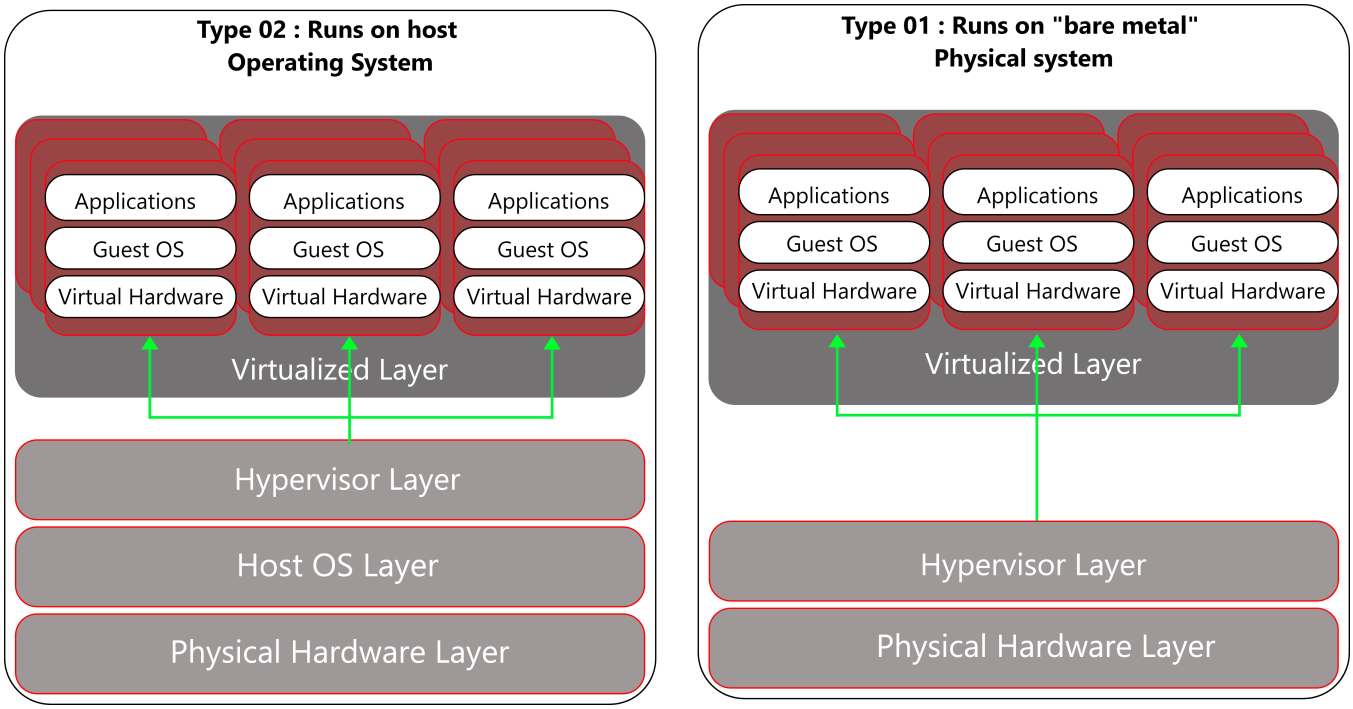


Fig. 1. Virtualization architecture.

physical hardware layer. This scenario type usually performs better in terms of speed and security because of less layer complexity. The physical server gets boots up wih this type of hypervisor.

- **Paravirtualization (Type 2) :** Type 2 (hosted hypervisors) is a form of software that does not require any specific hardware support for its execution as it is installed on top of the operating system. The two most popular used softwares VMWare and VirtualBox, are widely adopted by various organizations and individuals for training and testing purposes.

B. Containerization Technology

Containerization is an application layer virtualization concept that uses images and containers to run multiple distributed applications on top of the host operating system. Unlike hypervisor virtualization, as virtual machines run virtually on the physical hardware via an intermediate layer, containers runs in the user space on top of an operating system's kernel. Thus, Docker handles the deployment of an unlimited isolated user space instance applications at the same OS level[18].

As shown in figure 2 Containerization solutions provides many advantages in terms of applications deployment as it encapsulates projects into so called containers. Containers depend on one single host and use the same kernel. Thus, running multiple application instances can be performed on a single host and use it's same kernel operating system at the host level. To launch and manage containers, users' need to install the docker engine on the host[18].

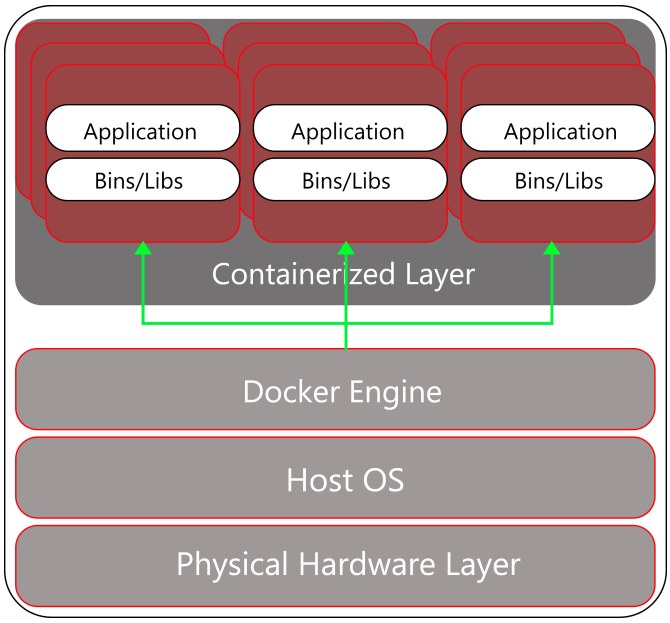


Fig. 2. Architecture of Docker.

Docker [24] is an open source tool [25] for OS-level container-based virtualization technologie that has been adopted by most cloud providers such as Google and Amazong[10]. Docker is a container engine based on LXC (Linux Container) developed and opensourced by the team at Docker, Inc (DotCloud) [26] and released by them

under the Apache 2.0 license [27].There are four main internal components of Docker, including Docker Client and Server, Docker Images, Docker Registries, and Docker Containers. These components constructs Docker's main architecture[27][28].

- **Docker Client and Server :** Is where the communication exchanged between docker server (i.e., Docker daemon) and clients (i.e., users) through API's. Docker daemon manage images, containers, networks, and volumes by listening for Docker API calls. Docker also allows the communication between client and multiple daemon's as also support's communication between multiple daemons to manage desired Docker services[28].
- **Docker Images :** Docker images can be refered to as "source code" of containers[27]. That is This is a read-only template providing instructions to create Docker containers[28] (e.g., add a file, run a command and open a port[27]).
- **Docker Registries :** Docker images are stored in a Docker registry. Docker Hub is a public registry that anyone may use, and Docker is set up by default to seek for images on Docker Hub [28].
- **Docker Containers :** is by default isolated from other containers, it's purpose considered as a runnable instance of an image[28].

V. Comparison between Virtual Machines and Docker Containers

Docker and traditional virtualization technologies are both based on the isolation and reallocation of the host resources, but Docker's containerization feature is slightly different from the traditional ones (i.e., VMs) [26]. Additionally, containers are more efficient and portable than VMs especially in terms of disk space utilization, and when deploying the same lab application to different environment [18]. For example, a class with more than 60 students, running individual VMs deployed by the school facilities requires too much computing processing and disk space utilization. As an example, if a VM requires 10 Gigs of storage, the total disk required for all students is 600 Gigs, without even mentioning the overloads. In contrary, Docker launch and deploy lightweight applications that runs and function properly [24] without requiring each application to have its own virtual machine environment because of its OS-level container-based virtualization which reduces the overhead[29] and it's lightweight features that requires only a couple of megabytes.

However, hardware virtualization is the top recommended virtualization for cybersecurity, among application virtualization as it is required for some sorts of attack types testing such as : malware, security breaches, network attacks, and other intrusions. Therefore, providing enough resources must be taken in consideration. In contrary some laboratories type requires less resources than VMs as it only needs the application (software) virtualization level which is sufficient due to the fewer resources that is required to run containers.

To Sum up, Hypervisor-based provisioning modules and Docker are two most widely used provisioning engine tools been in study in the last decade. Indeed, Docker is a promising solution right now because of the advantages of the state-of-the-art containerization technique of providing efficient fast container provisioning. VMs is no doubt still in its place because it's full virtualization feature and it's reliable and security which still in discussion in Docker, while VMs provision the guest operating systems Docker does not require this feature for containers which results in reducing both provisioning time and disk space compared to hypervisor-based provisioning modules.

As a result, we believe that the nature and scenario of the experiment for designing a virtual lab, and the amount of resources can be provided by educational institutions plays a significant role in making a final decision wither choosing to go with containerization or virtualization in designing a portable cybersecurity laboratories.

VI. Developing an efficient hands-on Cybersecurity Lab

Designing and developing a successful cybersecurity lab requires some significant efforts from the instructors and lab designers. Further, there is many requirements needs to be satisfied such as a good strategy to follow with less complexity as possible, Furthermore, a main focus on increasing student's retention, engagement and learning perspective that can encourage learner's to relate their experiments experience to industry practice.

To overcome the above-listed challenges and satisfy the requirements. In this section, we have combined and gathered multiple factors can be helpful for the researchers and instructors to consider while designing a particular cybersecurity lab.

Security: Security concerns and isolation shall be at its highest priority while designing a cybersecurity laboratory. A lab should guarantee a complete isolation from the main host operating system of the physical machine. As an example, when performing a blind analysis of malwares, it can dangerously harm the operating system of the potential learners or the physical hardware itself. Thus, sandboxing the environments helps to perform the lab remotely. Further, lab instances must be isolated so that the network traffic generated does not interfere with each other.

Documentation: Providing a well-documented documentation is highly recommended, it can be a document in which it describes how lab was designed (e.g., tools installed, configurations needed, important libraries required, and a well-documented source code) helps the designers after a period of time understand what have been done before, whenever the designers wants to improve the lab that document becomes handy to avoid further mistakes and complications encountered in the past. A user guide is also needed as it contains recommended instructions to follow while performing a particular lab. In fact, a study have been made by Zeng et al [4] shows that "reading lab materials plays a more significant role in hands-on learning than the acquired learning while working on lab tasks".

Licensing : Performing cybersecurity lab requires tools, and tools can be either commercial or open source, the main concerns of the creators of the lab is whether they need professional paid versions or open source alternative. Many commercial companies now provide academic licensing for students and for universities, sometimes it can be free and sometimes not. Today with the world changing and due to the limited budget universities cannot be able to afford or provide multiple licenses for the massive amount of students registered each year. For this reason, opens source is still preferred. Indeed, commercial products worth mention that they provide troubleshooting support for their products compared to open source that those not which requires manual intervention by lab creators in case any issues encountered.

Availability : A suitable place to store and save the pre-configured labs is also worth to mention. It can be uploaded on the university website or deployed on a cloud provider. Downloading lab anytime needed must be guaranteed for a long time period. Labtainer framework is an example that includes more than 50 cybersecurity lab exercises, labs are available for download online [13] They provide also a Docker image in wich the containers are hosted on the public Docker hub [30] and also two VM disk images for both options to be used by choice on VirtualBox or VMware hypervisors.

Learning perspective : It is widely approved by educators that "learning by doing [9]" is one of the most effective ways to learn[31]. Ma et al [32] investigated several laboratory instructional approaches and concluded that remote labs assist students develop designing ability and reasoning skills. According to Deshmukh et al [33] Laboratories should not be tied to a single technology, but rather should focus on many different aspects such as concepts, methodologies, and skills that guarantee full understanding of what is happening.

Reflective Practice : The main idea of reflective practice is the ability of students to develop and aquire knowledge through the lab by acquiring problem-solving skills and self-confident in decision-making [34]. According to Hinett's[35] four-step reflective approach that enables: 1) comprehend what they already know about the environment and tools they use 2) determine what they want and need to know in order to progress and accomplish the work 3) make logical sense of incoming information and knowledge in the perspective of their very own experiences 4) relevant potential learning possibilities.

Watermark : Unique watermark is preferably generated at start of performing the lab. Students watermark can be through a login access using a unique id for each user which later on be validated as part of the instructor-initiated assessment process. Further, watermarks help's to track student's activity which is designed for getting a diploma or qualification completion.

Traceability : Tracking student's activity through the process of completing a lab can be used as part of evaluation process as it is a good approach that makes development of cheating automations more challenging as it can be easily detected. In fact said by [36] cheating becomes a massive issue in schools when Academic Resources Information Service is used for learning.. According to a study done by Whitley [36] when students have been cheated on in the past, they are more likely to cheat again which considered as large effects. When they believe they aren't performing well or are performing poorly in a given task which considered as medium effect. While other reasons such as less parental control, particularly when a student lives far away from his parents, may also encourage students to think of cheating which considered as small effect [36].

Evaluation : at final stage students shall deliver a result which can be as a form of report a paper, presentation of the solution, or a project file depending of the lab nature. Thinking of giving a unique generated result for each user can prevent and detect wither the submission is really belonging to the user himself or not. As a scenario example, while performing a lab for computer networks students shall be given only an ip address to perform an SSH connection to a server using a given password, indeed, the ip address of the server and passwords are same for all students but they may be requested to use a particular tool to locate the port number wich assigned randomly and must be arbitrary given within a range to each student. As a result, guessing repeatedly is insufficient to finish the lab [22].

VII. Conclusion and Future Works

In this paper, we reviewed virtualization-based and containerization-based labs for cybersecurity learning. Next, we compared both virtual machines and Docker portable solutions in terms of flexibility, overhead, environment isolation, and their specific features. Finally, we presented several key aspects to be considered by instructors and education institutions while designing an efficient cybersecurity laboratory for educational purposes. Future research is needed to design suitable virtual labs in developing countries and explore its impacts on education. In the near future, we intend to use the findings of our study to create improved virtual laboratories for cybersecurity and computer science education in underdeveloped countries.

References

[1] K. Cabaj, Z. Kotulski, B. Ksiezopolski, and W. Mazurczyk, "Cybersecurity: trends, issues, and challenges," EURASIP J. Inf. Secur., vol. 2018, Dec. 2018, doi: 10.1186/s13635-018-0080-0.

[2] K. Cabaj, D. Domingos, Z. Kotulski, and A. Respício, "Cybersecurity education: Evolution of the discipline and analysis of master programs," Comput. Secur., vol. 75, pp. 24–35, Jun. 2018, doi: 10.1016/j.cose.2018.01.015.

[3] T. De Zan and fabio franco, "Cybersecurity skills development in the EU: The certification of cybersecurity degrees and ENISA's Higher Education Database," Mar. 2020.

[4] Z. Zeng, Y. Deng, Ih. Hsiao, D. Huang, and C.-J. Chung, "Improving student learning performance in a virtual hands-on lab system in cybersecurity education," in 2018 IEEE Frontiers in Education Conference (FIE), San Jose, CA, USA, Oct. 2018, pp. 1–5. doi: 10.1109/FIE.2018.8658855.

[5] H. Shahriar, K. Qian, and H. Zhang, "Learning Environment Containerization of Machine Leaning for Cybersecurity," in 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC), Madrid, Spain, Jul. 2020, pp. 1131–1132. doi: 10.1109/COMPSAC48688.2020.0-105.

[6] B. Ksiezopolski, D. Rusinek, M. Miśkiewicz, and A. Wroblewska, Hands-on Cybersecurity Labs in online learning. 2021.

[7] R. F. Babiceanu and R. Seker, "Big Data and virtualization for manufacturing cyber-physical systems: A survey of the current status and future outlook," Comput. Ind., vol. 81, pp. 128–137, Sep. 2016, doi: 10.1016/j.compind.2016.02.004.

[8] M. D. Workman, J. A. Luevanos, and B. Mai, "A Study of Cybersecurity Education Using a Present-Test-Practice-Assess Model," IEEE Trans. Educ., vol. 65, no. 1, pp. 40–45, Feb. 2022, doi: 10.1109/TE.2021.3086025.

[9] L. Carlson and J. Sullivan, "Hands-on Engineering: Learning by Doing in the Integrated Teaching and Learning Program*," undefined, 1999, Accessed: May 16, 2022.

[10] L. Catuogno, C. Galdi, and N. Pasquino, "An Effective Methodology for Measuring Software Resource Usage," IEEE Trans. Instrum. Meas., vol. 67, no. 10, pp. 2487–2494, Oct. 2018, doi: 10.1109/TIM.2018.2815431.

[11] Karagiannis, S., Magkos, E., Ntantogian, C., Ribeiro, L.L. (2020). Sandboxing the Cyberspace for Cybersecurity Education and Learning. In: , et al. Computer Security. ESORICS 2020. Lecture Notes in Computer Science(), vol 12580.

[12] "SEED Project." https://seedsecuritylabs.org/ (accessed May 13, 2022).

[13] "Labtainers - Center for Cybersecurity and Cyber Operations - Naval Postgraduate School." https://nps.edu/web/c3o/labtainers (accessed May 08, 2022).

[14] "Open source container-based virtualization for Linux.," OpenVz. https://openvz.org/ (accessed May 10, 2022).

[15] "What is a Container? - Docker," Nov. 11, 2021. https://www.docker.com/resources/what-container/ (accessed May 10, 2022).

[16] N. Dutta, N. Jadav, S. Tanwar, H. K. D. Sarma, and E. Pricop, "Design of a Virtual Cybersecurity Lab," in Cyber Security: Issues and Current Trends, N. Dutta, N. Jadav, S. Tanwar, H. K. D. Sarma, and E. Pricop, Eds. Singapore: Springer, 2022, pp. 143–157.

[17] "Comparison of platform virtualization software," Wikipedia. Apr. 04, 2022. Accessed: Mar. 22, 2022.

[18] A. Alsalamah and J. Kelly, Applying Virtualization and Containerization Techniques in Cybersecurity Education. 2018.

[19] P. M. Chen and B. D. Noble, "When virtual is better than real [operating system relocation to virtual machines]," in Proceedings Eighth Workshop on Hot Topics in Operating Systems, Elmau, Germany, 2001, pp. 133–138. doi: 10.1109/HOTOS.2001.990073.

[20] J. Sahoo, S. Mohapatra, and R. Lath, "Virtualization: A Survey on Concepts, Taxonomy and Associated Security Issues," in 2010 Second International Conference on Computer and Network Technology, Bangkok, Thailand, 2010, pp. 222–226. doi: 10.1109/ICCNT.2010.49.

[21] "Docker: Automated and Consistent Software Deployments," InfoQ. https://www.infoq.com/news/2013/03/Docker/ (accessed May 07, 2022).

[22] M. F. Thompson and C. E. Irvine, "Individualizing Cybersecurity Lab Exercises with Labtainers," IEEE Secur. Priv., vol. 16, no. 2, pp. 91–95, Mar. 2018, doi: 10.1109/MSP.2018.1870862.

[23] N. Gephart and B. Kuperman, "Design of a virtual computer lab environment for hands-on information security exercises," Journal of Computing Sciences in Colleges, 2010.

[24] "Home - Docker." https://www.docker.com/ (accessed May 07, 2022).

[25] "Open Source - Docker," Oct. 08, 2021. https://www.docker.com/community/open-source/ (accessed May 15, 2022).

[26] H. Ju, J. Wang, E. Zhu, X. Zhang, and F. Zheng, "Design Scheme of a Docker Container File Isolation against Computer Virus Spreading," Math. Probl. Eng., vol. 2022, p. e5348370, Mar. 2022, doi: 10.1155/2022/5348370.

[27] J. Turnbull and an O. M. C. Safari, The Docker Book. 2014. [Online]. Available: https://www.oreilly.com/library/view/the-docker-book/9780988820203/

[28] "Docker overview — Docker Documentation." https://docs.docker.com/get-started/overview/ (accessed May 18, 2022).

[29] J. T. Lim and J. Nieh, "Optimizing Nested Virtualization Performance Using Direct Virtual Hardware," in Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems, New York, NY, USA: Association for Computing Machinery, 2020, pp. 557–574.

[30] "labtainers's Profile — Docker Hub." https://hub.docker.com/u/labtainers (accessed May 16, 2022).

[31] M. Lombardi, "Authentic Learning for the 21st Century: An Overview," Jan. 2007.

[32] J. Ma and J. V. Nickerson, "Hands-on, simulated, and remote laboratories: A comparative literature review," ACM Comput. Surv., vol. 38, no. 3, pp. 7-es, Sep. 2006, doi: 10.1145/1132960.1132961.

[33] P. P. Deshmukh, C. D. Patterson, and W. T. Baumann, "A hands-on modular laboratory environment to foster learning in control system security," in 2016 IEEE Frontiers in Education Conference (FIE), Oct. 2016, pp. 1–9. doi: 10.1109/FIE.2016.7757669.

[34] J. Mezirow, "How Critical Reflection Triggers Transformative Learning," 2002.

[35] K. Hinett and T. Varnava, Developing reflective practice in legal education. Coventry, 2002.

[36] B. E. Whitley, "FACTORS ASSOCIATED WITH CHEATING AMONG COLLEGE STUDENTS: A Review," Res. High. Educ., vol. 39, no. 3, pp. 235–274, 1998, doi: 10.1023/A:1018724900565.

Web attack detection using ML algorithms and CIC-IDS 2017 dataset

Oumaima Chakir[1,2] and Yassine Sadqi[1,3]

[1] Laboratory LIMATI, FPBM, USMS University, Beni Mellal, Morocco
[2] chakir.oumaima@hotmail.com
[3] yassine.sadqi@gmail.com

Abstract. Because of the increased creation and use of web apps in recent years, there has been a surge in web-based attacks. A web application security-critical component has been identified as an intrusion detection and prevention system(IDPS) capable of identifying and preventing these attacks. In this respect, Machine Learning-based intrusion detection and prevention systems have been important in recent literature. As part of the evaluation of the effectiveness of machine learning algorithms in web-based attack detection, we continued in this paper the evaluation using one of the most recent datasets, namely CIC-IDS 2017. In which, we have investigated four types of machine learning algorithms commonly used in cybersecurity: KNN, Decision Tree, Multinomial, and Bernoulli Naive Bayes, SVM Linear, Sigmoid, and RBF. The experimental results based on the CIC-IDS 2017 dataset demonstrated that the decision tree (DT) surpassed all other classifiers in terms of accuracy, recall, precision, F1, FPR, and FNR with values of 99.75%, 96.86%, 85.86%, 91.03%, 0.2%, and 3.14%.

Keywords: Web application security · Web-based attacks · Machine Learning (ML) · CIC-IDS 2017 dataset · IDPS

1 Introduction

Because of the constant expansion in the number of threats targeting web apps, securing them has become one of the most challenging tasks in the cybersecurity area [1,2].To avoid such attacks, employ a security system capable of analyzing, detecting, and preventing any malicious HTTP/HTTPS Requests/Responses.

Web application firewalls, or WAFs, are a well-known and widely used intrusion detection and prevention systems for web application security. WAF uses two different types of detection approaches: signature-based and anomaly-based. On the one side, because of their unknown signatures, signature-based WAFs cannot identify new attacks, such as zero-day attacks. On the other side, while anomaly-based WAFs are efficient for detecting known and unknown attacks, they harm web app availability due to a high false-positive rate.

Machine learning and deep learning-based web intrusion detection systems have revealed an autonomous capacity to identify normal from abnormal user

behavior [3]. Experiments have shown that these systems can detect both known and novel web attacks with a very low false-positive rate.

As part of the evaluation of the effectiveness of machine learning algorithms in web-based attacks detection, we continue in this paper the evaluation using one of the most recent datasets, namely CIC-IDS 2017 [4]. In our previous work [5], we evaluated the potential effectiveness of KNN, Decision Tree (DT), SVM Linear (SVML), SVM RBF (SVMR), SVM Sigmoid (SVMS), Multinomial Naive Bayes (MNB), and Bernoulli Naive Bayes (BNB) classifiers in web application security using two open-source datasets generated for web-based attacks detection, namely ECML/PKDD 2007 [6] and CSIC HTTP 2010 [6].

This paper is structured as follows: Section 2 discusses related works. Section 3 presents the research methodology in detail. Section 4 presents and discusses the experimental outcomes. Section 5 is devoted to the conclusion.

2 Related Work

The experimental evaluation of the most prominent ML algorithms used to prevent web threats is the topic of this study. Web security techniques based on learning methods, in particular, have experienced a notable increase in attention since 2011 [7].

In [8], the authors compared the performance of three algorithms: J48, Adaboost, and Naive Bayes. J48 had a Detection Rate of 95.97%, whereas Naive-Bayes had a rate of 88.89%, and AdaBoost had a rate of 83.23%. Kozik et al. only offered accuracy, which is insufficient to demonstrate the performance of any machine learning algorithm in web intrusion detection. Other performance metrics such as Precision, Recall, F1-score, FPR, and FNR are discussed in this work.

The authors in [9] demonstrated that SVM and Logistic regression are effective in detecting web attacks using the CSIC HTTP 2010 dataset, with an accuracy of 95% and 97%, respectively.

Khan et al [10] investigated the performance of KNN, Naive Bayes, J48, and SVM classifiers, in detecting XSS-based attacks using a customized dataset of 1924 cases, which included 1515 bengin and 409 javascript codes as harmful instances. According to their experimental results, all four classifiers attained accuracy values of 97.14%, 95.06%, 99.22%, and 94.55%, respectively.

In [11],the authors examined the performance of KNN, SVM with polynomial and linear nuclei, and Random Forest classifiers in identifying XSS attacks. The experiments were conducted using a customized training data set of 2,000 harmful and benign scripts, and a test data set of 13,000 malicious and benign scripts. Based on their results, all these classifiers worked well in terms of accuracy, precision, sensitivity, and specificity.

SQLiGoT, a method for identifying injection threats by SQL queries modeling such as token graph using SVM, was introduced by Kar et al [12]. The authors developed different SVM classifier designs based on single and multiple SVMs. Their experimental results revealed that the proposed method performed

well in terms of accuracy, precision, recall, F-value, and FPR using single, two, and three SVMs.

3 Research Methodology

In this paper, we continue the evaluation of the most prominent ML algorithms in the cybersecurity sector, such as KNN, Decision Tree (DT), SVM Linear (SVML), SVM RBF (SVMR), SVM Sigmoid (SVMS), Multinomial Naive Bayes (MNB), and Bernoulli Naive Bayes (BNB), using the current dataset, namely CIC-IDS 2017. The following table present the parametre of each classifier used in this paper.

Table 1. The parameters of each classifier used in this work.

Classifier	Parameters
KNN	n_neighbors=5
SVM Linear	C=1 kernel='linear' gamma=20 decision_function_shape='ovr'
SVM RBF	kernel='rbf' random_state=0 gamma=20 decision_function_shape='ovr'
SVM Sigmoid	C=1 kernel='sigmoid' gamma=20 decision_function_shape='ovr'
DT	criterion='entropy' random_state=0

3.1 Dataset Description

CIC-IDS 2017 is a dataset created by the Canadian Institute for Cybersecurity at the University of New Brunswick. This dataset consists of a 5-day data stream. Each data is labeled as either a benign or a malicious data, with the specific type of attack [13]. In this work, we used the web-based attacks list collected on Thursday morning 6 July. It is a list that contains three types of attacks. SQL Injection is an injection attack that occurs when an attacker interferes with queries made by an application to its database in order to allow unauthorized users to read the data [14] [15]. Cross-Site Scripting attack occurs when an attacker injects malicious code into the victim's web app [14] [16]. Brute force attack is used to decipher the administrator's password, which tries a probabilistic collection of all possible passwords [14].

Table 2. The Training and testing CIC-IDS 2017 dataset size used in this work

	Train Data	Test Data
Normal	129614	32422
Attack	1729	414

3.2 Proposed Approach

We used the CIC-IDS 2017 Thursday Working Hours Morning Web Attacks dataset, as seen in the figure 1 . We observed during the preprocessing phase that this dataset contains some "Infinity," "NaN," and duplicated data, which could harm the classifier's performance. Furthermore, we discovered that this dataset has some irrelevant features that do not participate in the final classification because they have zero values for all rows. We deleted all of these undesirable instances to ensure that the data is clean. After converting the string labels (Bengin and the name of each attack) to numerical values, we divided the data into train and test sets. Finally, we used the training and testing data to train and test each classifier.

3.3 Evaluation Metrics

Evaluation metrics are a set of metrics used to assess the performance of machine learning classification algorithms. In this work we used the most known and important metrics:
•Accuracy (A): Refers to the ratio between the number of correctly detected attacks and the total number of attacks.

$$A = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

• Recall (R): Calculate the number of correctly detected attacks.

$$R = \frac{TP}{TP + FN} \tag{2}$$

• Precision (P): This metric is used to calculate the number of correctly classified legitimate requests.

$$P = \frac{TP}{TP + FP} \tag{3}$$

• F-value (F1): The harmonic average of recall and precision.

$$F1 = 2.\frac{R.P}{R + P} \tag{4}$$

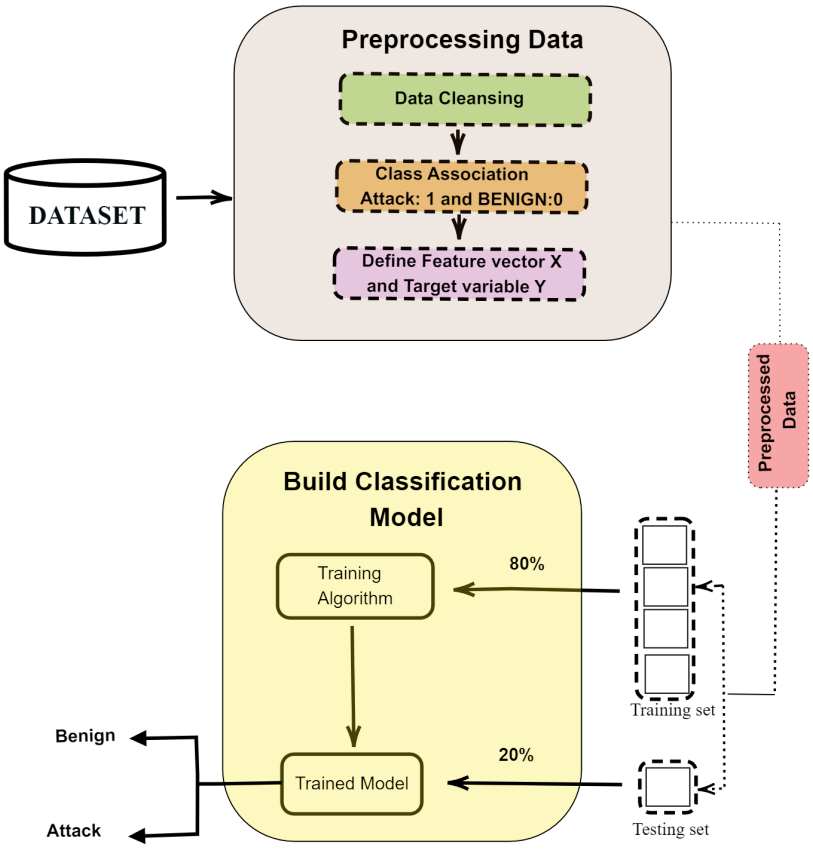


Fig. 1. Proposed Approach

• False positive rate(FPR): Calculate the number of legitimate requests detected as an attack by the classifier.

$$FPR = \frac{FP}{FP + TN} \tag{5}$$

• False negative rate(FNR): Calculate the number of attacks detected as a legitimate requests by the classifier.

$$FNR = \frac{FN}{FN + TP} \tag{6}$$

3.4 Experiment Environment

The experiments presented in this paper were carried out on a PC equipped Intel(R) Core (TM) i5-3320M*64 based CPU running at 2.6GHZ with 12G RAM installed and a 64-bit Windows 10 OS.

Table 3. Tools used in this work

Tool	Version
Conda	4.11.0
Sklearn	0.24.2
Jupyterlab	3.2.1
Python	3.9.7

4 Results and discussion

Table 4 presents the confusion matrix and Table 5 the performance evaluation of each classifier based on the CIC-IDS 2017 dataset.

Table 4. Confusion matrix of each classifier

Classifier	TN	FP	FN	TP
MNB	32390	32	414	0
BNB	29385	3037	44	370
KNN	32335	87	67	347
SVML	31178	1244	13	401
SVMR	32343	79	60	354
SVMS	32399	23	414	0
DT	32356	66	13	401

In our previous work [5], we found that SVM RBF(SVMR) performed better in terms of accuracy, recall, precision, F1, FPR, and FNR, with values of 99.97% , 100% , 99.88% , 99.93% , 0.03% , and 0% , respectively, on the ECML/PKDD 2007 dataset. Furthermore, the Decision Tree (DT) was the only classifier that performed well on the CSIC HTTP 2010 dataset, with accuracy, recall, precision, F1, FPR, and FNR values of 97.66%, 100% , 95.54% , 97.71% , 4% , and 0% , respectively.

Similarly to the results based on the CSIC HTTP 2010 dataset, the findings based on the CIC-IDS 2017 dataset revealed also that the decision tree(DT) is efficient for web-based attacks detection since it outperformed all the other classifiers in terms of accuracy, recall, precision, F1, FPR, and FNR with values of 99.75%, 96.86%, 85.86%, 91.03%, 0.2%, and 3.14%.

Table 5. The performance evaluation of each classifier based on the CIC-IDS 2017 dataset.

Classifier	A	R	P	F1	FPR	FNR
MNB	98.64%	0%	0%	0%	0.098%	100%
BNB	90.61%	89.37%	10.86%	19.36%	9.36%	10.62%
KNN	99.53%	83.81%	79.95%	81.84%	0.268%	16.18%
SVML	96.17%	96.86%	24.37%	38.95%	3.83%	3.14%
SVMR	99.57%	85.50%	81.75%	83.58%	0.24%	14.49%
SVMS	98.66%	0%	0%	0	0.07%	100%
DT	99.75%	96.86%	85.86%	91.03%	0.2%	3.14%

5 Conclusion

In this paper, we extended the study of the performance of machine learning methods in web-based attack detection, utilizing one of the most current datasets, CIC-IDS 2017. We researched and analyzed four types of machine learning algorithms commonly used in cybersecurity: KNN, Decision Tree, Multinomial, Bernoulli Naive Bayes, SVM Linear, Sigmoid, and RBF.

The experimental findings based on the CIC-IDS 2017 dataset revealed that the decision tree(DT) is efficient for web-based attacks detection since it outperformed all the other classifiers in terms of accuracy, recall, precision, F1, FPR, and FNR with values of 99.75%, 96.86%, 85.86%, 91.03%, 0.2%, and 3.14%.

References

1. (Xiaowei Li, and Yuan Xue), A Survey on Web Application Secruity,Nashville(2011).
2. (Y.Sadqi, and Y.MALEH), A systematic review and taxonomy of web applications threats. Information Security Journal: A Global Perspective (2021), p. 1-27.
3. (Maseer, Ziadoon Kamil, et al), Benchmarking of machine learning for anomaly based intrusion detection systems in the CICIDS2017 dataset, IEEE(2021)
4. Intrusion Detection Evaluation Dataset (CIC-IDS2017), https://www.unb.ca/cic/datasets/ids-2017.html
5. (OUMAIMA, Chakir, ABDESLAM, Rehaimi, YASSINE, Sadqi, et al), Experimental Study on the Effectiveness of Machine Learning Methods in Web Intrusion Detection, In : The International Conference on Information, Communication Cybersecurity. Springer, Cham(2021). p. 486-494.
6. ECML/PKDD 2007 and CSIC HTTP 2010 dataset. Available: https://gitlab.fing.edu.uy/gsi/web-application-attacksdatasets.
7. (S. M. Ghaffarian and H. R. Shahriari),Software Vulnerability Analysis and Discovery Using Machine-Learning and Data-Mining Techniques: A Survey, ACM(2017), https://doi.org/ 10.1145/3092566.
8. (R. Kozik, M. Choraś, R. Renk, and W. Hołubowicz), A Proposal of Algorithm for Web Applications Cyber Attack Detection, in Computer Information Systems and Industrial Management(2014), https://doi.org/10.1007/978-3-662-45237-0_61.
9. (R. Smitha, K. S. Hareesha, and P. P. Kundapur), A Machine Learning Approach for Web Intrusion Detection: MAMLS Perspective, in Soft Computing and Signal Processing(2019), https://doi.org/ 10.1007/978-981-13-3600-3_12.
10. (N. Khan, J. Abdullah, and A. S. Khan), Defending Malicious Script Attacks Using Machine Learning Classifiers, Wireless Communications and Mobile Computing(2017), https://doi.org/ 10.1155/2017/5360472.
11. (F. A. Mereani and J. M. Howe), Detecting Cross-Site Scripting Attacks Using Machine Learning, The International Conference on Advanced Machine Learning Technologies and Application(2018), https://doi.org/10.1007/978-3-319-74690-6_20.
12. (D. Kar, S. Panigrahi, and S. Sundararajan),SQLiGoT: Detecting SQL injection attacks using graph of tokens and SVM, Computers Security (2016),https://doi.org/10.1016/j.cose.2016.04.005
13. (R.Panigrahi,S.Borah), A detailed analysis of CIC-IDS 2017 dataset for designing Intrusion Detection Systems,International Journal of Engineering Technology(2018).
14. (A.Mohammed, F.Alenezi, and Z.Lu), Effectiveness of machine learning based intrusion detection systems, International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage. Springer(2019).
15. SQL Injection Attacks, Available: https://owasp.org/www-project-web-security-testing-guide/latest/4-Web_Application_Security_Testing/07-Input_Validation_Testing/05-Testing_for_SQL_Injection
16. Cross Site-Scripting attack, Available: https://owasp.org/www-project-top-ten/2017/A7_2017-Cross-Site_Scripting_(XSS).html

Task Offloading for the Single-User Single-Task Scenario in Mobile Edge Computing Networks

Marouane MYYARA[1], Anouar DARIF[2] and Abderrazak FARCHANE[3]
Laboratory of Innovation in Mathematics, Applications and Information Technologies (LIMATI)
Polydisciplinary Faculty, Sultan Moulay Slimane University
Beni Mellal, Morocco
[1]mar.myyara@gmail.com, [2]anouar.darif@gmail.com, [3]a.farchane@gmail.com

***Abstract*—Mobile Edge Computing (MEC) is a new model with great potential to extend the capabilities of the cloud and its services to the edge and close to mobile users, allowing compute-intensive and delay-sensitive applications to be quickly processed by edge servers to meet the needs of mobile users; this is referred to as computation offloading or task offloading. In MEC, users offload compute-intensive and delay-sensitive applications to the edge server for minimizing the transmission delay and power consumption, as it is very difficult for a mobile device with limited compute and storage resources to meet the requirements of these applications. Similarly, battery life is the main constraint for mobile users and with local computing, users may not get a better quality of experience. However, in edge computing context, the decision to offload a task is important, especially, it depends on the availability of communication resources, the resulting delays that must be respected during the offloading of the task and the energy consumption for both local and remote execution of the task to be offloaded. In this paper, we consider a single task and single user scenario where the mobile device has a single task at the moment, this task is usually heavy and computationally intensive. Therefore, the user decides whether or not to offload it to the MEC server. In fact, this decision must be carefully considered because it affects the energy consumption of the mobile device and the resulting delay.**

***Index Terms*—MEC, computation offloading, single task offloading, single user, offloading policy.**

I. Introduction

In recent years, the computing capabilities of user equipment have increased. However, they have limited computational and battery resources due to power-hungry and complex applications [1]. With the tremendous advancement of mobile devices, the considerable demands of critical resource-intensive and delay-sensitive applications such as real-time speech recognition, augmented and virtual reality, interactive games, video transformation, and image retrieval applications have attracted the attention of researchers. The delay sensitivity of these applications has led to an increase in computational demand and power consumption. Therefore, to reduce the processing time and energy consumption of mobile users, a new paradigm, known as mobile edge computing (MEC), has been introduced [2]. MEC provides computing and storage capacity to users at the edge of wireless networks, and is one of the main enablers of cloud computing, as it keeps services at the user's edge. The main objective of edge computing is to offload computations to overcome the limitations of mobile devices, such as computational capacity, storage and also to improve application performance by executing tasks that are processing intensive and resource heavy. Thus, it is a key and important component of the new 5G architecture, which supports a variety of services and applications that require very low latency, and does not accept some delay certainly large.

Edge computing is an open distributed computing architecture that provides mobile users with decentralized processing power. It is also used in the era of the Internet of Things, where part of the data collected from connected devices is processed locally at the mobile user level, to reduce network traffic to the edge server, and also to provide real-time data processing and analysis. Thus MEC can be further considered as a network architecture bringing the computing resources from the cloud to the edge of the cellular network, and instead of sending all the data to be processed to the cloud, the edge network analyzes, processes and stores the data from the mobile users, as the cloud data centers cannot guarantee acceptable response time and transfer rate. As MEC architecture, we find the mobile devices are connected to the base station (case of cellular networks) or to the cloudlet (small scale data centers or clusters) via local networks, these base stations are connected to the edge server, this one is connected to the cloud via the core network.

In this work, we study a scenario considering a single-task single-user offloading where the user executes a heavy computational task that must be processed in its associated MEC server. We organized the rest of this work as follows, first, in section 2, we review the related work and literature that have discussed the computational offloading. In section 3, we discuss the system model that constitutes our MEC system, which includes the offloading model, the network model, the task computation model and the energy consumption model. Then in section 4, we proceed to simulate the effect of the size of the computational task and the computation CPU-cycles frequency of mobile device on the energy consumption and on the execution delay, for the two offloading decisions that will be discussed next. And finally, section 5 concludes the article.

II. Related Works and Motivation

The problem of using the MEC network for computation offloading has been the subject of much research. However, the efficient offloading of computations remains a challenging

problem. In order to meet the requirements of low latency, saving mobile device costs, and improving the quality of experience of emerging applications in the MEC network, designing an efficient computational offloading strategy is important.

Recently, Mobile Edge Computing MEC (also known as Multi-access edge computing) [14] has been receiving attention more and more, and a variety of offloading policies have been proposed [3]–[5], [7], [11], [12]. These policies can be categorized into two types : centralized and distributed offloading. An additional classification can be made based on the number of users, the number of MEC servers, and the available resource sharing among MEC servers.

The key issue in edge computing is to make the offloading decision. The results of this decision are either local execution where the entire computational task is performed locally, full offloading where the entire computational task is offloaded and processed by an edge server, or partial offloading where part of the computational task is processed locally while the rest is offloaded to an edge server. By a combination of local execution and full offloading, the problem can be modeled as a binary offloading problem. Partial offloading can be modeled as a continuous offloading decision making problem.

We focus in our study only on binary offloading, as partial is not involved in the single-user single-task offloading scenario. Binary offloading is mainly concerned with small-scale computational tasks that require large computing resources, these tasks are completely transferred to the edge server. Since offloading migrates computations to a more resource-rich computing device, it involves making a decision about whether and what computations to migrate. There is a large amount of research on offloading decisions to improve performance and save energy.

To enhance the performance of mobile devices, offloading is one of the attractive solutions to meet the response time requirements of mobile devices, as applications become increasingly complex. Another objective is to address real-time constraints, where multiple data streams from different sources (GPS, maps, accelerometers, sensors, etc.) must be analyzed together to obtain real-time information about the context of the user. In many of these scenarios, the limited computing speeds of mobile systems can be improved by offloading. Smartphones are no longer only used for vocal communication, but also for acquiring and watching videos, gaming, surfing the web, and many other things. As a result, these systems will likely consume more power and reduce battery life. Even though battery technology has been constantly improving, it has not been able to keep up with the rapid growth in power consumption of these mobile devices. Offloading can extend battery life by shifting power-intensive parts of the computation to peripheral servers [16].

The table I below, classifies the different work, research and literature articles on offloading especially the binary one, that deal with offloading techniques to achieve the two objectives mentioned before that will contribute to an optimal offloading decision.

TABLE I: Summary of previous works on binary offloading

Optimization parametres	Years	Paper
Task execution delay	2015	[8]
	2016	[6], [9]
	2018	[33]
	2019	[38]
Energy consumption	2015	[10], [15], [18]
	2016	[19]
	2019	[37]
Trade-off between execution delay and energy consumption	2014	[30]
	2017	[20], [21], [31]
	2018	[13], [34], [36]
	2019	[32], [35]

III. System Model and Problem Formulation

The single-user MEC system contains a mobile device that is connected to an edge server (Figure 1). The mobile device performs a computationally intensive and delay-sensitive task, so the MEC server helps perform the task computation and provides the necessary computing power. Generally, a MEC system is composed of mobile devices and MEC servers. The mobile devices and servers are separated by the radio interface where wireless reliable connections can be established using advanced wireless communication and networking technologies.

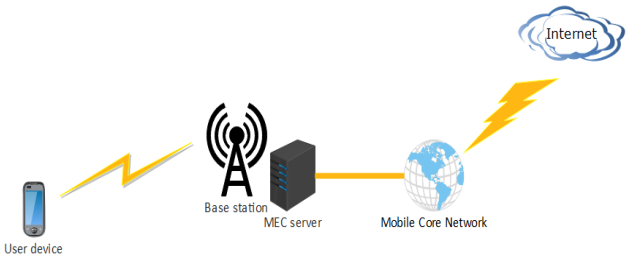


Fig. 1: Single-user single-task Mobile Egde Computing architecture.

The MEC server is deployed on the mobile access network side and communicates with the mobile device over the wireless channel. Thus we assume that the MEC server has sufficient computational resources and the computational speed is faster than the mobile device, so it can handle multiple devices simultaneously.

Based on the proposed MEC architecture, we schematize the process of computation offloading in MEC systems.

The overall offloading procedure of our system model are shown in Figure 2, which shows all the processes of offloading task into the MEC server or local executing in the mobile device. In the same figure it can be seen that when a mobile user's task arrives, the offloading decision takes into consideration the energy constraints of the mobile user's battery and also the size of the task's input data. After the decision is made, the task's input data will be transmitted either to the central computing unit of the mobile device or to the transmitter that will offload the task to the edge server.

After the calculation and processing of the task is completed, the result will be transferred to the mobile device. In short, this is the process of task offloading in a typical MEC system.

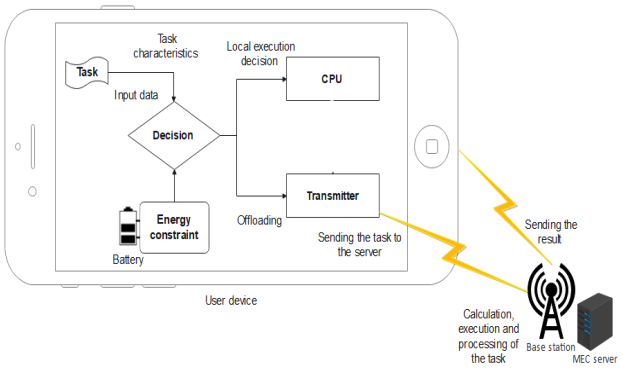


Fig. 2: Processes of task offloading in a typical MEC system

In the following subsections we will discuss the models that constitute our MEC system, we note here the offloading model, the network model, the task computation model and finally the energy consumption model of the mobile devices.

A. *Offloading Model*

Offloading resource- and computation-intensive tasks to resource-rich edge servers for processing can help to increase the capabilities of mobile devices. The reduction in energy consumption of mobile devices through computational offloading has been demonstrated in experimental tests and analytical studies [28]. The main issues in offloading computations are deciding whether to offload, how much of the computational task should be offloaded, and to which server. Basically, as already mentioned, the offloading of computations can give rise to all three types of decisions [22], local execution, full offloading and partial offloading. The decision to offload computations is very difficult, as it requires consideration of several factors, such as the application requirements, the quality of the communication link, and the capabilities of the computing resources of the edge servers.

For each task, a decision must be made whether to run it locally on the mobile device, send it to the mobile edge server, or forward it to remote cloud server. Even assuming that the processing times and costs of each task at different locations are known in advance, this leads to an integer scheduling problem that is often extremely difficult to solve [23]. In addition, the processing times and costs are likely to be unknown before a task is executed, due to the lack of exact information about the number of operations required for the task, as well as the randomness of the available computational cycles of each processor [29]. New analytical tools are needed to deal with this uncertainty in order to design MEC systems that guarantee average performance (the focus of our future work).

In our case, each mobile device U has to make decisions about remote processing or local processing of its task T_u. We denote $\alpha_{T_u} \in \{0,1\}$, the binary offloading decision variable for the task T_u of user U. Specifically, $\alpha_{T_u} = 1$ means that user U decides to execute his task T_u at the edge server and $\alpha_{T_u} = 0$ indicates a local processing decision. In other words :

$$\alpha_{T_u} = \begin{cases} 1, & T_u \text{ must be offloaded to the edge server} \\ 0, & T_u \text{ will be executed locally} \end{cases}$$

B. *Network Model*

The network model for MEC systems is generally considered to be an OFDMA system with an uplink multiple access scheme [24], where the operating frequency bandwidth is divided into equal subbands of size $w\,[Hz]$. Since the mobile device transmit on only one subband, according to [25] the achievable rate R_u of the mobile device when sending his task data to the base station is given by :

$$R_u = w.\log_2\left(1 + \frac{P_u g_i}{\sigma^2}\right) \quad (1)$$

where P_u is the user's transmit power of the transmission unit, g_i is the channel gain and σ^2 is the variance of the channel's background or complex white Gaussian noise. The product is noted as the signal interference withe noise ratio (SINR) on the associated subband at the base station.

C. *Task Computation Model*

This subsection presents the main computational components of our MEC system. In the beginning, we assumed that we have a single mobile device U that has a single computational task T_u, must be executed, which is atomic and cannot be divided into subtasks. The task T_u is characterized by the two main parameters $T_u(d_{T_u}, \lambda_{T_u})$, d_{T_u} denotes the size of offloading computation task data; λ_{T_u} denotes the number of Central Processing Unit (CPU) cycles needed for executing computation task T_u. The values of d_{T_u} and λ_{T_u} are dependent on the nature of the computational task and can be obtained by offline measurements by carefully profiling the task execution.

Each task can be executed locally on the mobile device or offloaded to a MEC server. By offloading the computational task to the MEC server, the mobile device saves its energy for task execution, but consumes extra time to send the task input over the uplink. Thus the primary computational engine of mobile devices is the CPU, the architecture of advanced mobile CPUs adopts an advanced technique of dynamic frequency and voltage scaling, which allows the CPU cycles to be increased or decreased, thereby increasing and decreasing power consumption, respectively. In practice, the value of f_c, the user's local computational capacity in terms of CPU cycles per second, is limited by a maximum value, which reflects the computational capacity limitation of the mobile. In the following all expressions of time are in seconds.

In the case of a local execution of the task T_u, the local execution time $t_{T_u}^{Local}$ is expressed as follows,

$$t_{T_u}^{Local} = (1 - \alpha_{T_u}).\frac{\lambda_{T_u}}{f_c} \quad (2)$$

this indicates that more CPU cycles are needed to reduce the execution latency.

Therefore, in the case of remote processing where the task T_u is going to be offloaded to the edge server, the total completion time of the offloading process is written,

$$t_{T_u}^{Off} = t_{T_u}^{Trans} + t_{T_u}^{Wait} \tag{3}$$

the offload time includes the transmission time $t_{T_u}^{Trans}$ and the waiting time $t_{T_u}^{Wait}$.

The transmission time $t_{T_u}^{Trans}$ to offload the task into the uplink can be calculated as,

$$t_{T_u}^{Trans} = \alpha_{T_u} . \frac{d_{T_u}}{R_u} \tag{4}$$

The waiting time $t_{T_u}^{Wait}$ includes the link delay $t_{T_u}^{Delay}$, the execution time $t_{T_u}^{Exec}$ and the reception time $t_{T_u}^{Res}$, is given by

$$t_{T_u}^{Wait} = t_{T_u}^{Delay} + t_{T_u}^{Exec} + t_{T_u}^{Res} \tag{5}$$

Since the size of the computation result from the edge server is usually much smaller than that of the input, and the downlink rate is much higher than that of the uplink, we omit the output transfer time in our computation, as is also the case in [6],[11]. Then the time to receive the output from the MEC server to the user on the downlink is ignored. Thus, the waiting time can be written as follows

$$t_{T_u}^{Wait} = \alpha_{T_u} . \left(\Phi . d_{T_u} + \frac{\lambda_{T_u}}{f_s} \right) \tag{6}$$

where Φ is the downlink transmission delay scale factor of the mobile device.

Finally the total offloading time $t_{T_u}^{Off}$ is written :

$$t_{T_u}^{Off} = \alpha_{T_u} . \left(\frac{d_{T_u}}{R_u} + \Phi . d_{T_u} + \frac{\lambda_{T_u}}{f_s} \right) \tag{7}$$

D. Energy Consumption Model

In this subsection, the energy consumption of the mobile device is discussed, whether it is to offload the computational task to the edge server or to process it locally. All energy expressions discussed below are given in Joule.

Since the mobile devices are energy constrained, the energy consumption of the local execution is a critical performance metric for computational efficiency. According to [26], the energy consumption of each CPU cycle is given by $\xi_L f_c^2$, where ξ_L is the effective switched capacitance, depending on the chip architecture. The power consumption for the local execution of task T_u with f_c CPU cycles can be calculated as follows

$$e_{T_u}^{Local} = \xi_L . f_c^2 . \lambda_{T_u} \tag{8}$$

Thus, from (Eq. 2) and (Eq. 8), if the execution time $t_{T_u}^{Local}$ is greater than a maximum tolerated latency or if the battery capacity of the mobile device is less than $e_{T_u}^{Local}$, the mobile device must decide to offload the task to the edge server for processing. Otherwise, if it executes the task locally, it may lose its energy before completing the task execution. If the task T_u is offloaded to the MEC server, then its offload energy consumption is written as:

$$e_{T_u}^{Off} = e_{T_u}^{Trans} + e_{T_u}^{Exec} \tag{9}$$

where $e_{T_u}^{Trans}$ is the transmission power consumption occurred at the user and $e_{T_u}^{Exec}$ is the power consumption occurred during processing of the task T_u in the MEC server. Furthermore, the energy consumption during transmission of T_u can be obtained by multiplying each term of offload time $t_{T_u}^{Off}$ by the transmission power allocated to the user P_u

$$e_{T_u}^{Trans} = P_u . t_{T_u}^{Trans} + P_u . t_{T_u}^{Wait} \tag{10}$$

Thus the total energy consumption in the offloading case of the task T_u is written

$$e_{T_u}^{Off} = P_u . t_{T_u}^{Trans} + P_u . t_{T_u}^{Wait} + \alpha_{T_u} . \xi_S . \lambda_{T_u} \tag{11}$$

where the product $\xi_S . \lambda_{T_u}$ is the energy consumption during the execution of the task T_u in the edge server to which the mobile device is associated [27].

Combining the execution time (Eq. 2) and the energy consumption (Eq. 8) of local execution, the total cost of local computing can be given as

$$C_{T_u}^{Local} = X_u^t . t_{T_u}^{Local} + X_u^e . e_{T_u}^{Local} \tag{12}$$

where X_u^t and X_u^e represent the weights of execution time and energy consumption cost of task T_u. We assume that $X_u^e \in [0,1]$, and $X_u^t \in [0,1]$ to avoid the case that a task experiences a huge delay. These parameters depend on the user and the type of application. For example, for a delay–sensitive application, the user may set $X_u^t = 1$ and $X_u^e = 0$.

Simularly, Combining the execution time 7 and the energy consumption 11 of offloading task, the total cost of offloading computing can be given as

$$C_{T_u}^{Off} = X_u^t . t_{T_u}^{Off} + X_u^e . e_{T_u}^{Off} \tag{13}$$

and the sum cost of the task T_u in the MEC offloading system is expressed as

$$C_{T_u}^{Total} = (1 - \alpha_{T_u}) . C_{T_u}^{Local} + \alpha_{T_u} . C_{T_u}^{Off} \tag{14}$$

IV. Results and Discussions

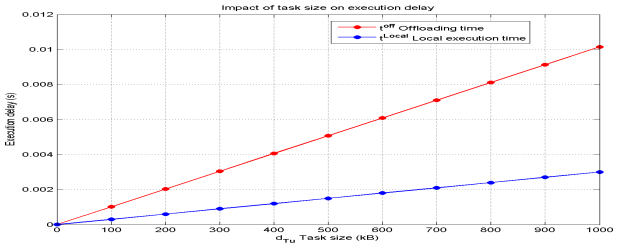
In this section, a simulation experiment is provided regarding task offloading for mobile edge computing, and we present simulation results to study the impact of the size of the computational task on the execution delay and on the energy consumption in the mobile device. For this, we opted to program the equations of the different model of our system modeled above. The programming was done using MATLAB R2013a simulator. The program was implemented on an Intel Core i7-5600U CPU 2.60 GHz laptop, with 8 GB RAM, 64-bit operating system. The simulation parameters are listed in the table II .

We first illustrate the effect of increasing the size of the task on the execution time, on the energy consumed during the execution of the task and on the sum cost of the MEC system in Fig. 3. The simulation has been performed on both

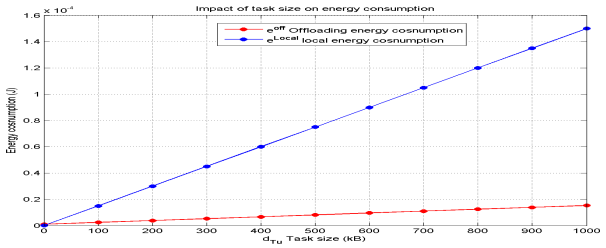
TABLE II: Simulation parameter values.

Parameter	Value
W	5 MHz
P_u	0.1 W
g_i	10^{-6}
σ^2	10^{-9} W
d_{T_u}	$100 - 1000$ kB
λ_{T_u}	300 cycles CPU
f_c	$0.1 - 2$ GHz
Φ	$0.00001 - 0.00002$ sec/kB
ξ_L, ξ_S	$10^{-11} - 10^{-28}$
f_s	$3 - 5$ GHz

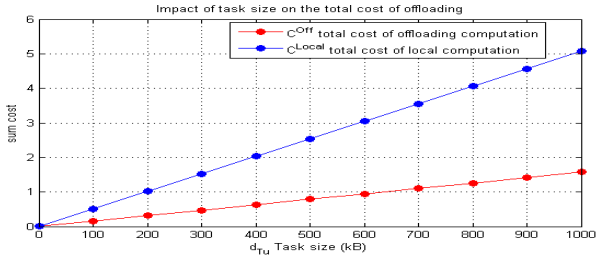
offloading decisions, either local execution or offloading to a server edge. We observe the trade-off between the energy consumption and the execution time of mobile device for different task size in same figure.



(a)



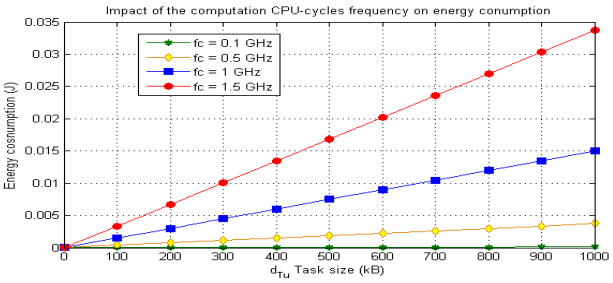
(b)



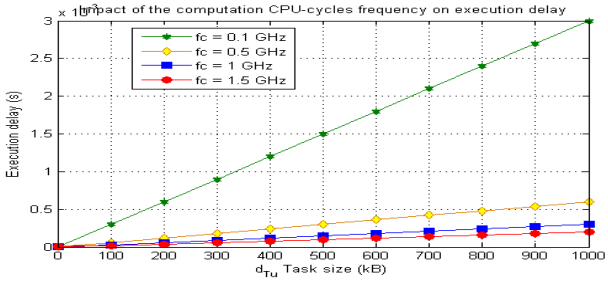
(c)

Fig. 3: Effect of task size on (a) execution delay, (b) energy consumption and sum cost (c).

Figure 3 shows that more energy is consumed when the task is executed locally. The task offloaded to the MEC sever has higher latency, resulting in the bad user experience. It can be said that this is due to the fact of local computation in the process of the mobile device itself requires more energy than that needed to offload the task to the server, and the same goes for the execution time, we can see that the time required for local execution is smaller than the time required for offloading. From the perspective of the mobile device, the overhead of the mobile device mainly depends on its transmission power and computation CPU-cycle frequency.



(a)



(b)

Fig. 4: The impact of the computation CPU-cycles frequency of mobile device on (a) the energy consumption and (b) the execution delay under different task size.

Figure 4 shows the impact of the computation CPU-cycle frequency of mobile device on the energye consumption and execution latency as a function of the task size. We can observe that the energy consumption under $f_c = 1.5$ GHz is higher than under $f_c = 0.1$ GHz and the execution delay under $f_c = 1.5$ GHz is lower than under $f_c = 0.1$ GHz with an increase in the task size, which is consistent with the definition of local energy consumption and latency. We can deduce that a range of CPU frequency is desirable to have a good compromise between energy consumption and latency that corresponds to the frequency range between 0.5 and 1 GHz.

V. Conclusion

In this paper, we have presented the scenario of single user single task offloading in a edge server. Firstly, we have cited recent works that have dealt with the case of task offloading for MEC systems especially the binary offloading. After that we moved to the system model which is composed by different models, offloading model, network model, task computation model and energy consumption model. Simulation results demonstrate the effect of the size of the computational task on the energy consumption and on the execution delay. To extend

this study, we can study a multi-user multi-task scenario with several edge servers to have a general vision on the MEC networks in general. Moreover, we can improve the above mentioned models, especially in the part of the offloading decision, instead of modeling it mathematically, we try to make it dynamic and adapt to all kind of constraints.

REFERENCES

[1] R.S. Pereira, D.D. Lieira, M.A. da Silva, A.H. Pimenta, J.B. da Costa, D. Rosário, L. Villas, R.I. Meneguette, "RELIABLE : Resource allocation mechanism for 5G network using mobile edge computing". Sensors 2020, 20, 5449.

[2] Q.V. Pham, F. Fang, V.N. Ha, M.Le, L.B. Le, W.J. Hwang, Z. Ding, "A survey of multi-access edge computing in 5G and beyond: Fundamentals, technology integration, and state-of-the-art". IEEE Access, 2020, 8, 116974–117017.

[3] X. Zhang, H. Huang, H. Yin, D. O. Wu, G. Min, Z. Ma, "Resource provisioning in the edge for iot applications with multi-level services," IEEE Internet of Things Journal, 2018.

[4] Z. Yu, J. Hu, G. Min, H. Lu, Z. Zhao, H. Wang, N. Georgalas, "Federated learning based proactive content caching in edge computing," IEEE Global Communications Conference. IEEE, 2018, pp. 1–6.

[5] E.N. Pencheva, I.I. Atanasov, P.K. Penchev, V.G. Trifonov, "Web service interfaces for intra-cell terminal activity," 13th International Conference on Advanced Technologies, Systems and Services in Telecommunications (TELSIKS). IEEE, 2017, pp. 124–127.

[6] J. Liu, Y. Mao, J. Zhang, K.B. Letaief, "Delay-optimal computation task scheduling for mobile-edge computing systems", IEEE Int. Symp. Inf. Theory (ISIT), Jul. 2016, pp. 1451–1455.

[7] P. Mach, Z. Becvar, "Mobile edge computing: A survey on architecture and computation offloading", IEEE Communications Surveys & Tutorials, vol. 19, no. 3, pp. 1628–1656, 2017.

[8] J. O. Fajardo, I. Taboada, F. Liberal, "Radio-aware service-level scheduling to minimize downlink traffic delay through mobile edge computing", International Conference on Mobile Networks and Management, Santander, 16–18 Sep 2015 (Springer, 2015), pp. 121–134.

[9] Y. Mao, J. Zhang, K.B. Letaief, "Dynamic computation offloading for mobile-edge computing with energy harvesting devices" IEEE J. Sel. Areas Commun. 34(12), 3590–3605, 2016.

[10] M. Kamoun, W. Labidi, M. Sarkiss, "Joint resource allocation and offloading strategies in cloud enabled cellular networks", IEEE International Conference on Communications, London, 8–12 June 2015 (IEEE, 2015), pp. 5529–5534.

[11] M. T. Beck, M. Werner, S. Feld, and S. Schimper, "Mobile edge computing: A taxonomy", Proceedings of the Sixth International Conference on Advances in Future Internet. Citeseer, 2014, pp. 48–55.

[12] H. El-Sayed, S. Sankar, M. Prasad, D. Puthal, A. Gupta, M. Mohanty, and C.-T. Lin, "Edge of things: The big picture on the integration of edge, iot and the cloud in a distributed computing environment", IEEE Access, vol. 6, pp. 1706–1717, 2018.

[13] J. Zhang, X. Hu, Z. Ning, E. C.H. Ngai, L. Zhou, J. Wei, J. Cheng, and B. Hu, "Energy-latency tradeoff for energy-aware offloading in mobile edge computing networks", IEEE Internet Things J., vol. 5, no. 4, pp. 2633–2645, Aug. 2018.

[14] J.H. Anajemba, T. Yue, C. Iwendi, M. Alenezi, M. Mittal, "Optimal cooperative offloading scheme for energy efficient multi-access edge computation", IEEE Access, vol. 8, pp. 53931–53941, 2020.

[15] W. Labidi, M. Sarkiss, M. Kamoun, "Energy-optimal resource scheduling and computation offloading in small cell networks", IEEE 22nd International Conference on Telecommunications, Sydney, 27–29 Apr 2015 (IEEE, 2015), pp. 313–318.

[16] Kumar K, Lu YH, "Cloud computing for mobile users: can offloading computation save energy", IEEE Comput 43, pp. 51–56 2010.

[17] E. Cuervo, A. Balasubramanian, D. Cho, A. Wolman, S. Saroiu, R. Chandra, P. Bahl, "MAUI: making smartphones last longer with code offload", International conference on mobile systems, applications, and services, pp 49–62, 2010.

[18] W. Labidi, M. Sarkiss, M. Kamoun, "Joint multi-user resource scheduling and computation offloading in small cell networks", IEEE International Conference on Wireless and Mobile Computing, Network, and Communications, Abu Dhabi, 19–21 Oct 2015, pp. 794–801.

[19] K. Zhang, Y. Mao, S. Leng, Q. Zhao, L. Li, X. Peng, L. Pan, S. Maharjan, Y. Zhang, "Energy-efficient offloading for mobile edge computing in 5G heterogeneous networks", IEEE Access 4, 5896–5907, 2016.

[20] M. Chen, B. Liang, M. Dong, "Joint offloading and resource allocation for computation and communication in mobile cloud with computing access point", IEEE International Conference on Computer Communications, Atlanta, Georgia, 1–4 May 2017 (IEEE, 2017), pp. 1–6.

[21] T.Q. Dinh, J. Tang, Q.D. La, T.Q.S. Quek, "Offloading in mobile edge computing: task allocation and computational frequency scaling", IEEE Trans. Commun. 65(8), 3571–3584 2017.

[22] P. Mach, Z. Becvar, "Mobile edge computing : a survey on architecture and computation offloading", in IEEE Communications Surveys & Tutorials, vol. 19, no. 3, 2017, pp. 1628–1656.

[23] M. Drozdowski, "Scheduling for Parallel Processing", Springer, 2009.

[24] E. Dahlman, S. Parkvall, J. Skold, "4G: LTE/LTE-advanced for mobile broadband", Academic press, 2013.

[25] P. Zhao, H. Tian, C. Qin, G. Nie, "Energy-saving offloading by jointly allocating radio and computational resources for mobile edge computing", IEEE Access, vol. 5, pp. 11 255–11 268, 2017.

[26] X. Chen, L. Jiao, W. Li, X. Fu, "Efficient multi-user computation offloading for mobile-edge cloud computing", IEEE/ACM Transactions on Networking, vol. 24, no. 5, pp. 2795–2808, 2015.

[27] K. Zhang, Y. Mao, S. Leng, Q. Zhao, L. Li, X. Peng, L. Pan, S. Maharjan, Y. Zhang, "Energy-efficient offloading for mobile edge computing in 5g heterogeneous networks", IEEE access, vol. 4, pp. 5896–5907, 2016.

[28] K. Kumar Y.H. Lu, "Cloud Computing for Mobile Users: Can Offloading Computation Save Energy ?" IEEE Computer, vol. 43, no. 4, pp. 51-56, Apr. 2010.

[29] D.B. Shmoys, J. Wein, D.P. Williamson, "Scheduling Parallel Machines On-Line, SIAM Journal on Computing", vol. 24, no. 6, pp. 1313-1331, Dec. 1995.

[30] S. Barbarossa, S. Sardellitti, P.D. Lorenzo, "Communicating while computing: Distributed mobile cloud computing over 5G heterogeneous networks", IEEE Signal Process. Mag., vol. 31, no. 6, pp. 45–55, Nov. 2014.

[31] Y. Mao, J. Zhang, K.B. Letaief, "Joint task offloading scheduling and transmit power allocation for mobile-edge computing systems", in Proc. IEEE Wireless Commun. Netw. Conf, Mar. 2017, pp. 1–6.

[32] F. Shan, J. Luo, J. Jin, W. Wu, "Offloading delay constrained transparent computing tasks with energy-efficient transmission power scheduling in wireless IoT environment", IEEE Internet Things J., vol. 6, no. 3, pp. 4411–4422, Jun. 2019.

[33] A. Yousefpour, G. Ishigaki, R. Gour, J.P. Jue, "On reducing IoTservice delay via fog offloading", IEEE Internet Things J., vol. 5, no. 2, pp. 998–1010, Apr. 2018.

[34] H. Guo, J. Liu, J. Zhang, W. Sun, N. Kato, "Mobile-edge computation offloading for ultradense IoT networks", IEEE Internet Things J., vol. 5, no. 6, pp. 4977–4988, Dec. 2018.

[35] T. Yang, H. Feng, C. Yang, Y. Wang, J. Dong, M. Xia, "Multivessel computation offloading in maritime mobile edge computing network", IEEE Internet Things J., vol. 6, no. 3, pp. 4063–4073, Jun. 2019.

[36] L. Liu, Z. Chang, X. Guo, "Socially aware dynamic computation offloading scheme for fog computing system with energy harvesting devices", IEEE Internet Things J., vol. 5, no. 3, pp. 1869–1879, Jun. 2018.

[37] G. Zhang, F. Shen, Z. Liu, Y. Yang, K. Wang, M.T. Zhou, "FEMTO : Fair and energy-minimized task offloading for fog-enabled IoT networks", IEEE Internet Things J., vol. 6, no. 3, pp. 4388–4400, Jun. 2019.

[38] Y. Wu, J. Shi, K. Ni, L. Qian, W. Zhu, Z. Shi, L. Meng, "Secrecybased delay-aware computation offloading via mobile edge computing for Internet of Things", IEEE Internet Things J., vol. 6, no. 3, pp. 4201–4213, Jun. 2019.

Social engineering: common attacks and countermeasures

Mohamed Zaoui and Yassine Sadqi

Laboratory LIMATI, FPBM, USMS University, Beni Mellal, Morocco

Abstract

Traditionally, cybersecurity is viewed through the window of a technical issue to be handled in the technological space, but in the fact, it is more than a technical issue. Hackers recognize that one of the easiest routes to break into an organization, stealing information, or hacking a system is by misleading the user into taking actions that are not in their best interest, such as clicking on a malicious link or opening an infected file. Social engineering is one of the most prolific and effective means of gaining access to secure systems and obtaining sensitive information, however, it requires minimal technical knowledge. This paper provides an overview of social engineering and explores the common attacks used by social engineers. In closing, we recommend various preventive measures and possible tips to encounter threats of social engineering.

Keywords: social engineering; social engineering attacks; cybersecurity; phishing; Social Engineering Measures.

1. Introduction

After the emergence of the COVID-19 pandemic, remote working has become common in many organizations and firms. Employees began working from home and conducting all external interactions via remote communication solutions such as instant messaging, video calling, and video conferencing. This resultant shift of remote work, as well as education online, can be seen as a catalyst for an increase in cyber-attacks, especially in the field of social engineering [16]. Social engineering can be defined as the art of influencing people to divulge sensitive information and the process of doing so is known as a social engineering attack [17]. Social engineering is extremely successful because it relies on human emotions such as fear, greed, curiosity, rushing, trust, and so on. Based on a relative comparison between technical problems and human errors, it is reasonable to characterize humans as the weakest link in the computer security chain [1]. Hackers recognize that the easiest way to gain access to a system or reach sensitive data is by manipulating people into performing actions or disclosing confidential information. Cybercriminals choose social engineering techniques when there is no way to hack a system with no technical vulnerabilities [4]. furthermore, people generally believe that they are adept at detecting such attacks. Research, however, shows that people perform poorly in detecting deceptions and lies [14]. Whereas technical solutions alone cannot provide complete protection, the awareness of necessary security requirements plays an additional role in encountering some of the most common and effective forms of

social engineering attacks. Therefore, the impacts of the lack of awareness and training against social engineering techniques of deceiving and misleading ordinary users may lead to financial losses, reputational damage, and large legal fees for organizations and firms [15].

The main contribution of this article is to provide a comprehensive overview of social engineering in the first section.

The remainder of this paper is structured as followed: Section 2 highlights the most common attacks and techniques used by attackers. Section 3 provides a classification of Social Engineering Attacks. Section 4 suggests several recommendations and countermeasures that may reduce the probability of success of these attacks, followed by a Conclusion in Section 5.

2. Overview of Social Engineering

Social Engineering is a term used to describe any cyberattack that targets a person rather than a machine; for this reason, it is frequently referred to as « the art of people hacking ». It manipulates users' minds to deceive them into making security mistakes or divulging important information.

Although social engineering attacks differ from each other, they have a common pattern with similar phases. The common pattern involves four phases: (1) collect information about the target; (2) develop a relationship with the target; (3) exploit the available information and execute the attack; and (4) exit with no traces [12]. Figure 1

Figure 1. Social engineering attack stages

illustrates the different stages of a social engineering attack. [4]

A. Investigation

In this initial phase, also called information gathering or the research phase, the attacker collects information about the victim. The attacker may start by obtaining a small amount of publicly available information from a victim's social media presence, which he could then use to get more information from, say, phone number, account information, email, friends, and so on. The likelihood of success for most attacks depends on this phase, so it is only natural that attackers invest the majority of their time and attention here. [2][3]

B. HOOK

In the hook phase, the attacker starts to gain the trust of the potential victim through direct contact, over the phone, or email communication. It can also be as building an online relationship with the target through a fake profile on a dating or social networking site. [3][4]

C. Play

The play phase aims to accomplish the purpose of the attack. In this phase, the

attacker exploits the available information and executes the attack by influencing the victim emotionally to provide sensitive information or perform security mistakes by misleading him /her into taking actions that are not in their best interest, such as clicking on a malicious link or opening an infected file. [4][5]

D. Exit

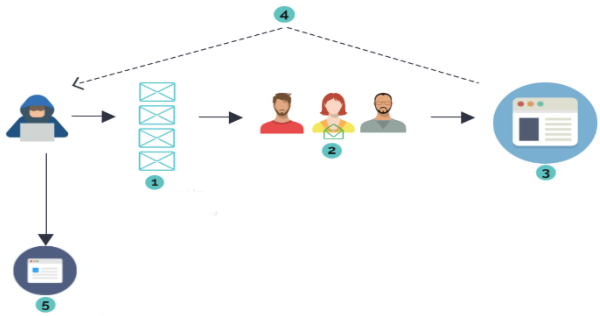
Lastly, after achieving his ultimate goal, the social engineer finalizes the interaction with the victim, then quits without leaving any proof or traces. After this last phase, the attacker is typically very difficult to track down. [5]

3. Common Social Engineering Attacks

Social engineering attacks have no limit, and they only depend on the creativity of social engineers. This section provides a brief description of the different kinds of common social engineering attacks.

3.1 Phishing

Phishing is the most common attack conducted by social engineers. Whereas social engineering is a very general term used to describe any attack that takes advantage of a human rather than a computer system, phishing specifically describes attacks whereby an attacker manipulates people to elicit their personal data. Phishing is the combination of social engineering and technical methods to mislead targets to obtain sensitive and confidential information. Phishing is typically carried out by email spoofing or instant messaging. It targets the user who has no knowledge about social engineering attacks, and internet security, like



persons who do not take care of the privacy of their account's details such as Facebook, Gmail, credit banks accounts, and other financial accounts. [4][6]

This process of phishing attack is shown in the following diagram (figure 2):

1. The attacker sends out a malicious phishing email campaign.
2. Prospective victims receive the emails, some of them open the email and click the link.

Figure 2. The process of phishing attack.

3. The victims enter their credentials into the attacker's fake web page.
4. The web page stores the credentials or sends them directly to the attacker.
5. The attacker uses the credentials to access the site, thus taking over the victims' accounts.

There are five primary types of phishing attacks:

General Phishing: This is the traditional method of phishing, which encourages the email recipient to respond or follow up via other means. The email might include malicious links or attachments.

Spear phishing: This is a targeted version of the phishing attack, in this instead of targeting too many people the focus is on a specific person who possesses critical information.

Smishing: Also known as SMS phishing, this attack involves sending text messages or SMS asking users to click on a malicious link or provide personal information.

Vishing: Also known as voice phishing. The attacker uses a fake caller ID and pretends to be a reputable organization, such as your credit card company or bank, to obtain information.

Whaling: Whaling is identical to spear phishing, except for the size of the fish. The attacker employs spear-phishing tactics but goes after high-profile targets such as an executive within an organization or company.

3.2 Baiting

Baiting is similar to a phishing attack, but in baiting the attacker promises something in exchange. For example, the hacker promises free music, phones, prize money, etc. The attacker can use this physical attack vector by infecting a storage medium with malware, leaving it at the workplace to be found by the targeted victim, who may naively plug it into the system. The victim, out of curiosity, inserts the storage device and activates the attack unintentionally [9,10,11].

3.3 Quid pro quo

The Latin phrase means "Something for something". A quid pro quo attack is a type of baiting method. A common threat to an attacker is to impersonate a technology expert and offer assistance to a victim who may be facing technical challenges that conveniently require sensitive information to be solved [13,5].

3.4 Watering hole

The name watering hole attack is inspired by a real-life situation in which the preditor lurk near the water holes so that they can attack a targeted prey. Similarly, in this attack, the attacker compromises a website that is likely to be of interest to the chosen victim [10,12].

3.5 Dumpster Diving

consist of gathering sensitive documents from the company's trash or discarded equipment such as old computer materials, drives, CDs, and DVDs. Often, individuals and organizations, do not adequately dispose of documents, papers, and even hardware from which can be retrieved confidential data [4,5].

3.6 Pretexting

This type of social engineering attack focuses on creating a fabricated scenario. The attacker impersonates an authority figure or a trustworthy entity to steal a victim's personal information. This attack is mostly done through mobile number spoofing by making up and using a fake scenario (the pretext) to force the target to look for an urgent solution. The strength of this type of attack depends on the awareness and intelligence of the target [4,5,12].

3.7 Shoulder surfing

Refers to using direct observation techniques to collect personal information by looking over someone's shoulder at their screen or keyboard, typically used for extracting authentication data [5,10,11].

3.8 Tailgating

In this type of social engineering attack, an attacker simply follows closely behind a person who has legitimate access to a specific restricted area. The attacker may ask the victim to hold the door or can simply reach for it and enter before it closes. [5,9].

3.9 Reverse Social Engineering

Describes an attack that usually involves the establishment of trust between the attacker and the victim. The attacker creates a situation in which the victim requires help and then presents himself as someone the victim will consider someone who can both solve their problem and is allowed to receive privileged information.

Figure 3. Basic classification of social engineering

3.10 Popup Window

A malicious pop-up refers to windows appearing on the victim's screen telling him that he has lost his network connection and needs to reenter their username and password. The user reacts by re-entering the login information. A program previously installed, remotely forwards back the login information to the attacker.[4]

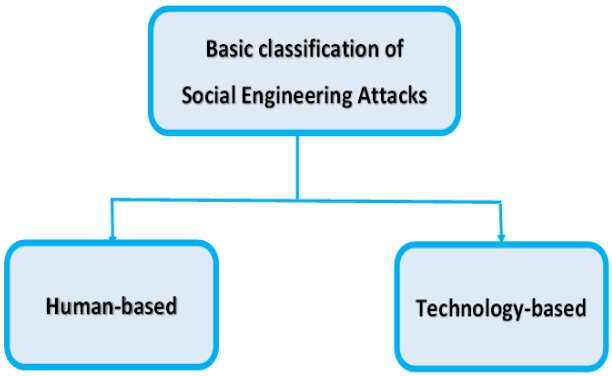
4. Classification of Social Engineering Attacks

Social engineering attacks can be classified into several categories depending on different perspectives. They can be classified into two categories according to which entity is involved: human or technology (Basic classification). They can also be classified into four categories depending on how the attack is conducted (Advanced classification): physical, technical, social, and socio-technical attacks as illustrated in Figures 3 and 4, respectively.

Figure 4. Advanced classification of social engineering

4.1 Basic classification

Social engineering attacks can be classified into two categories: human-based or technology-based as illustrated in Figure 3.

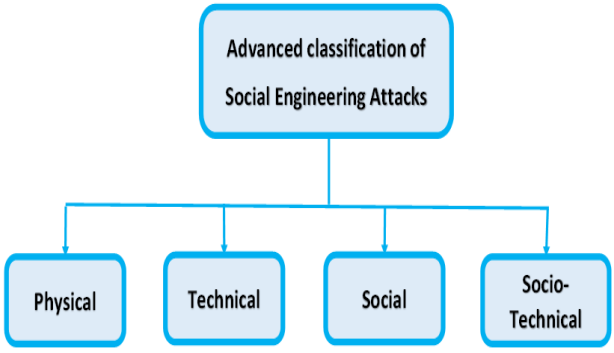


Human-based attack refers to a person-to-person interaction used to obtain the desired information. This type of attack can be carried out using third-party authorization, impersonating another employee, shoulder surfing, persuasion, or dumpster diving. The technology-based attack is performed using an electronic interface such as computers, mobile phones, software, and programs to gather information from the targets. this kind of attack attacks

can be carried out using emails, pop-ups, online scams, phishing, or vishing.[4][8]

4.2 Advanced classification

Social engineering attacks can also be classified into four distinct categories, i.e., physical, technical, social, and socio-technical attacks, as shown in Figure 4 below.



The physical type of social engineering attack involves physical actions on the part of the attacker, such as gaining access to the personal information of the target or any other credential details for a system. An example of such attacks is searching in dumpsters for valuable documents (dumpster diving attack). Technical attacks are conducted through the internet via social networks and online services websites, and they gather desired information such as passwords, credit card details, and security questions. Social-based attacks employ psychological techniques conducted through relationships with the victims to play on their psychology and emotion. These attacks are the most dangerous and successful as they involve human interactions. Examples of these attacks are baiting and spear phishing. The socio-technical type of social engineering attack is a combination of social and technical approaches [4][9].

5. Counteracting social engineering attacks

AS mentioned above, social engineers use a variety of tools and methods to gain access to users' computers and corporate networks, they are also continuously identifying new means to get sensitive information from their victims, Security awareness is a must-have skill to counter social engineering attacks. No matter your role at your company or organization, no matter whether you are an employee or not, you could be a potential target for cybercriminals. Being more security-aware will significantly help mitigate potential threats and risks. A review of the guidance available from different websites and experts led us to propose the best practices listed below, which everyone should know and follow.

- **Security Awareness Training:** Since social engineering is based on exploiting flaws in human behavior, security awareness should be an ongoing activity at any organization. Conducting, and continuously refreshing, security awareness among employees is the first line of defense against social engineering attacks.
- **Check the source:** Consider where the communication is coming from before putting your trust in it. Always check that the sender is legitimate before acting. A credible source will never email you or call asking for personal information or your login credentials.
- **Avoid making your personal information public:** As previously mentioned, the attacker starts the

- attack by collecting information about the victim. Studies consistently show that the majority of social networking site users continue to share personal information online and take little precautions to understand who is collecting and monitoring their information.
- **Use strong passwords:** Use strong passwords in logins including longer passwords, uppercase, numbers, and special symbols. Two-factor authentication must be made mandatory.
- **Don't use the same password for different accounts:** Your password should be both complex and unique, never reuse it for more than one account. If a social engineering attack obtains the password for your social media account, you don't want them to be able to unlock all of your other accounts too.
- **Secure your devices and services:** Install firewall and antivirus programs and keep them up to date. scan frequently the system to detect the possible infection, and block malicious websites and files.
- **Check and Update your Security Patches:** Cybercriminals are generally looking for weaknesses in your software, application, or systems to gain access to your credentials. As a preventive measure, regularly maintain your security patches up to date and keep your web browsers & systems up to date with the latest versions.
- **Adjust Your Spam Filter:** The majority of email service providers come up with spam filters. Any email that is deemed suspicious shall automatically be thrown away in the spam folder.
- **Never plug external media** (e.g. USBs/CDs/etc) into a computer that you care about or that is connected to any other devices.
- **Delete unknown or untrusted emails without opening them**: If you can see anything suspicious in the email, don't open it and don't click on embedded links in emails, also report it as spam to your email provider, or forward it to your IT Security department if you received the email at work.
- **Never open attachments from untrusted emails**: This includes any attachments from a legitimate contact that you were not expecting.

6. Conclusion

In this paper, we provided an overview of social engineering and the most common attacks used by social engineers. Humans are considered the weakest link in a cybersecurity chain, but they are also the first line of defense, and the way to reduce risks and mitigate potential threats in cyberspace is to make people more security-aware. The effectiveness of organizational information systems in combating social engineering threats requires a combination of advanced technical measures and managerial efforts to raise employee awareness. For these reasons, we suggested a list of security countermeasures that may help mitigate potential threats and risks of social engineering, however, it is not enough, due to the fact that social engineers are continuously

identifying new means to get sensitive information from their victims.

References :

[1] Yan, Z., Robertson, T., Yan, R., Park, S. Y., Bordoff, S., Chen, Q., & Sprissler, E. (2018). Finding the weakest links in the weakest link: How well do undergraduate students make cybersecurity judgment? Computers in Human Behavior, 84, 375–382. https://doi.org/10.1016/j.chb.2018.02.019

[2] Yasin A, Fatima R, Liu L, Wang J, Ali R, Wei Z. Understanding and deciphering of social engineering attack scenarios. Security and Privacy. 2021;4:e161. https://doi.org/10.1002/spy2.161

[3] https://www.social-engineer.org/framework/attack-vectors/attack-cycle/

[4] Salahdine, F., & Kaabouch, N. (2019). Social engineering attacks: A survey. In Future Internet (Vol. 11, Issue 4). MDPI AG. https://doi.org/10.3390/FI11040089

[5] Breda F, Barbosa H, & Morais T. (n.d.). SOCIAL ENGINEERING AND CYBER SECURITY.

[6] A Literature Survey on Social Engineering Attacks: Phishing Attack. International Conference on Computing, Communication and Automation (ICCCA2016)

[7] K. Krombholz, H. Hobel, M. Huber, and E. Weippl, "Advanced social engineering attacks," Journal of Information Security and Applications, Article vol. 22, pp. 113-122, 2015, DOI: 10.1016/j.jisa.2014.09.005.

[8] T. R. Peltier, "Social engineering: Concepts and solutions," Information Security Journal, vol. 15, no. 5, p. 13, 2006.

[9] Aldawood, H., & Skinner, G. (2020). An Advanced Taxonomy for Social Engineering Attacks. International Journal of Computer Applications, 177(30), 1–11. https://doi.org/10.5120/ijca2020919744

[10] Krombholz, K., Hobel, H., Huber, M., & Weippl, E. (2013). Social engineering attacks on the knowledge worker. SIN 2013 - Proceedings of the 6th International Conference on Security of Information and Networks, 28–35. https://doi.org/10.1145/2523514.2523596

[11] Krombholz, K., Hobel, H., Huber, M., & Weippl, E. (2015). Advanced social engineering attacks. Journal of Information Security and Applications, 22, 113–122. https://doi.org/10.1016/j.jisa.2014.09.005

[12] Shivam Lohani. Social Engineering: Hacking into Humans.

[13] Conteh, N. Y., & Schmick, P. J. (2016). Cybersecurity: risks, vulnerabilities, and countermeasures to prevent social engineering attacks. International Journal of Advanced Computer Research, 6(23), 31–38. https://doi.org/10.19101/ijacr.2016.623006

[14] T. Qin and J. K. Burgoon, "An Investigation of Heuristics of Human Judgment in Detecting Deception and Potential Implications in Countering Social Engineering," 2007 IEEE Intelligence and Security Informatics, 2007, pp. 152-159, DOI: 10.1109/ISI.2007.379548.

[15] Aldawood, H., & Skinner, G. (2019). Reviewing cyber security social engineering training and awareness programs-pitfalls and ongoing issues. In Future Internet (Vol. 11, Issue 3). MDPI AG. https://doi.org/10.3390/fi11030073

[16] Venkatesha, S., Reddy, K.R. & Chandavarkar, B.R. Social Engineering Attacks During the COVID-19 Pandemic. SN COMPUT. SCI. 2, 78 (2021). https://doi.org/10.1007/s42979-020-00443-1

[17] Mouton, Francois & Leenen, Louise & Venter, H.s. (2016). Social engineering attack examples, templates, and scenarios. Computers & Security. 59. 186 - 209. 10.1016/j.cose.2016.03.004.

Mask R-CNN for Weeds Detection:

training in different technical environments

1st Brahim JABIR
LIMATI Laboratory
Polydisciplinary Faculty, Sultan Moulay Slimane University.
Beni Mellal, Morocco
ibra.jabir@gmail.com

2nd Khalid EL MOUTAOUAKIL
LIMATI Laboratory
Polydisciplinary Faculty, Sultan Moulay Slimane University.
Beni Mellal, Morocco
elmoutaouakil.kh@gmail.com

3rd Noureddine FALIH
LIMATI Laboratory
Polydisciplinary Faculty, Sultan Moulay Slimane University.
Beni Mellal, Morocco
nourfald@yahoo.fr

***Abstract*—Automatic weed location using deep learning algorithms has been studied in recent years. These deep learning-based weed classification systems require detailed information about farmland, fields, and the characteristics of individual weeds and crops. With the development in the field of computer vision, it is possible to do wide processing and regular observation with existing advanced algorithms. A number of factors, the most important of which are the technical factors of the training environment, affects the results of training these algorithms. In this chapter, we have proposed a pipeline to detect and mask the presence of weeds in an image. The image data is preprocessed and fed into a Mask R-CNN model. In the training stage, we will see two approaches. The first is to train the model using a local machine and resources. In the second approach, we will rely on the cloud to train our model using remote resources. At the end, we will compare the returns of each approach and look at how each approach performs in terms of time and memory allocation.**

Keywords—Deep learning; CNN; Mask R-CNN; Precision Agriculture; Weed detection

I. INTRODUCTION

Weeds are clearly ranked as the major pest problem in ancient and modern agriculture [1]. Weeds are plants that grow spontaneously on agricultural soils where they are ubiquitous and undesirable among all crops in the agricultural soil. The growth of these plants, if unmanaged, has a great capacity to reduce productivity, to have a huge impact on biodiversity values, to cause damage, to make it difficult to operate harvesting machinery and to increase the impurity and moisture of the grains [2]. The negative effects of weeds on crops include competition for water, light, nutrients and space, increased production costs, difficulty in harvesting, impairment of product quality, increased risk of pests and diseases and reduced commercial value of cultivated areas [3].

Applying herbicides works best if the treatment is targeted to the specific category of weed. Knowledge of weed infestation is a fundamental procedure for the use of preventive measures in their control [4]. In order to achieve area-specific weed control and reduce the overall cost and negative impact of unnecessary herbicide use on human health and crops, it is important to have methods to quantify and to analyze the distribution of weed infestation quickly and economically, and a more practical method than systematic observation of crops is needed.

Site-specific weed management is one of the practices associated with precision agriculture. Site-specific weed management is one of the effective measures to reduce pesticide use, save herbicide costs, improve weed control and avoid unnecessary environmental pollution [5].

Automatic weed location using deep learning algorithms has been studied in recent years. These deep learning-based weed classification systems require training on powerful machines using GPUs or CPUs to yield good results.

Mask R-CNN [6] is a state-of-the-art object detection and instance segmentation system, which adds a mask prediction branch to the latest stage of Faster R-CNN, developed in Python by Facebook. Therefore, Mask R-CNN has three outputs, a class label, a bounding box offset, and a third branch that produces the object's mask.

The ROI clustering layer is also replaced by the ROI alignment layer, which performs better at mask prediction than a ROI clustering layer. Mask R-CNN is also divided into two parts, the backbone network which is responsible for feature extraction (the first layer detects low-level features and the next layer higher-level features), the region, the network head that performs classification, regression, and mask prediction. In this article, we have built an algorithm of MASK R-CNN to train it on the data we have on two different environments, the first on our own computer using GPU, and the second on the Paperspace Gradient cloud, in order to that we can compare the results in the two environments and also compare these results with the results obtained in the literature review.

II. METHODS AND TOOLS

In this section, we will present the different tools, techniques and environments that we used to build and train the algorithm, as well as the data that we used for this purpose. This part is important because it falls within the objectives of the article, which is to determine the extent of the impact of the methods and tools used on the results obtained.

A. Algorithm used

Mask R-CNN is an object detection algorithm; it is divided into two steps. At the first level or first stage, the framework takes an input image, scans it to find areas that likely contain an object using a region proposal network (RPN). Based on the areas proposed in the first stage, the framework predicts the

classes in the second stage, refines the bounding box and generates the segmentation masks for each object [7].

In Mask R-CNN [7], we differentiate between: (i) the convolutional back-end architecture which is composed of the following elements: back-end network, region proposal network and object classification module, and (ii) the headend which contains the bounding box regression module, and the mask segmentation module. Each input image is converted into a feature map by the backend network, which is a standard convolutional neural network (CNN) that extracts the features. The feature map is therefore used as input for the next step where the region proposal network (RPN) analyzes the entire images to detect candidate areas that contain objects. The implementation of the feature extraction step is based on the original implementation of Faster R-CNN with ResNet-101 [8]. The following figure shows the different component parts of the Mask R-CNN architecture.

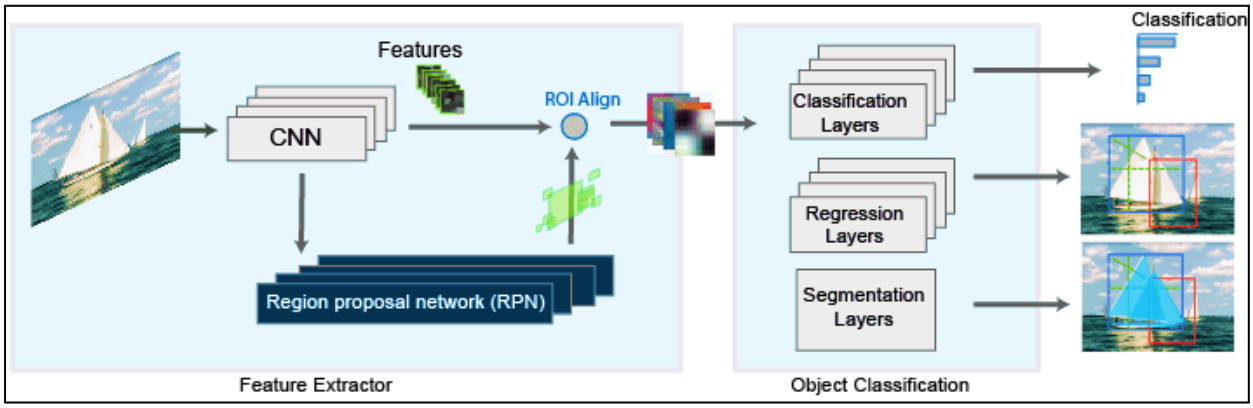


Fig. 1. Figure 4- 1 : Mask R-CNN architecture [9]

B. Dataset

The main dataset used by Mask R-CNN is an MS COCO dataset, which has 80 classes and one hundred and fifteen thousand training images [10].

The bounding box and segmentation mask evaluation metric is based on intersection versus union. The preprocessing weights learned on the MS COCO dataset are used as pre-trained weights to train our model with our own datasets. In the first step, the main dataset we use is composed of 150 weed images taken in Morocco.

After the stage of collecting images for our study. We need to make sure that we have a large variation in angles, brightness, scale, etc., and to make sure of that there are several data augmentation techniques, in our case we are increasing the amount images by adding slightly modified copies of already existing images by adjusting capture angles and brightness. This process is used mainly in the case where we only have small datasets to train our deep learning models. The goal behind feeding the model with varied data is to improve the overall training procedure and performance, as well as generalization goals [11].

After the increase, it became a set of 300 images. We divided our dataset used in this study into three sets, the first consisting of 200 images to process our model, the second consisting of 20 images for validation, and the last consisting of 80 images for testing. All images used in this study are photos of weeds found in fields and bodies in Morocco [12].

After that and to make the model bigger and give it more classes to differentiate weeds and bodies more accurately. Therefore, we are expanding our dataset by adding new weeds, and some crop types like sugar beet and wheat. The new dataset contains almost 620 images.

Our study involved some image pre-processing steps, before the particular image, features or statistics of the image are fed into the Mask R-CNN model. Our preprocessing procedure was to create pixel-level mask annotations to define the boundaries of objects in the dataset [13].

Image annotation is essential for a wide range of applications, including computer vision, robotic vision, facial recognition, and solutions that rely on machine learning to interpret images. To train these solutions, metadata must be assigned to images in the form of identifiers, captions or keywords.

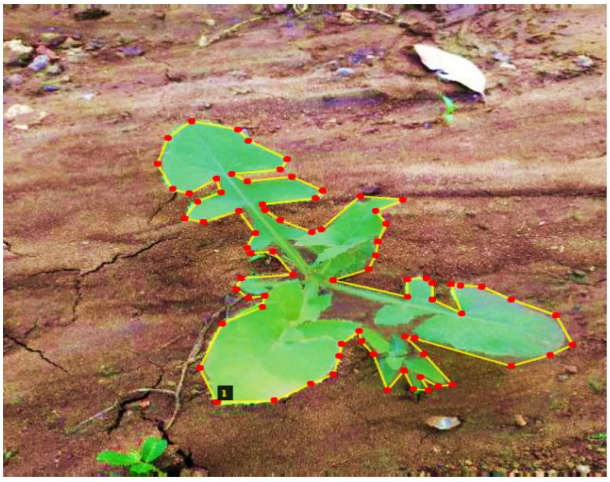


Fig. 2. Images annotation using VGG Image Annotator

C. Tools

Any training for any deep learning algorithm needs tools and libraries, the table 1 below summarizes the most important methods that we used in this study.

TABLE I. TABLE 1: NECESSARY TOOLS FOR TRAINING THE MODEL

Tools	Details
TENSORFLOW	An open source Machine Learning library, created by Google, to develop and run Machine Learning and Deep Learning applications. Its name is notably inspired by the fact that current operations on neural networks are mainly carried out via multidimensional data arrays, called Tensors. A two-dimensional tensor is the equivalent of a matrix [251]. In this study, we use TensorFlow 1.15 for local training, and TensorFlow-gpu 1.5 for cloud training [14].
KERAS	An open source library written in Python (under the MIT license) based mainly on the work of Google developer François Chollet as part of the ONEIROS (Open-ended Neuro-Electronic Intelligent Robot Operating System) project. A first version of the cross-platform software was released on March 28, 2015. The goal of this library is to enable rapid building of neural networks. In this context, Keras does not function as its own framework, but as an application programming interface (API) for accessing and programming different machine learning frameworks [262]. In this study, we use Keras 2.2.5 [15].
TENSORBOARD	Visualization tools for understanding, debugging and optimizing TensorFlow programs, also visualizes the TensorFlow graph, plots quantitative metrics on execution, and additional data [16].

III. L'APPRENTISSAGE DU MODÈLE

Training can take place at two levels. At the first level, since we have a small data set, we will use transfer learning. Transfer learning is a method of machine learning in which a model developed for one task is reused as a starting point for a model in a second task. It is a popular deep learning approach where pre-trained models are used as a starting point for computer vision and natural language processing tasks, due to the huge computational resources and time required. To develop neural network models. At this level, therefore, we can only train the heads. Here we freeze all layers of the backbone and only train the randomly initialized layers (i.e. those for which we did not use the MS COCO pretrained weights). The second level is to train all layers of the entire model.

For this study, it is not necessary to train the model completely, since we start from the weights trained by COCO. In addition, we have a double dataset, the first being composed of 300 images of which 200 images were used for training and 20 images for validation. The second being 620 frames, we do not need to train all the layers, which would take too much time.

In this study, we will train in two environments. In the first, we will use a local machine and a processor (TensorFlow), but in the second, we will use a cloud and a processor (TensorFlow-gpu).

A. CPU and GPU

The central processing unit (CPU) is the main component that provides the arithmetic logic and control of every computer. The main function of the CPU is to sequentially execute the instructions stored in the computer's memory. The CPU plays a vital role in the computation of neural networks, as it handles general arithmetic calculations during the learning phase. CPUs are typically built with a few powerful processing cores clocked between 2GHz and 3GHz, making them ideal for performing sequential tasks [17]. Furthermore, all the inputs/outputs of the training phase, such as the loading of the training data, are handled by the CPU, independently of the use of the GPU for the calculations [18]. With all these tasks, training a large model with the CPU can be risky and time-consuming, and for this reason, we believe that using the GPU instead of the CPU in the training phase will be helpful [19].

The Graphics Processing Unit (GPU), like the CPU, is the main component of a computer used to process instructions, except the GPU can execute multiple instructions at the same time using parallelization. A GPU often consists of several weak processing cores with a much lower clock speed than the CPU. The system of multiple processing cores has a major objective and the reason for its development, to parallelize the calculations by using threads to process the instructions and thus speed up the calculations usually performed over a longer period of time in the CPU [17] [18].

Since the GPU has the ability to run many processes at the same time, it has proven useful in training neural networks that involve computationally intensive matrix multiplications. Training a neural network requires a large number of computations and the GPU is used to optimize the computations by using multiple memory channels and streaming processors [20].

GPUs were originally designed for rendering graphics. Therefore, easily executing custom code on the GPU requires APIs to achieve a higher level of abstraction, from low-level programming languages to high-level languages. Since the GPU architecture has enabled all types of parallelism such as multithreading, MIMD, SIMD, and instruction level, CUDA was developed to use low-level instructions to exploit these types of parallelism. TensorFlow can, with CUDA, use the entire GPU architecture to further optimize computation time.

B. Training the model locally in the cloud

First, we will clone a repository that contains some of the building code blocks for implementing Mask RCNN. The dataset we are using found weed species (which vary in size and shape). The images in the dataset were of different sizes, so we are scaling the images to 512x512 pixels [21].

Then we choose the version of TensorFlow and Keras and this is an important step, we have encountered many errors in the 2.x version of TensorFlow. So we used TensorFlow 1.5. Also, version 2.2.5 of Keras saved us from many errors.

After that, we will import the essential libraries. Then we give the path to the trained weights file. This can be the COCO weight file or your last saved weight file (checkpoint). The log

directory is where all our data will be stored when training begins.

Next, we will create the CustomConfig class, which contains our custom configurations for the model. We are simply overwriting the original Config class data from the config.py file that was imported earlier. The number of classes is assumed to be total classes + 1 (for background). The steps per epoch are fixed at 60, but we can increase them if we have access to higher computing resources. The detection threshold is 90%, which means that all proposals with a confidence lower than 0.9 will be ignored. This threshold is different from the one used to test an image.

The next step is the creation of the *CustomDataset* class, this class inherits from "*utils.Dataset*" that we imported in the first step. This class contains three crucial methods, the "load _custom" method is for saving the annotations with the image. The "*load_mask*" method loads masks based on polygon coordinates. The mask of an image is nothing but a list containing binary values. *Skimage.draw.polygon()* does the work for us and returns the indices for the mask coordinates.

Then, creating the Train() function, we will first create an instance of the *CustomDataset* class for the training dataset. Similarly, we create another instance for the validation dataset. Then we will call the *load_custom()* method passing the name of the directory where our data is stored. The "Layers" parameter is set to "Heads" here, because we do not plan to train all the layers of the model. This will only form some of the top layers of the architecture. If you want you can set "Layers" to "all" to train all layers in the model. We run the model for only 10 epochs. Finally, we start training.

We train a model for a binary classification (it is weed or not). So we decide to expand the model and give it more classes to differentiate weeds from bodies more accurately. Therefore, we are expanding our dataset by adding new weeds and some crop types like sugar beet and wheat. The new dataset contains nearly 620 images. After preparing the new dataset, we retrain the model with this new dataset.

To train the model, we use a local machine with the following configuration: INTEL CORE i5-2520M CPU @2.50 GHz, 8GB RAM. In a second step, we use the cloud service. There are several platforms and servers that we can use to train our model like: Amazon Web Services AWS, Google Colab, Kaggle and Paperspace Gradient [22]. The last is the solution we used in this study. It is an end-to-end platform that includes a free-hosted cloud service for Jupyter laptops, with lots of preconfigured environments and free GPUs and CPUs. Gradient simplifies building, training, and deploying deep learning models. It comes with web UI, CLI and SDK. One of the advantages of this Gradient cloud is that it offers useful features for beginners as well as experts, with an impressive web user interface and an extremely low barrier to entry.

In this study, we will use the free environment. Gradient offers us two choices: the first gives us access to 2GB of RAM, 2CPU and no GPU; this environment can operate for 12 hours without interruption. The second, as the figure 3 shows, gives us access to 30GB of RAM, 8CPU, Nvidia Quadro M4000 as GPU and 6 hours of non-stop execution. We choose to work with the second environment.

Fig. 3. Our cloud training environment

IV. Result and Discussion

In this section, we will discuss the results obtained while training the model, i.e. evaluating our model using known metrics. Evaluation of a Deep Learning model is an essential step in any project process to assess and estimate the accuracy and performance of our model. There are several parameters and matrices that can be used to evaluate a model. In our study, we will use the logarithmic loss as an evaluation parameter.

Logarithmic loss or Log Loss works well for multiclass classification. It works by penalizing false classifications [23]. When working with Log Loss, the model must assign a probability to each class for all samples. Suppose we have n samples belonging to M classes, the logarithmic loss is calculated as shown below:

$$\text{Logarithmic Loss} = \frac{-1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M} y_{ij} \times \log\left(p_{ij}\right)$$

Where:

y_{ij} : Indicates whether sample **i** belongs to class **j** or not.

p_{ij} : Indicates the probability that sample **i** belongs to class **j**.

The logarithmic loss has no upper limit and is in the interval [0, ∞). A log loss closer to 0 indicates higher accuracy, while if the log loss is far from 0, it indicates lower accuracy.

In general, decreasing log loss gives better accuracy to the classifier. The goal of training a model is to find a set of weights and biases that have low loss, on average, over all examples. In the R-CNN mask, we have 5 main small losses, each one has a specific meaning, so we use the losses to calculate a big loss Loss: a combination (surely an addition) of all the small losses.

In our study, in local mode it took about 9 hours for 10 epochs with an average of 2400 seconds for each epoch (Figure 4), each epoch having 60 training batches and a threshold of 0.9 for the prediction. The training loss is 0.6350 and the validation loss is 0.9447. Training the model in our machine

requires more cost and performance to increase the accuracy and decrease the training time.

```
Epoch 8/10
60/60 [==============================] - 2352s 39s/step - loss: 0.9585 -
rpn_class_loss: 0.0154 - rpn_bbox_loss: 0.2431 - mrcnn_class_loss: 0.1323
- mrcnn_bbox_loss: 0.2517 - mrcnn_mask_loss: 0.3159 - val_loss: 0.9990 -
val_rpn_class_loss: 0.0083 - val_rpn_bbox_loss: 0.2482 - val_mrcnn_class_
loss: 0.1229 - val_mrcnn_bbox_loss: 0.2972 - val_mrcnn_mask_loss: 0.3225
Epoch 9/10
60/60 [==============================] - 2341s 39s/step - loss: 1.1262 -
rpn_class_loss: 0.0367 - rpn_bbox_loss: 0.3401 - mrcnn_class_loss: 0.1571
- mrcnn_bbox_loss: 0.2644 - mrcnn_mask_loss: 0.3278 - val_loss: 0.9970 -
val_rpn_class_loss: 0.0112 - val_rpn_bbox_loss: 0.2654 - val_mrcnn_class_
loss: 0.1310 - val_mrcnn_bbox_loss: 0.2955 - val_mrcnn_mask_loss: 0.2939
Epoch 10/10
60/60 [==============================] - 2350s 39s/step - loss: 0.8486 -
rpn_class_loss: 0.0116 - rpn_bbox_loss: 0.2418 - mrcnn_class_loss: 0.0841
- mrcnn_bbox_loss: 0.2189 - mrcnn_mask_loss: 0.2923 - val_loss: 1.0131 -
val_rpn_class_loss: 0.0086 - val_rpn_bbox_loss: 0.2669 - val_mrcnn_class_
loss: 0.1050 - val_mrcnn_bbox_loss: 0.3047 - val_mrcnn_mask_loss: 0.3280
```

Fig. 4. Le réentraînement du modèle en local

On the cloud, we are going to take the same dataset and the same model that we trained locally the second time around and we are going to run the training of this model again, but this time in the cloud. This time it took about 28 minutes for 10 epochs with an average of 170 seconds for each epoch, with each epoch having 60 training batches and a threshold of 0.9 for prediction. The learning loss is 0.8406 and the validation loss is 0.964. To decrease the loss, we increase the number of training batches to 100 and re-train the model again, this time the model takes 45 minutes for 10 epochs by the average of 280 seconds for each epoch, with each epoch having 100 training batches and a threshold of 0.9 for prediction. The learning loss is 0.6896 and the validation loss is 0.9113.

However, we see that the loss is still large, so we re-train our model this time with 40 epochs, as shown in figure 5, each epoch having 100 training lots and a threshold of 0.9 for the prediction. The learning loss is 0.2871 and the validation loss is 0.9663. This time the model takes almost 3 hours.

```
Epoch 36/40
100/100 [==============================] - 255s 3s/step -
loss: 0.2739 - rpn_class_loss: 0.0044 - rpn_bbox_loss: 0.
0452 - mrcnn_class_loss: 0.0314 - mrcnn_bbox_loss: 0.0373
- mrcnn_mask_loss: 0.1556 - val_loss: 0.9307 - val_rpn_cl
ass_loss: 0.0078 - val_rpn_bbox_loss: 0.3262 - val_mrcnn_
class_loss: 0.0848 - val_mrcnn_bbox_loss: 0.2101 - val_mr
cnn_mask_loss: 0.3018
Epoch 37/40
100/100 [==============================] - 259s 3s/step -
loss: 0.2954 - rpn_class_loss: 0.0045 - rpn_bbox_loss: 0.
0508 - mrcnn_class_loss: 0.0355 - mrcnn_bbox_loss: 0.0437
- mrcnn_mask_loss: 0.1609 - val_loss: 0.9927 - val_rpn_cl
ass_loss: 0.0072 - val_rpn_bbox_loss: 0.4082 - val_mrcnn_
class_loss: 0.0675 - val_mrcnn_bbox_loss: 0.2174 - val_mr
cnn_mask_loss: 0.2922
```

Fig. 5. Entraînement du modèle sur le serveur de gradient (cloud)

After training the model locally and on the Cloud, based on Figures 6(a) and 6(b), we are sure that the training on the Cloud is faster and better where it gives a very low loss which reaches 0.26 in the last era.

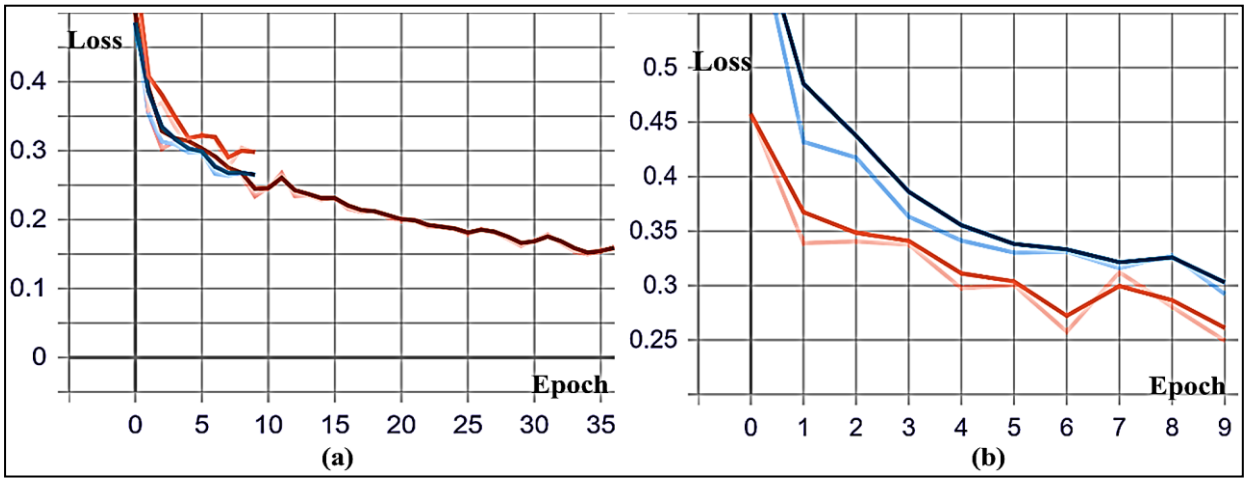


Fig. 6. (a) Model loss in local training, (b) Model loss in cloud training

We can see that 10 epochs are not enough to train the model and give good weights to generalize the model and use it in the prediction of any new image. Also, 40 epochs is excessive to train this model, because after epoch 13 we start talking about Overfitting. As we can see from the figures, after epoch 13, the loss reduction rate becomes low, which warns us that we are definitely starting to overfit.

The reason this happens is that the model always tries to minimize the loss in the sample. At some point, the model

starts memorizing the input data rather than learning general things about the actual data, which means that any new data we try to feed into the model, will perform poorly. Thus, in the inference, we will use the weights collected after epoch 13.

The table 2 below provides a summary of the comparison between training the model locally and on the cloud.

TABLE II. COMPARISON BETWEEN LOCAL MODEL TRAINING AND CLOUD MODEL TRAINING

	Epochs	Lots	Dataset	Time	Duration/ epochs		
Local	**10**	**60**	**300 images**	**7 hours**	**2100 seconds**	36 seconds	0.8486
Local	**10**	**60**	**620 images**	**9 hours**	**2300 seconds**	39 seconds	0.635
Cloud	**10**	**60**	**620 images**	**28 minutes**	**170 seconds**	2 seconds	0.8406
Cloud	**10**	**100**	**620 images**	**45 minutes**	**280 seconds**	3 seconds	0.6896
Cloud	**40**	**100**	**620 images**	**3 hours**	**280 seconds**	3 seconds	0.26

From the results discussed above, we can conclude that cloud training is more powerful than computer-based training, in terms of precision, loss, time spent, data processing, etc. This means that model training should always choose the appropriate environment and the appropriate tools, regardless of the structure of the model. Gradient offers useful features for beginners as well as experts, with an impressive web user interface and an extremely low barrier to entry. Here are some advantages of Gradient over other solutions:

- One of the advantages of Gradient is that it offers useful features for both beginners and experts, with an impressive web user interface and an extremely low barrier to entry. Here are some advantages of Gradient over other solutions:
- Faster and more persistent storage (no more reinstalling libraries and reloading files every time your laptop starts.).
- Sessions are guaranteed, so you do not risk having your instance closed in the middle of your work. Either way, you do not need to be logged in all the time, so you do not need to be interactive; start your session, log out, come back later and your session will resume where you left off.
- Preconfigured containers and templates. You can choose from a variety of different popular frameworks with all necessary dependencies pre-installed (e.g. PyTorch, TensorFlow or Data Science Stack), or use your own custom container. There is also an ML showcase that includes sample projects you can build (free) and run on your own account.
- A public dataset repository with a wide selection of popular datasets mounted on every laptop and available for free use. The ability to easily scale to add more storage and high-end dedicated GPUs for the same environment, depending on your needs.

CONCLUSION

Scientists and researchers are still trying to come up with an effective model for an intelligent system to detect weeds in crops in real time using computer vision, they are still trying every way to get a good performance deep learning algorithms. The objective of this study was to create, train and optimize a detection-qualified mask R-CNN model using two different training approaches. The first using a local machine, and the second using remote servers in the cloud and comparing the results, and using the best training weights in a web application for prediction, segmentation and identification of the presence of weeds in certain images, with the best possible precision and the least possible loss. This model can help optimize the use of herbicides and control the sprxwxwead of these weeds.

REFERENCES

[1] Young, S. L., Pierce, F. J., & Nowak, P. (2014). Introduction: Scope of the problem—rising costs and demand for environmental safety for weed control. In Automation: The future of weed control in cropping systems (pp. 1-8). Springer, Dordrecht.

[2] Voll, E., Gazziero, D. L. P., Brighenti, A. M., Adegas, F. S., Gaudêncio, C. D. A., & Voll, C. E. (2005). A dinâmica das plantas daninhas e práticas de manejo (Vol. 260). Londrina: Embrapa Soja.

[3] Rizzardi, M. A., & Fleck, N. G. (2004). Métodos de quantificação da cobertura foliar da infestação de plantas daninhas e da cultura da soja. Ciência Rural, 34, 13-18.

[4] Jabir, B., & Falih, N. (2022). Deep learning-based decision support system for weeds detection in wheat fields. International Journal of Electrical and Computer Engineering, 12(1), 816. Contreras, D. J. (2020).

[5] Critical Period of Grass Weed Control in Grain Sorghum (Sorghum bicolor) and Wheat (Triticum aestivum) (Doctoral dissertation, North Carolina State University).

[6] Mask, R. (2017). Kaiming he georgia gkioxari piotr dollr and ross girshick. In IEEE International Conference on Computer Vision (ICCV).

[7] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969).

[8] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[9] Bharati, P., & Pramanik, A. (2020). Deep learning techniques—R-CNN to mask R-CNN: a survey. In Computational Intelligence in Pattern Recognition (pp. 657-668). Springer, Singapore.

[10] Ahonen, T., Hadid, A., & Pietikainen, M. (2006). Face description with local binary patterns: Application to face recognition. IEEE transactions on pattern analysis and machine intelligence, 28(12), 2037-2041.

[11] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25, 1097-1105.

[12] Jiang, H., & Learned-Miller, E. (2017, May). Face detection with the faster R-CNN. In 2017 12th IEEE international conference on automatic face & gesture recognition (FG 2017) (pp. 650-657). IEEE.

[13] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 580-587).

[14] Dillon, J. V., Langmore, I., Tran, D., Brevdo, E., Vasudevan, S., Moore, D., ... & Saurous, R. A. (2017). Tensorflow distributions. arXiv preprint arXiv:1711.10604.

[15] Ketkar, N. (2017). Introduction to keras. In Deep learning with Python (pp. 97-111). Apress, Berkeley, CA.

[16] Ketkar, N. (2017). Introduction to keras. In Deep learning with Python (pp. 97-111). Apress, Berkeley, CA.

[17] Lind, E., & Pantigoso Velasquez, Ä. (2019). A Performance Comparison between CPU and GPU in TensorFlow.

[18] Memon, Z. A., Samad, F., Awan, Z. R., Aziz, A., & Siddiqi, S. S. (2017). Cpu-gpu processing. Int. J. Comput. Sci. Netw. Secur, 17(9), 188-193.

[19] Chien, S. W., Markidis, S., Sishtla, C. P., Santos, L., Herman, P., Narasimhamurthy, S., & Laure, E. (2018, November). Characterizing deep-learning I/O workloads in TensorFlow. In 2018 IEEE/ACM 3rd International Workshop on Parallel Data Storage & Data Intensive Scalable Computing Systems (PDSW-DISCS) (pp. 54-63). IEEE.

[20] Wei, J., Zhang, Y., Zhou, Z., Li, Z., & Al Faruque, M. A. (2020, June). Leaky DNN: Stealing deep-learning model secret with GPU context-switching side-channel. In 2020 50th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN) (pp. 125-137). IEEE.

[21] Jabir, B., Falih, N., Sarih, A., & Tannouche, A. (2021). A strategic analytics using convolutional neural networks for weed identification in sugar beet fields. AGRIS on-line Papers in Economics and Informatics, 13(1), 49-57.

[22] Marques, L., Lopes, L., Ferreira, M., Wanzeller, C., Martins, P., & Abbasi, M. (2022). Image Processing: Impact of Train and Test Sizes on Custom Image Recognition Algorithms. In Marketing and Smart Technologies (pp. 365-380). Springer, Singapore.

[23] Montaez, C. A. C., Fergus, P., Montaez, A. C., Hussain, A., Al-Jumeily, D., & Chalmers, C. (2018, July). Deep learning classification of polygenic obesity using genome wide association study SNPs. In 2018 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE.

Effect of Buoyancy Ratio on Natural, Mixed, and Forced Convection Fluid Flow and Heat and Mass Transport in a Single Lid-Driven Rectangular Cavity

Youssef Tizakast, Mourad Kaddiri
Industrial Engineering Laboratory
Sultan Moulay Slimane University
BP 523, Beni-Mellal 23000, Morocco
youssef.tizakast@usms.ma

Mohamed Lamsaadi
Research Laboratory in Physics and Sciences for Engineers (LRPSI)
Polydisciplinary Faculty, Sultan Moulay Slimane University
BP 592, Beni-Mellal, Morocco

***Abstract*—This paper reports an analytical and numerical investigation of double-diffusive mixed convection in a lid-driven rectangular cavity filled with a Newtonian fluid and subjected to uniform heat and mass fluxes along the short vertical walls, while the horizontal ones, with the top one sliding, are insulated and impermeable. The numerical solution of the governing equations adopts the finite difference method, while the analytical solution is derived based on the parallel flow approximation valid in the case of a shallow cavity. Both solutions show a good agreement for the explored ranges of dimensionless parameters governing fluid flow and heat and mass transfer, namely Peclet number, thermal Rayleigh number, Lewis number, and buoyancy ratio. In order to gain further insights into the effects of said parameters, a practically and scientifically valid process is used to delimit natural, mixed, and forced convections dominance regions. The obtained results can be expressed in the form of mixed convection parameter independently of the buoyancy ratio value, where Peclet number and thermal Rayleigh number strongly influences the transition between the convective regimes. As for buoyancy ratio, increasing it enhances natural regime input in overall convection due to intensified mass buoyancy force, leading to accelerate fluid circulation and enhance heat and mass transport phenomena while diminishing sliding walls' shear effect. In contrast, intensifying thermal buoyancy force have no significant effect on flow characteristics resulting in a dominant forced regime.**

Keywords—double-diffusive mixed convection; heat and mass transfer; lid-driven cavities; parallel flow, finite difference method; buoyancy ratio

I. INTRODUCTION

Double-diffusive convection phenomenon manifests in numerous vital industrial sectors including solar energy [1], crystal growth [2], solidification processes [3], ventilation [4], and many others. Double-diffusive convection designates fluid flows induced by buoyancy forces due to applied temperature and concentration gradients with different rates of diffusion, where in reality heat and mass transfer occur together. Double-diffusive natural convection is one of the most visited subjects in thermos-fluids. Arani et al. [5] numerically simulated the problem natural within a square cavity with active vertical walls and insulated horizontal ones. They reported found compared to pure thermal convection, thermosolutal one enhances heat and mass transfer. Furthermore, increasing Rayleigh number or buoyancy ratio also enhances heat and mass transfer rates. Liang et al. [6] numerically investigated double-diffusive convection within a rectangular cavity with applied horizontal temperature and concentration gradients. The results showed that while increasing buoyancy ratio from 1 to 2, the flow of transition is a complex series of lasting processes that continuously excite different frequency waves. Hu et al. [7] also investigated double-diffusive natural convection, this time focusing on cavity inclination and applied magnetic field effects. They concluded that heat and moisture transfer rates enhanced for both vertical and inclined cavities as thermal Rayleigh number or buoyancy ratio increased. As for the magnetic field, its capacity to suppress convection currents diminished heat and moisture transfer.

Double-diffusive mixed convection flow in lid-driven cavities is driven simultaneously by the sliding walls' shear effect and buoyancy forces generated by heat and mass gradients. For such fundamental fluid flows, Reynolds number Re and Grashof number Gr are appropriate to model as they describe shear and buoyancy forces magnitudes, respectively. Most studies published in the literature consider either flow and heat transfer (mono-diffusive) or double-diffusive mixed convection within square cavities. Nosonov and Sheremet [8] investigated mixed convection inside a rectangular cavity with the presence of a local heater. Thee results showed that increasing Richardson number associated with stronger buoyancy force enhances convection. Zhou et al. [9] numerically studied fluid flow and heat transfer within a double lid-driven cubic cavity. They found that Richardson and Reynolds numbers have a significant impact on velocity profiles and heat transfer rate. Sivasankaran et al. [10] studied double-diffusive mixed convection inside a lid-driven square cavity while focusing on sinusoidal heating effect on transport characteristics. The authors reported that independently of Richardson number, phase deviation and amplitude ratio affect heat and mass transfer rates. Kumar et al. [11] adopted the velocity-vorticity form of Navier Stokes's equations to model double-diffusive convection inside a lid-driven square cavity. They reported that for opposing temperature and concentration gradients, negative buoyancy ratios give rise to aiding flows; while for positive buoyancy ratios opposing flows generated separated flow regimes at the core of the cavity. Kefayati [12]

investigated double diffusive mixed convection within two-sided lid-driven square cavities filled with shear-thinning fluids. The study presented the effects of governing parameters: Richardson number, Lewis number, power-law index, and buoyancy ratio on heat and mass transfer. Other studies are listed for further details on mixed convection inside lid-driven cavities [13-16].

In contrast, little interest has been given to rectangular cavities despite the fact that a wide range of components can be modeled with such geometries especially in the highly complicated electronics industry. Soufiene Bettaibi et al. [17] studied double-diffusive mixed convection inside a single-lid driven rectangular cavity with applied constant temperature and concentration. The multiple relaxation time lattice Boltzmann method was adopted to solve fluid while finite difference method was used to compute temperature and concentration. The goal was to test validity of such model for predicting heat and mass transfer thermodynamics in lid-driven cavities. Tizakast et al. [18] considered double-lid driven rectangular cavities filled with non-Newtonian fluids to investigate double-diffusive mixed convection. They concluded that non-Newtonian fluids rheological behavior influence on flow characteristics is strongly associated to the dominant convection regime. Teamah and El-Maghlany [19] numerically investigated mixed convection in a rectangular cavity subjected to constant temperature and concentration with sliding top wall. The results show that heat and mass transfer rates increased as Richardson number decreased for both assisting flow (surface moves to right) and opposing flow case (surface moves to left), while increasing the absolute value of buoyancy ratio enhanced both heat and mass transfer.

All in all, the literature contains an insufficient number of published papers investigating double-diffusive mixed convection within lid-driven rectangular cavities with imposed heat and mass fluxes, particularly when accounting the high number of parameters governing fluid flow and heat and mass transfer, where rigorous investigations are required to highlight the effect of each parameter alone and the correlated effects of different parameters on flow characteristics. Accordingly, the present paper investigates, numerically and analytically, double-diffusive mixed convection within a single lid-driven rectangular cavity filled with a Newtonian fluid and submitted to uniform heat and mass fluxes from the short vertical walls, while the horizontal ones are insulated and impermeable. The combined effects of governing parameters on flow structure, heat and mass transfer, and bounds delimiting natural, mixed, and forced regimes are amply discussed while attributing special attention to buoyancy ratio effect.

II. PHYSICAL CONFIGURATION AND MATHEMATICAL MODEL

The physical configuration considered in the present study along with the associated boundary conditions investigated are given in Fig.1. A horizontal rectangular enclosure of height H' and length L' filled with Newtonian fluid and subjected to uniform horizontal density of heat and mass fluxes q' and j', respectively. The horizontal boundaries are maintained insulated and impermeable, with the top wall sliding from left to right (i.e., the same direction as the imposed heat and mass fluxes) with constant velocity u_0', while the bottom wall is motionless.

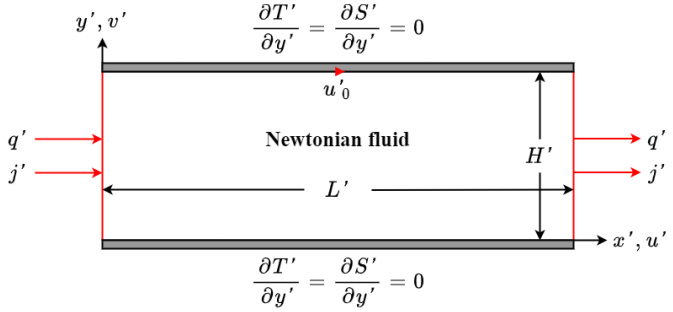


Fig. 1. Geometry of the enclosure along with the associated boundary conditions and coordinates system.

This work adopts the following commonly used main assumptions, i.e.,

- Flow is laminar. The fluid velocities are small enough due to small fluxes applied to the cavity.
- The fluid is incompressible. For pressures close to atmospheric, liquids can be considered as incompressible with a good approximation.
- Viscous dissipation contribution is negligible compared to applied heat and mass fluxes effect.
- The physical properties are considered independent of temperature and concentration except for the viscosity and density in the buoyancy term. The density obeys the Boussinesq approximation.
- The third dimension of the cavity is large enough to consider the problem as two-dimensional. This assumption is generally well satisfied and provides insights into the more complicated three-dimensional flows.

Based on the above-mentioned assumptions, dimensionless equations describing conservation of mass (1), momentum (2) and (3), energy (4), and concentration (5) written in terms of velocity components (u, v), pressure (p), temperature (T), and concentration (S) are given as follows:

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \tag{1}$$

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} = -\frac{\partial p}{\partial x} + Pr\left[\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right] \tag{2}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} = -\frac{\partial p}{\partial y} + Pr\left[\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right) + Ra_T(T + NS)\right] \tag{3}$$

$$\frac{\partial T}{\partial t} + u\frac{\partial T}{\partial x} + v\frac{\partial T}{\partial y} = \left[\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right] \tag{4}$$

$$\frac{\partial S}{\partial t} + u\frac{\partial S}{\partial x} + v\frac{\partial S}{\partial y} = \frac{1}{Le}\left[\frac{\partial^2 S}{\partial x^2} + \frac{\partial^2 S}{\partial y^2}\right] \tag{5}$$

The dimensionless expressions of the governing equations are obtained using the following characteristic scales H', $\rho(\alpha^2/H'^2)$, H'^2/α, α/H', $q'H'/\lambda$, and $j'H'/D$ corresponding to length, pressure, time, velocity, characteristic temperature, and characteristic concentration, respectively.

The associated boundary conditions:

$$u = v = 0 \text{ and } \frac{\partial T}{\partial x} + 1 = \frac{\partial S}{\partial x} + 1 = 0 \quad \text{for } x = 0 \text{ and } x = A; \tag{6}$$

$$u = v = 0 \text{ and } \frac{\partial T}{\partial y} = \frac{\partial S}{\partial y} = 0 \text{ for } y = 0; \tag{7}$$

$$\mathrm{u} - Pe = v = 0 \text{ and } \frac{\partial T}{\partial y} = \frac{\partial S}{\partial y} = 0 \text{ for } y = 1 \tag{8}$$

The following dimensionless governing parameters arise:

$$A = \frac{L'}{H'};\ Pr = \frac{\nu}{\alpha};\ Pe = \frac{u_0' H'}{\alpha}; Ra_T = \frac{g\beta_{\mathrm{t}} H'^4 q'}{\nu\alpha\lambda};$$
$$Le = \frac{\alpha}{D};\ N = \frac{\beta_S \Delta S^*}{\beta_T \Delta T^*} \tag{9}$$

representing: aspect ratio of the enclosure A, Prandtl number Pr, Peclet number Pe, thermal Rayleigh number Ra_T, Lewis number Le, and buoyancy ratio N, respectively.

It is worth to mention that:

$$Pe = RePr \text{ and } Ra_T = GrPr \tag{10}$$

where Re and Gr are the Reynolds and Grashof numbers, respectively.

The stream function Ψ and the vorticity Ω can be written using the horizontal component u and the vertical component v of the dimensionless velocity as follows:

$$u = \frac{\partial \psi}{\partial \mathrm{y}}; \qquad v = -\frac{\partial \psi}{\partial \mathrm{x}} \tag{11}$$

$$\Omega = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \tag{12}$$

Thus, the governing dimensionless equations reduce to:

$$\frac{\partial \Omega}{\partial t} + \frac{\partial (u\Omega)}{\partial x} + \frac{\partial (v\Omega)}{\partial y} = Pr\left[\left(\frac{\partial^2 \Omega}{\partial x^2} + \frac{\partial^2 \Omega}{\partial y^2}\right) + Ra_T\left(\frac{\partial T}{\partial x} + N\frac{\partial S}{\partial x}\right)\right] \tag{13}$$

$$\frac{\partial T}{\partial t} + \frac{\partial (uT)}{\partial x} + \frac{\partial (vT)}{\partial y} = \left[\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right] \tag{14}$$

$$\frac{\partial S}{\partial t} + \frac{\partial (uS)}{\partial x} + \frac{\partial (vS)}{\partial y} = \frac{1}{Le}\left[\frac{\partial^2 S}{\partial x^2} + \frac{\partial^2 S}{\partial y^2}\right] \tag{15}$$

and:

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = -\Omega \tag{16}$$

Local Nusselt and Sherwood numbers are used to quantify heat and mass transfer, respectively:

$$Nu(y) = \frac{q'L'}{\lambda \Delta T'} = \frac{1}{(\Delta T/A)} \tag{17}$$

$$Sh(y) = \frac{j'L'}{D\Delta S'} = \frac{1}{(\Delta S/A)} \tag{18}$$

where $\Delta T = T(0,y) - T(A,y)$ and $\Delta S = S(0,y) - S(A,y)$ represent the local dimensionless temperature and concentration difference between the two vertical walls $x = 0$ and $x = A$, respectively.

Given that numerical results will be compared to analytical ones, valid in the central part of the enclosure, and in order to avoid problems related to complex flow near vertical boundaries, Nu and Sh are quantified far from the end sides. Accordingly, and considering Eqs. (25)-(26) and two symmetrical sections with respect to the central section $(x = A/2)$, Nu and Sh become:

$$Nu(y) = \lim_{\delta x \to 0} \frac{\delta x}{\delta T} = \lim_{\delta x \to 0} \frac{1}{\left(\frac{\delta T}{\delta x}\right)} = -1/(\partial T / \partial x)_{x=A/2} \tag{19}$$

$$Sh(y) = \lim_{\delta x \to 0} \frac{\delta x}{\delta S} = \lim_{\delta x \to 0} \frac{1}{\left(\frac{\delta S}{\delta x}\right)} = -1/(\partial S / \partial x)_{x=A/2} \tag{20}$$

where, δx is the horizontal distance between the two infinitesimally close sections considered above.

The average horizontal Nusselt number, representing the overall horizontal heat transfer, and the average horizontal Sherwood number, representing the overall horizontal mass transfer, can thus be deduced respectively:

$$\overline{Nu} = \int_0^1 Nu(y)dy \tag{21}$$

$$\overline{Sh} = \int_0^1 Sh(y)dy \tag{22}$$

III. Numerical solution

To numerical solve the two-dimensional governing equations, a second-order central finite-difference scheme with a uniform mesh is used. The alternating-direction implicit method (ADI) frequently used for Newtonian fluids is implemented for the integration of vorticity, energy, and concentration equations 13-15. Further, and in order to satisfy the conservation of mass, Poisson equation (16) is solved using a point successive over-relaxation method (PSOR) with an optimum relaxation factor calculated by the Franckel formula [30]. The solution convergence is verified by the criterion $\sum_{i,j}\left|\Psi_{\mathrm{i,j}}^{k+1} - \Psi_{\mathrm{i,j}}^{k}\right| < 10^{-5}\sum_{i,j}\left|\Psi_{\mathrm{i,j}}^{k+1}\right|$, where $\Psi_{\mathrm{i,j}}^{k}$ is the value of the stream function at the k^{th} iteration level. Finally, to evaluate the vorticity on the rigid boundaries, the Woods' relation is used for its accuracy and stability.

A trial-and-error calculations are carried to determine the grid size offering the optimal compromise between computation time and solution accuracy. As a result, for $A = 24$ (found as the smallest value of A beyond which mixed convection heat and mass transfer do not change), a uniform mesh size of 341×81

is found adequate to accurately model fluid flow, temperature, and concentration distributions within the enclosure.

IV. ANALYTICAL SOLUTION: PARALLEL FLOW APPROXIMATION

Streamlines, isotherms, and iso-concentrations obtained numerically show that the flow is parallel with respect to the x-direction while the temperature and concentration profiles show linear stratification in the same horizontal direction, and that independently of the value of N. The mentioned observations confirm the existence of an analytical solution and lead to the following simplifications:

$$u(x,y) = u(y), v(x,y) = 0, \psi(x,y) = \psi(y),$$

$$T(x,y) = C_T\left(x - \frac{A}{2}\right) + \theta_T(y)$$

$$\text{and} \quad S(x,y) = C_S\left(x - \frac{A}{2}\right) + \theta_S(y) \tag{23}$$

where, C_T *and* C_S are the unknown but constant horizontal dimensionless temperature gradient and horizontal dimensionless concentration gradient, respectively. Based on the above simplifications, the dimensionless governing equations are simplified as follows:

$$\frac{d^3u(y)}{dy^3} = Ra_T E \tag{24}$$

$$C_T u(y) = \frac{d^2\theta_T(y)}{dy^2} \tag{25}$$

$$C_S u(y) = \frac{1}{Le}\frac{d^2\theta_S(y)}{dy^2} \tag{26}$$

with:

$$E = C_T + NC_S \tag{27}$$

The associated boundary conditions also reduce to:

$$u = \Psi = \frac{d\theta_T(y)}{dy} = \frac{d\theta_S(y)}{dy} = 0 \; for \; y = 0 \tag{28}$$

$$u - Pe = \Psi = \frac{d\theta_T(y)}{dy} = \frac{d\theta_S(y)}{dy} = 0 \; for \; y = 1 \tag{29}$$

The following expressions:

$$\int_0^1 u(y)dy = 0 \tag{30}$$

$$\int_0^1 \theta_T(\mathrm{y})dy = 0 \tag{31}$$

$$\int_0^1 \theta_S(\mathrm{y})dy = 0 \tag{32}$$

define the return flow, mean temperature, and mean concentration conditions, respectively.

Consequently, solving equations (24) - (25) satisfying conditions (28) - (31) results in the following expressions:

$$u(y) = \frac{Ra_T E}{12}(2y^3 - 3y^2 + y) + Pe(3y^2 - 2y) \tag{33}$$

$$\theta_T(y) = \frac{C_T \mathrm{Ra_T} E}{1440}(12y^5 - 30y^4 + 20y^3 - 1) + \frac{C_T \mathrm{Pe}}{60}(15y^4 - 20y^3 + 2) \tag{34}$$

The expression of $\theta_S(\mathrm{y})$ can be expressed in terms of $\theta_T(\mathrm{y})$, using Eqs. (25) - (26) and boundary conditions (28), (31), and (32), as follows:

$$\theta_S(y) = \frac{LeC_S}{C_T}\theta_T(y) \tag{35}$$

The stream function $\Psi(y)$ is obtained by integrating Eq. (11) considering boundary conditions (28) - (29) and equation (33):

$$\psi(y) = \frac{Ra_T}{24}E\left(y^4 - 2y^3 + y^2\right) + Pe(y^3 - y^2) \tag{36}$$

The Flow near the vertical walls is much more complicated and cannot be described with the parallel flow approximation. Thus, equation (6) describing thermal and solutale boundary conditions along the vertical walls cannot be reproduced accurately with this approximation. However, an equivalent energy and concentration flux conditions can be imposed:

$$\int_0^1 -\frac{\partial T}{\partial x}dy + \int_0^1 uTdy = \int_0^1 -\left(\frac{\partial T}{\partial x}\right)_{x=0 \; or \; x=A} dy \tag{37}$$

$$\int_0^1 -\frac{\partial S}{\partial x}dy + Le\int_0^1 uSdy = \int_0^1 -\left(\frac{\partial S}{\partial x}\right)_{x=0 \; or \; x=A} dy \tag{38}$$

In particular, in the parallel flow region and with the application of conditions (6), Eqs. (37) and (38) become:

$$C_T + 1 = \int_0^1 u(y)\theta_T(y)dy \tag{39}$$

$$C_S + 1 = Le\int_0^1 u(y)\theta_S(y)dy \tag{40}$$

By replacing $u(y)$, $\theta_T(\mathrm{y})$, and $\theta_S(\mathrm{y})$ with their respective expressions we find:

$$C_T = \frac{1}{-\frac{Ra_T^2}{362{,}880}E^2 + \frac{Ra_T Pe}{3360}E - \frac{Pe^2}{105} - 1} \tag{41}$$

$$C_S = \frac{1}{-\frac{Le^2 Ra_T^2}{362{,}880}E^2 + \frac{Le^2 \mathrm{Ra_T} Pe}{3360}E - \frac{Le^2 Pe^2}{105} - 1} \tag{42}$$

Finally, using Eq. (27), the following transcendental equation is achieved:

$$\frac{\mathrm{Le^2 Ra_T^4}}{362{,}880^2}\mathrm{E}^5 - \frac{\mathrm{Le^2 Ra_T^3 Pe}}{609{,}638{,}400}\mathrm{E}^4 + \mathrm{Ra_T^2}\left[\frac{\mathrm{Le^2 Pe^2}}{19{,}051{,}200} + \frac{1}{362{,}880}(1 + \mathrm{Le^2}) + \frac{\mathrm{Le^2 Pe^2}}{11{,}289{,}600}\right]\mathrm{E}^3 + Ra_T\left[-\frac{\mathrm{Le^2 Pe^3}}{176{,}400} - \frac{\mathrm{Pe}}{3{,}360}(1 + \mathrm{Le^2}) + \frac{\mathrm{Ra_T}}{362{,}880}(\mathrm{Le^2} + \mathrm{N})\right]\mathrm{E}^2 + \left[\frac{\mathrm{Le^2 Pe^4}}{11{,}025} + \frac{\mathrm{Pe^2}}{105}(1 + \mathrm{Le^2}) - \frac{\mathrm{Ra_T Pe}}{3{,}360}(\mathrm{Le^2} + \mathrm{N}) + 1\right]\mathrm{E} + \frac{\mathrm{Pe^2}}{105}(\mathrm{Le^2} + \mathrm{N}) + \mathrm{N} + 1 = 0 \tag{43}$$

Newton-Raphson method is used to solve Eq. (43) for the value of E and the values of C_T and C_S are deduced from Eqs. (41) and (42) for fixed values of controlling parameters $Pe, Ra_T, Le,$ and N.

Considering Eqs. (19) - (23), mean Nusselt and Sherwood numbers are constant:

$$\overline{Nu} = \frac{-1}{C_T}; \qquad \overline{Sh} = \frac{-1}{C_S} \tag{44}$$

For pure natural convection, results are obtained by considering a new cavity configuration with both walls being motionless ($u' = 0$ for $y' = 0$ and $y' = H'$). Here, the dimensionless governing equations are defined by (1)-(5) with dimensionless dynamical boundary conditions ($u(0) = u(1) = 0$). Accordingly, the mean Nusselt and Sherwood numbers in pure natural convection are expressed as follows:

$$\overline{Nu_n} = \frac{Ra_T^2}{362{,}880} E^2 + 1 \tag{45}$$

$$\overline{Sh_n} = \frac{Le^2 Ra_T^2}{362{,}880} E^2 + 1 \tag{46}$$

For $Ra_T = 0$, the convection is assured only by shear force with no buoyancy effects and the regime is qualified as pure forced convection. In this case, mean Nusselt and Sherwood numbers are given by:

$$\overline{Nu_f} = \frac{Pe^2}{105} + 1 \tag{47}$$

$$\overline{Sh_f} = \frac{Le^2 Pe^2}{105} + 1 \tag{48}$$

V. Results and Discussion

For enclosures with aspect ratio A sufficiently large and with thermal and solutal boundary condition of Neumann type, flow characteristics become independent of the aspect ratio. Thus, the analytical solution based on the parallel flow assumption is valid asymptotically in the limit of $A >> 1$. Numerical tests are conducted to determine the smallest value of A leading to numerical results reasonably close to analytical ones. The results show that beyond $A = 24$, numerical flow characteristics show an asymptotic behavior while reaching a good agreement with analytical results. This finding can be further confirmed by examining the agreement between numerical and analytical curves depicted below in the limit of the explored ranges of governing parameters ($Pe \leq 1100$, $Ra_T \leq 10^7$, $Le = 1, 3$ and $10^{-2} \leq N \leq 10^2$).

Prandtl number does not produce any change in double-diffusive mixed convection providing that $Pr \geq 10$. Accordingly, the couple (Ra_T, Pe) in which Pr is incorporated is used in the present study instead of (Gr, Re), allowing to reduce the number of governing parameters by explicitly ignoring the role of Pr and not needing its value. Mixed convection heat and mass transfer is thus governed by Peclet number Pe, thermal Rayleigh number Ra_T, Lewis number Le, and buoyancy ratio N.

A. Mixed Convection Parameter

Mixed convection flows are characterized by the important mixed convection parameter $\eta = Gr/Re^n$ (known as the modified Richardson number) providing a measure of the influence of natural convection input magnitude compared to that of forced convection on the fluid flow, with the exponent $n\,(n > 0)$ is a constant which depends on the flow configuration and the associated thermal and solutal boundary conditions. Mixed convection regime is commonly defined in the range of $\eta_f \leq \eta \leq \eta_n$, where η_f and η_n are the lower and the upper bounds of the mixed convection flow regime, respectively. Outside this mixed convection region, either pure forced convection ($\eta \leq \eta_f$) or the pure natural convection ($\eta_f \leq \eta$) analysis can be used to describe flow, temperature, and concentration fields accurately.

Note that by setting $Le = 1$, heat and mass transfer end up with similar diffusion characteristics which permits highlighting the inferences of mixed convection parameter alone.

In the literature, most of the studies only report results for pure natural or pure forced convection, making it crucial to distinguish the conditions under which a given convection may be regarded as dominant natural or dominant forced convection from those under which it may be considered as mixed. In the present paper, a convection will be considered to be effectively dominated by natural or forced regime if the transfer rate deviates by no more than 5% from the value associated with the completely pure natural or forced convection, respectively. Accordingly, to differentiate the three convective regimes natural, mixed, and forced, the following relative differences are introduced:

$$\varepsilon_{Nu_n} = \frac{|\overline{Nu} - \overline{Nu_n}|}{\overline{Nu_n}}; \qquad \varepsilon_{Nu_f} = \frac{|\overline{Nu} - \overline{Nu_f}|}{\overline{Nu_f}} \tag{49}$$

with $\overline{Nu_n}$ and $\overline{Nu_f}$ are the average Nusselt numbers for pure natural and pure forced convections, respectively. Natural regime dominates the overall convection for $\varepsilon_{Nu_n} < 5\%$ while forced regime dominates for $\varepsilon_{Nu_f} < 5\%$. If neither is the case, the convective regime is considered mixed.

Based on Eq. (49), the calculations performed for Pe, Ra_T and N varying in their respective range allow to construct the diagram in Fig. 2. The locations of points $(\log(Ra_T), \log(Pe))$ obtained analytically (solid lines) and numerically (symbols) and corresponding approximately to $(\varepsilon_{Nu_n} = 5\%)$ and $(\varepsilon_{Nu_f} = 5\%)$ are organized in parallel straight lines. These results can be correlated in the form of the mixed convection parameter as follows:

$$\frac{Ra_T}{Pe^{3.0}} = \eta_n \text{ and } \frac{Ra_T}{Pe^{3.0}} = \eta_f \tag{50}$$

for natural and forced convection, respectively, where η_n and η_f values are given in Table 1 and illustrated in Fig. 3 with dashed lines.

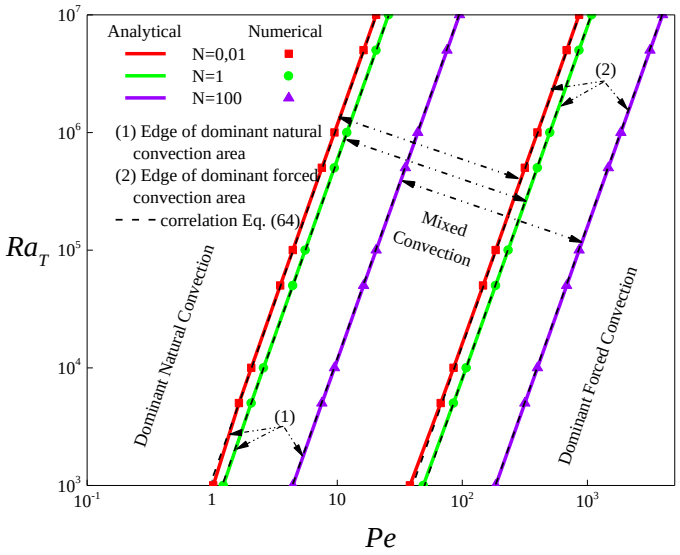


Fig. 2. Diagram illustrating the dominance regions of natural, mixed, and forced convective regimes for $Le = 1$ and different characteristic values of N.

TABLE I. VALUES OF η_n AND η_f FOR DIFFERENT VALUES OF N.

Convection regime	Natural convection η_n	Forced convection η_f
$N = 0.01$	1129.482	0.015313
$N = 1.0$	575.5498	0.008051
$N = 100$	11.88542	0.00016

First thing to notice is the good agreement between the analytical and numerical results, where both straight lines defined by Eq. (50) delimit three regions. The first zone below the straight line (2) experiences a dominant forced convection. The second region above the straight line (1) is characterized by a dominating natural convection regime. The third and final zone is delimited by the two straight lines where both natural and forced convection are of comparable magnitudes (mixed convection regime). As for the effect of governing parameters, it is obvious that the transition from dominant natural regime to mixed regime then to dominant forced regime requires higher Pe values as thermal Rayleigh number increases. This due to the effect of Ra_T on enhancing buoyancy forces leading to strengthen natural convection regime; thus, the transition requires higher shear effect (higher Pe). Increasing buoyancy ratio, which corresponds to strengthening solutal buoyancy force compared to thermal buoyancy force, also delays the shift from dominant natural regime towards forced one as the figure shows that the transition for the same value of Ra_T necessitates higher shear force as N rises. One more thing is that strengthening thermal buoyancy force by a factor of 100 compared to solutal one ($N = 0.01$) have smaller effect on the convective regimes' dominance areas in contrast to strengthening mass buoyancy force by the same factor ($N = 100$) which strongly affect the mixed regime boundaries. All in all, for cooperating buoyancy forces case ($N > 0$), higher buoyancy ratios ($N > 1$) corresponding to higher solutal volume force strongly promote natural convection contribution in the overall convective regime.

To recap, mixed convection regime in a single lid-driven rectangular cavity filled with Newtonian fluid and submitted to uniform horizontal density of heat and mass fluxes is bounded as follows:

$$0.0153 < \frac{Ra_T}{Pe^{3.0}} < 1129.482 \quad (51)$$

$$0.0080 < \frac{Ra_T}{Pe^{3.0}} < 575.5498 \quad (52)$$

$$0.00016 < \frac{Ra_T}{Pe^{3.0}} < 11.8854 \quad (53)$$

for $N = 0.01, 1,$ and 100, respectively.

B. *Effect of Buoyancy Ratio*

The effects of buoyancy ratio N on flow characteristics $|\Psi c|$, $\overline{Nu}$, and $\overline{Sh}$ are illustrated in Fig.3 for $Ra_T = 10^4$, $Le = 3$, and various values of Pe. Three convective regimes can be deciphered; first, for small values of N ($N \leq 1$), $|\Psi c|$, $\overline{Nu}$, and $\overline{Sh}$ are indifferent to variation of N as thermal buoyancy force dominates the solutal one, while the enhancing effect of shear force magnitude (Pe) is noticeable indicating the strong forced convection. Next, as the value of the buoyancy ratio moves away from 1, a transitional regime is observed as the three flow characteristics simultaneously begin to increase slowly. This can be attributed to the enhanced mass buoyancy force compared to the thermal buoyancy force. The value of N signaling this change is related to the value of Pe expressing the shear force magnitude, where as Pe increases the transition is delayed as stronger solutal volume force (higher value of N) is required to impact fluid flow and heat and mass transfer. The slight improvement in transfer rates is allowed at this early stage (the solutal buoyancy force is not fully dominant) given the cooperating roles of sliding wall and solutal buoyancy force acting together from left to right. Finally, a regime for which solutal volume force completely dominates leading $|\Psi c|$, $\overline{Nu}$, and $\overline{Sh}$ to increase linearly with N while the effect of the moving wall vanishes. These results are in line with the early findings concerning the role of increasing buoyancy ratio beyond unity in promoting natural regime contribution in the overall convection; while for low values of N ($N < 1$), the sliding wall strongly influences the convective regime indicating a dominating forced convection.

VI. CONCLUSIONS

This work investigates double-diffusive mixed convection within a single lid-driven rectangular cavity filled with Newtonian fluid and subjected to uniform heat and mass fluxes along the vertical short boundaries (Neumann type conditions) while the horizontal walls are considered insulated and impermeable. The study establishes two solutions, a numerical one based on the finite difference method is used to solve the governing equations and an analytical solution based on the

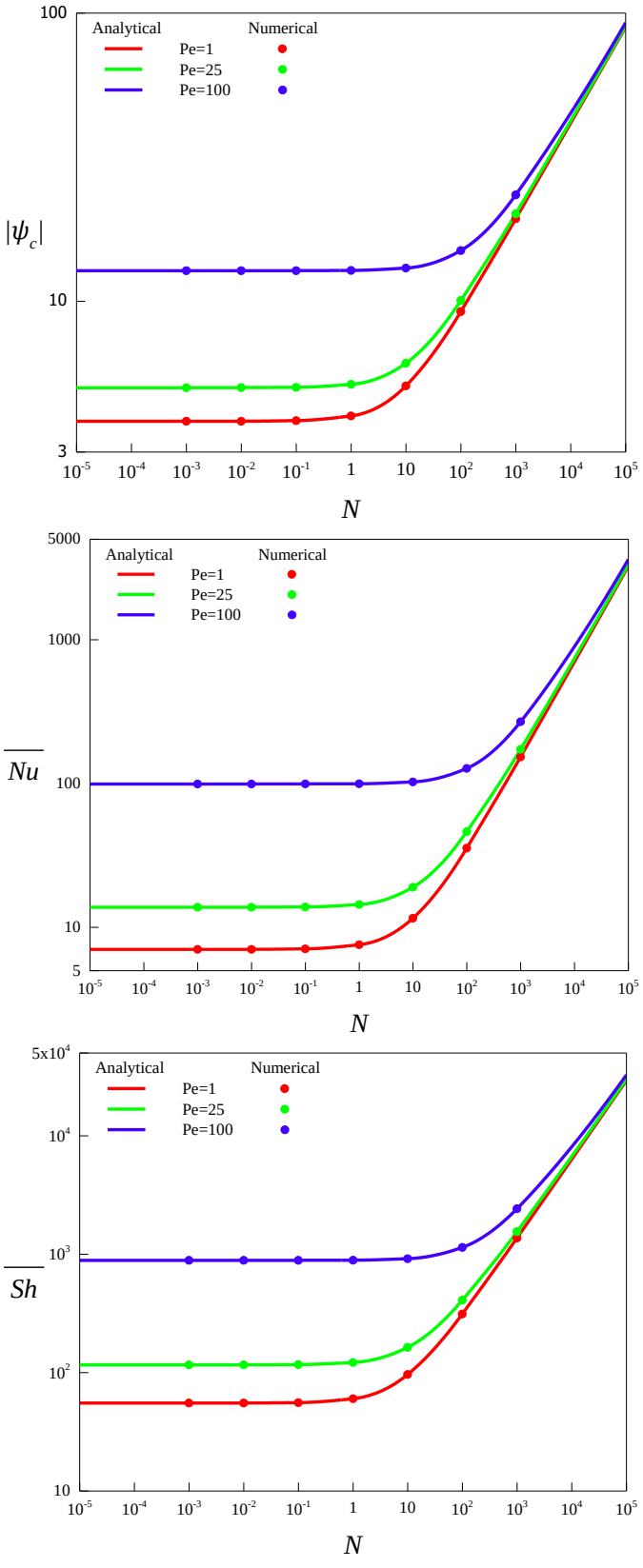


Fig. 3. Evolutions of stream function (top), Nusselt number (middle), and Sherwood number (bottom) with buoyancy ratio for $A = 24, Ra_T = 10^4, Le = 3, N = 1$, and various values of Pe.

parallel flow assumption. For a sufficiently large value of the aspect ratio ($A \geq 24$), fluid flow and heat and mass transfer become invariant with respect to A. Consequently, double-diffusive mixed convection in a shallow rectangular enclosure filled with Newtonian fluid is mainly governed by Peclet number Pe, thermal Rayleigh number Ra_T, Lewis number Le, and buoyancy ratio N. Both solutions, numerical and analytical, agree perfectly for a wide range of governing parameters: $(0.1 \leq Pe \leq 4 \times 10^3), (1 \leq Ra_T \leq 10^7), (Le = 1 \; and \; 3)$, and $(10^{-3} \leq N \leq 10^3)$, validating the numerical code, the analytical approach, and justifies the choice of $A = 24$ as the large aspect ratio approximation value.

The ensuing key findings can be summarized as follows:

- The mixed convection parameter $\frac{Ra_T}{Pe^{3.0}}$ accurately delimits natural, mixed, and forced convection dominance regions, with the following limits:

$$0.015313 < \frac{Ra_T}{Pe^{3.0}} < 1129.482, 0.008051 < \frac{Ra_T}{Pe^{3.0}} < 575.5498, \text{and}$$
$$0.00016 < \frac{Ra_T}{Pe^{3.0}} < 11.88542$$

  are found to bound mixed convection regime for $N = 0.01, N = 1.0,$ and $N = 100,$ respectively. While natural and forced convection are found to be dominant for higher and lower values of the mixed convection parameter, respectively.
- Peclet number and thermal Rayleigh number strongly influence the transition from one convective regime to another, where increasing Ra_T delays the transition towards forced convection as buoyancy forces driving natural convection enhances. In contrast, increasing Pe accelerates such trasition given the associated high shear force responsible for forced flow.
- Increasing buoyancy ratio enhances fluid circulation and heat and mass transfer owing to strong solutal buoyancy force.
- The buoyancy ratio N increases the input of natural regime in overall convection given the dominating role of solutal buoyancy force ($N > 1$) leading to diminish the sliding wall effect for higher values of N. However, strengthening thermal buoyancy force ($N < 1$) have no significant impact on flow characteristics allowing shear force to dominate.

REFERENCES

[1] K. B. Saleem, L. Koufi, A. K. Alshara, and L. Kolsi, "Double-diffusive natural convection in a solar distiller with external fluid stream cooling," Int. J. Mech. Sci. 181, 2020.

[2] L. Jun and B. Bofeng, "Thermosolutal convection and solute segregation during the vertical Bridgman growth of Hg1-xCdxTe single crystals," J. Cryst. Growth 311, pp. 38-46, 2008.

[3] S. Yang-Cheng and T. Shu-Min, "PIV study on the development of double diffusive convection during the solidification effected by lateral cooling for a super-eutectic binary solution," Appl. Therm. Eng. 29, pp. 2773–2782, 2009.

[4] J. Serrano-Arellano, J. Xamán and G. Álvarez, "Optimum ventilation based on the ventilation effectiveness for temperature and CO2 distribution in ventilated cavities," Int. J. Heat Mass Transfer, 62, pp. 9–21, 2013.

[5] A. A. Abbasian Arani, E. Kakoli, and N. Hajialigol, "Double-diffusive natural convection of Al_2O_3-water nanofluid in an enclosure with partially active side walls using variable properties," Journal of Mechanical Science and Technology 28 (11), pp. 4681-4691, 2014.

[6] X. Liang, B. Peng, and Z. F. Tian, "Complex transition of double-diffusive convection in a rectangular enclosure with height-to-length ratio equal to 4: Part II," International Journal of Heat and Mass Transfer 135, pp. 247–261, 2019.

[7] J. T. Hu, S. J. Mei, D. Liu, F. Y. Zhao, and H. Q. Wang, "Hydromagnetic double diffusive moisture convection from an inclined enclosure inserted with multiple heat-generating electronic modules," International Journal of Thermal Sciences 159, 2021.

[8] I. I. Nosonov, and M. A. Sheremet, "Conjugate mixed convection in a rectangular cavity with a local heater," International Journal of Mechanical Sciences 136, pp. 243–251, 2018.

[9] W. Zhou, Y.Yan, X. Liu, H. Chen, and B. Liu, "Lattice Boltzmann simulation of mixed convection of nanofluid with different heat sources in a double lid-driven cavity," International Communications in Heat and Mass Transfer 97, pp. 39-46, 2018.

[10] S. Sivasankaran, S. S. Ananthan, M. Bhuvaneswari and A. K. Abdul Hakeem, "Double diffusive mixed convection in a lid-driven cavity with non- uniform heating on sidewalls," Sādhanā 42, pp. 1929–41, 2017.

[11] D. Senthil Kumar, K. Murugesan and H. R. Thomas, "Numerical Simulation of Double Diffusive Mixed Convection in a Lid-Driven Square Cavity Using Velocity-Vorticity Formulation," Numerical Heat Transfer, Part A: Applications: An International Journal of Computation and Methodology 54(9), pp. 837-65, 2008.

[12] GH.R. Kefayati, "Double-diffusive mixed convection of pseudoplastic fluids in a two-sided lid-driven cavity using FDLBM," J. Taiwan Inst. Chem. Eng. 45 (5), pp. 2122–2139, 2014.

[13] M. A. Waheed, "Mixed convective heat transfer in rectangular enclosures driven by a continuously moving horizontal plate," Int. J. Heat Mass Transfer 52, pp. 5055–5063, 2009.

[14] M. M. Abdelkhalek, "Mixed convection in a square cavity by a perturbation technique," Comput. Mater. Sci. 42, pp. 212–219, 2008.

[15] A.A. Abbasian Arani et al., "Numerical simulation of double-diffusive mixed convection in an enclosure filled with nanofluid using Bejan's heatlines and masslines," Alexandria Eng. J., 2017.

[16] GH. R. Kefayati, "Mesoscopic simulation of magnetic field effect on double-diffusive mixed convection of shear-thinning fluids in a two-sided lid-driven cavity," J. Mol. Liq., 2014.

[17] M. A. Teamah, and W. M0 El-Maghlany, "Numerical simulation of double-diffusive mixed convective flow in rectangular enclosure with insulated moving lid," Int J Therm Sci. 49, pp. 1625–38, 2010.

[18] S. Bettaibi, et al., "Hybrid LBM-MRT model coupled with finite difference method for double-diffusive mixed convection in rectangular enclosure with insulated moving lid," Physica, 2015.

[19] Y. Tizakast, M. Kaddiri, and M. Lamsaadi, "Double-diffusive mixed convection in rectangular cavities filled with non-Newtonian fluids," International Journal of Mechanical Sciences 208, 2021.

Charctersitics of Double-Diffusive Mixed Convection Opposing Mechanism Inside a Double Lid-Driven Cavity Submited to Heat and mass Fluxes

Youssef Tizakast, Mourad Kaddiri
Industrial Engineering Laboratory
Sultan Moulay Slimane University
BP 523, Beni-Mellal 23000, Morocco
youssef.tizakast@usms.ma

Mohamed Lamsaadi
Research Laboratory in Physics and Sciences for Engineers (LRPSI)
Polydisciplinary Faculty, Sultan Moulay Slimane University
BP 592, Beni-Mellal, Morocco

***Abstract*—The present paper investigates numerically and analytically double-diffusive mixed convection within a shallow rectangular cavity filled with Newtonian fluid and subjected to uniform heat and mass fluxes along the short vertical walls while the insulated horizontal boundaries slide in opposite directions. The sliding walls and imposed gradients act in opposite directions resulting in an opposing flow case. The study established two separate approaches to solve the problem; a numerical solution using finite volume method and analytical one derived based on the parallel flow assumption. The two approaches showed a good agreement for the examined range of governing parameters, namely: Peclet number, thermal Rayleigh number, Lewis number, and buoyancy ratio. A mixed convection parameter is defined to separate the predominance zones of natural, mixed, and forced convective regimes. The opposing flow nature of the problem lead to decreasing fluid flow and heat and mass transfer in the mixed convection regime, while increasing Peclet number results in a dominant forced convection; thus, inducing a more active convection and demolish the enhancing role of thermal Rayleigh number observed in the natural regime.**

Keywords—double-diffusive; mixed Convection opposing flow; heat and mass transfer; lid-driven cavities; parallel flow; finite volume method

I. INTRODUCTION

Double-diffusive convection describes fluid flows driven by buoyancy effects due to simultaneously applied temperature and concentration gradients with different diffusion rates. Such phenomenon occurs in many industrial applications including, but not limited to, solar energy, ventilation, crystal growth, and food processing. On the other hand, Double-diffusive mixed convection results from combined buoyancy forces generated by heat and mass gradients and shear force driven by the moving walls (or other external source). Yet, despite the importance of double-diffusive fluid flows as we mentioned before, the literature review shows that an important part of published papers covers mixed convection heat transfer. Munshi et al. [1] investigated mixed convection inside a lid-driven square cavity subjected to a constant heat flux on the bottom wall in the presence of an internal elliptic body. The results show that the studied phenomenon is mainly governed by Richardson number (Ri), Grashof number (Gr), and Reynolds number (Re). Cheng [2] numerically examined mixed convection within a two-dimensional square cavity. The horizontal walls are maintained isothermally with lower temperature on the top sliding wall, while the vertical walls are considered adiabatic. The authors reported on the effect of simultaneously increasing Grashof and Reynolds numbers on heat transfer. Moreover, the study provides a validation of Nusselt number (Nu) correlations found in the literature. Aydm [3] conducted a numerical study of the same phenomenon this time in a square cavity with left wall sliding at a constant velocity from bottom to top. Authors used mixed convection parameter Gr/Re^2 to separate natural, mixed, and forced heat transport regimes; thus, found that the mixed regime zone in the aiding buoyancy case is narrower compared to opposing-buoyancy case. Nosonov and Sheremet [4] considered a rectangular cavity to investigate mixed convection in the presence of local heater. The study showed that while increasing Richardson number, convection enhances within the cavity the increasing buoyancy force deforms the forced flow. Other works studied the same phenomenon while considering different combinations of physical configurations and associated temperature gradients [5-8].

On the other hand, fewer studies (compared to the studies covering mixed convection flow and heat transfer) investigated double-diffusive mixed convection fluid flows. Kumar et al. [9] studied double-diffusive convection inside square lid-driven cavity. They reported that negative buoyancy ratios give rise to aiding flows in the case of imposed opposing thermal and solutal gradients along the vertical direction. Ababaei et al. [10] considered mixed convection heat and mass transfer within a right-angled trapezoidal cavity with the bottom wall partially heated and salted. The results showed that for Lewis number $Le = 0.1$, the conduction mass transfer dominates the mass transfer phenomenon leading to mass-lines virtually perpendicular to the iso-concentration lines. Moreover, increasing Le decreases the average Nusselt number. Nath et al. [11] studied double-diffusive mixed convection inside a backward facing channel filled with a Newtonian nanofluid. The authors found that while heat transfer enhances while increasing buoyancy ratio from -10 to $+10$ owing to strengthened buoyancy force, mass transfer showed an opposite trend as the nanoparticles slow down the mass transport. Tizakast et al. [12]

numerically and analytically studied mixed convection heat and mass transfer inside a rectangular cavity filled with non-Newtonian fluids. The study considered the case where buoyancy forces and shear one act in the same direction (aiding flow case). The results showed that while the fluid behavior strongly affects heat and mass transfer rates in the natural regime, such effect vanishes when forced regime dominates the overall convection.

In summary, the literature shows a striking lack of studies considering shallow rectangular cavities with associated thermal and solutal boundary conditions of Neumann type (i.e., imposed heat and mass fluxes to the boundaries) when investigating double-diffusive mixed convection. Furthermore, the opposing flow case (natural and forced convection acts in opposite directions) of double-diffusive mixed convection within rectangular double lid-driven cavity with boundary conditions of Neumann type has not been considered yet. To fill in the gap, the present paper investigates opposing flow mixed convection heat and mass transfer in a double lid-driven horizontal rectangular cavity with applied thermal and solutal fluxes to the motionless boundaries. The study establishes two separate solutions, numerical and analytical, and compare them in order to validate the numerical code and the analytical approach.

II. Physical Problem And Governing Equations

The present work studies the cavity in Fig. 1, where a shallow horizontal rectangular cavity of height H' and length L' subjected to uniform density of heat and mass fluxes q' and j', respectively, along the vertical boundaries while the horizontal walls are considered insulated and impermeable. Furthermore, the bottom wall slides from left to right with constant velocity u_0', while the top wall moves in the opposite direction with the same uniform velocity u_0' (i.e., the moving walls acts in the opposite direction of the imposed heat and mass fluxes); whereas the vertical ones are motionless.

The dimensionless governing equations are as follows:

$$\frac{\partial \mathrm{u}}{\partial \mathrm{x}} + \frac{\partial \mathrm{v}}{\partial \mathrm{y}} = 0 \tag{1}$$

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} = -\frac{\partial p}{\partial x} + Pr\left[\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right] \tag{2}$$

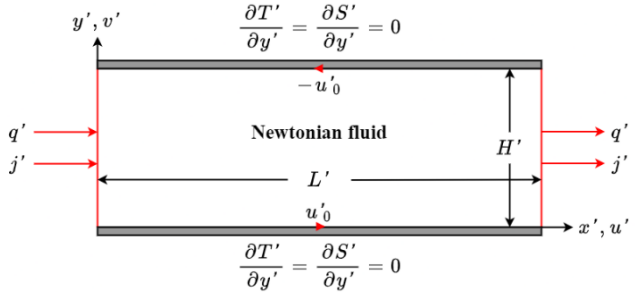


Fig. 1. Physical configuration of the cavity along the associated boundary conditions and coordinates system.

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} = -\frac{\partial p}{\partial y} + Pr\left[\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right) + Ra_T(T + NS)\right] \tag{3}$$

$$\frac{\partial T}{\partial t} + u\frac{\partial T}{\partial x} + v\frac{\partial T}{\partial y} = \left[\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right] \tag{4}$$

$$\frac{\partial S}{\partial t} + u\frac{\partial S}{\partial x} + v\frac{\partial S}{\partial y} = \frac{1}{Le}\left[\frac{\partial^2 S}{\partial x^2} + \frac{\partial^2 S}{\partial y^2}\right] \tag{5}$$

associated with the boundary conditions:

$$u = v = 0 \text{ and } \frac{\partial T}{\partial x} + 1 = \frac{\partial S}{\partial x} + 1 = 0 \quad \text{for } x = 0 \text{ and } x = A; \tag{6}$$

$$u - Pe = v = 0 \text{ and } \frac{\partial T}{\partial y} = \frac{\partial S}{\partial y} = 0 \text{ for } y = 0; \tag{7}$$

$$\mathrm{u} + Pe = v = 0 \text{ and } \frac{\partial T}{\partial y} = \frac{\partial S}{\partial y} = 0 \text{ for } y = 1 \tag{8}$$

Consequently, the following dimensionless parameters appear:

$$A = \frac{L'}{H'};\ Pr = \frac{\nu}{\alpha};\ Pe = \frac{u_0' H'}{\alpha};\ Ra_T = \frac{g\beta_t {H'}^4 q'}{\nu\alpha\lambda};\ Le = \frac{\alpha}{D};$$

$$N = \frac{\beta_S \Delta S^*}{\beta_T \Delta T^*} \tag{9}$$

representing: aspect ratio of the enclosure A, Prandtl number Pr, Peclet number Pe, thermal Rayleigh number Ra_T, Lewis number Le, and buoyancy ratio N, respectively.

To study the flow structure, the stream function ψ is introduced:

$$u = \frac{\partial \psi}{\partial y}; \quad v = -\frac{\partial \psi}{\partial x}\ (\psi = 0 \text{ on all boundaries}) \tag{10}$$

III. Numerical Approach

To numerically solve the governing equations (1)-(5) associated with boundary conditions (6)-(8), the finite volume method with SIMPLE algorithm is adopted. The discretized equations system is solved using a line-by-line tridiagonal matrix algorithm. A threshold $MAX\left(\frac{f^{k+1}-f^k}{f^{k+1}}\right) < 10^{-7}$, where $\mathrm{f^k}$ is the value of u, v, p, T, or S at the k^{th} iteration level, is implemented to test convergence.

The average horizontal Nusselt number, describing the overall horizontal heat transfer, and the average horizontal Sherwood number, describing the overall horizontal mass

transfer, are given, respectively, in terms of temperature T and concentration S:

$$\overline{Nu} = \int_0^1 -1/(\partial T \,/\, \partial x)_{x=A/2} \tag{11}$$

$$\overline{Sh} = \int_0^1 -1/(\partial S \,/\, \partial x)_{x=A/2} \tag{12}$$

IV. Analytical Solution: Parallel Flow Assumption

The approximate analytical solution adopts the following simplifications:

$$u(x,y) = u(y), v(x,y) = 0, \psi(x,y) = \psi(y),$$

$$T(x,y) = C_T\left(x - \frac{A}{2}\right) + \theta_T(y)$$

$$\text{and} \quad S(x,y) = C_S\left(x - \frac{A}{2}\right) + \theta_S(y) \tag{13}$$

where, C_T *and* C_S are the unknown but constant horizontal dimensionless temperature and concentration gradient, respectively. Based on these approximations, the ensuing simplified non-dimensional governing equations are:

$$\frac{\mathrm{d}^3 u(y)}{\mathrm{d}y^3} = Ra_T(C_T + NC_S) \tag{14}$$

$$C_T u(y) = \frac{\mathrm{d}^2\theta_T(y)}{\mathrm{d}y^2} \tag{15}$$

$$C_S u(y) = \frac{1}{Le}\frac{\mathrm{d}^2\theta_S(y)}{\mathrm{d}y^2} \tag{16}$$

while the related boundary conditions become as follows:

$$u - Pe = \Psi = \frac{d\theta_T(y)}{dy} = \frac{d\theta_S(y)}{dy} = 0 \text{ for } y = 0 \tag{17}$$

$$u + Pe = \Psi = \frac{d\theta_T(y)}{dy} = \frac{d\theta_S(y)}{dy} = 0 \text{ for } y = 1 \tag{18}$$

The following relations:

$$\int_0^1 u(y)dy = 0 \tag{19}$$

$$\int_0^1 \theta_T(\mathrm{y})dy = 0 \tag{20}$$

$$\int_0^1 \theta_S(\mathrm{y})dy = 0 \tag{21}$$

describe return flow and mean temperature and concentration conditions, respectively.

Solving the dimensionless governing equations lead to finding $\overline{Nu}$ and $\overline{Sh}$ with:

$$\overline{Nu} = \frac{-1}{C_T}; \qquad \overline{Sh} = \frac{-1}{C_S} \tag{22}$$

V. Results and Discussion

The present study uses the couple (Ra_T, Pe) in which Pr, who's effect on mixed convection is well established, is incorporated instead of (Gr, Re), allowing us to explicitly ignore the role of Pr and not need its value, thus, reducing the number of governing parameters. On the other hand, numerical tests showed that after an aspect ratio value of $A = 24$, the fluid flow characteristics $|\psi_c|$, $\overline{Nu}$ and $\overline{Sh}$ become insensitive to A variation while achieving a good agreement between numerical and analytical results. Consequently, Newtonian double-diffusive mixed convection with opposing flow developed within closed cavities is governed by Peclet number Pe, thermal Rayleigh number Ra_T, buoyancy ratio N, and the Lewis number Le.

A. Mixed convection parameter

The ratio Gr/Re^{n} (modified Richardson number known also as the mixed convection parameter) compares the buoyancy forces generating natural convection to shear force driven by the moving walls. The exponent n is related to the cavity geometry, the fluid nature, and the associated boundary conditions. The mixed convection parameter allows to separate the three convective regimes, namely: natural, mixed, and forced convection. Note that, we set $Le = 1$ (same diffusion rate for heat and mass transfer) to only focus on the implications of mixed convection parameter.

In order to separate the three convective regimes, a criterion is chosen as to be acceptable in practice, where a convection regime will be considered pure (either natural or forced) if transfer rate deviates by no more than 5% from the value obtained in the case of the completely pure convection. Thus, to separate natural, mixed, and forced convection dominance areas, the following relative differences are adopted:

$$\varepsilon_{Nu_n} = \frac{|\overline{Nu} - \overline{Nu_n}|}{\overline{Nu_n}}; \qquad \varepsilon_{Nu_f} = \frac{|\overline{Nu} - \overline{Nu_f}|}{\overline{Nu_f}} \tag{23}$$

where $\overline{Nu_n}$ and $\overline{Nu_f}$ are the mean Nusselt numbers corresponding to pure natural and pure forced convections, respectively. It is assumed that natural convection dominates the overall convective regime for $\varepsilon_{Nu_n} < 5\%$ while forced convection is predominant when $\varepsilon_{Nu_f} < 5\%$. If neither is the case, the convective regime is qualified as mixed.

Based upon Eq. (18), we construct the diagram shown in Fig. 2 with Pe *and* Ra_T varying in their inspected ranges. The

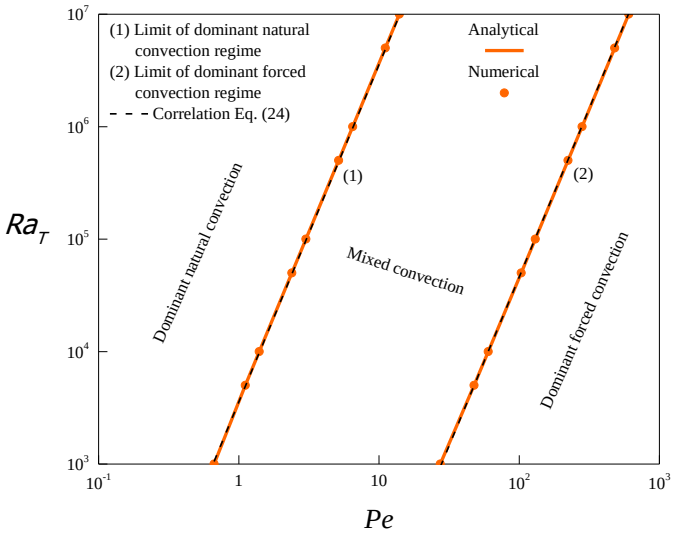


Fig. 2. Diagram showing the dominance regions of natural, mixed, and forced convective regimes at $Le = 1$ and $N = 1$.

locations of points $(log\ (Ra_T), log(Pe))$ obtained analytically (solid lines) and numerically (symbols) and corresponding approximately to $(\varepsilon_{Nu_n} = 5\%)$ and $(\varepsilon_{Nu_f} = 5\%)$ are aligned in parallel straight lines. The results can be correlated in the form of the mixed convection parameter as follows:

$$\frac{Ra_T}{Pe^{3.0}} = \eta_n = 3553.367 \text{ and } \frac{Ra_T}{Pe^{3.0}} = \eta_f = 0.046 \quad (24)$$

corresponding to natural and forced convection, respectively, where η_n and η_f values are presented in Fig. 2 with dashed lines.

The excellent agreement between numerical and analytical solutions is clearly visible. The first zone, located above the straight line (1) experiences a dominating natural convection regime. The second zone, located below the straight line (2) is characterized by a dominating forced convection regime. Finally, the third zone in between the two straight lines, is associated with a regime where both buoyancy forces and shear one are of comparable magnitudes (mixed convection regime).

B. Effect of Peclet Number

The variations of flow characteristics $|\psi_c|$, $\overline{Nu}$, and $\overline{Sh}$ with Peclet number Pe are illustrated in Fig. 3 for $Le = 2, N = 1$, and different values of Ra_T. For low values of Pe associated with weak shear force magnitude, natural convection dominates the overall convection given that the results agree with the ones obtained in the case of pure natural convection presented with dotted lines where Pe augmentation do not influence fluid circulation intensity and heat and mass transfer rates. After a given value of Pe, the three characteristics begin to decrease as the shear force and buoyancy forces, acting in opposite directions, compete against each other and hurt the transport phenomenon indicating a regime shift from natural to mixed convection. The said value of Pe varies with Ra_T, as increasing

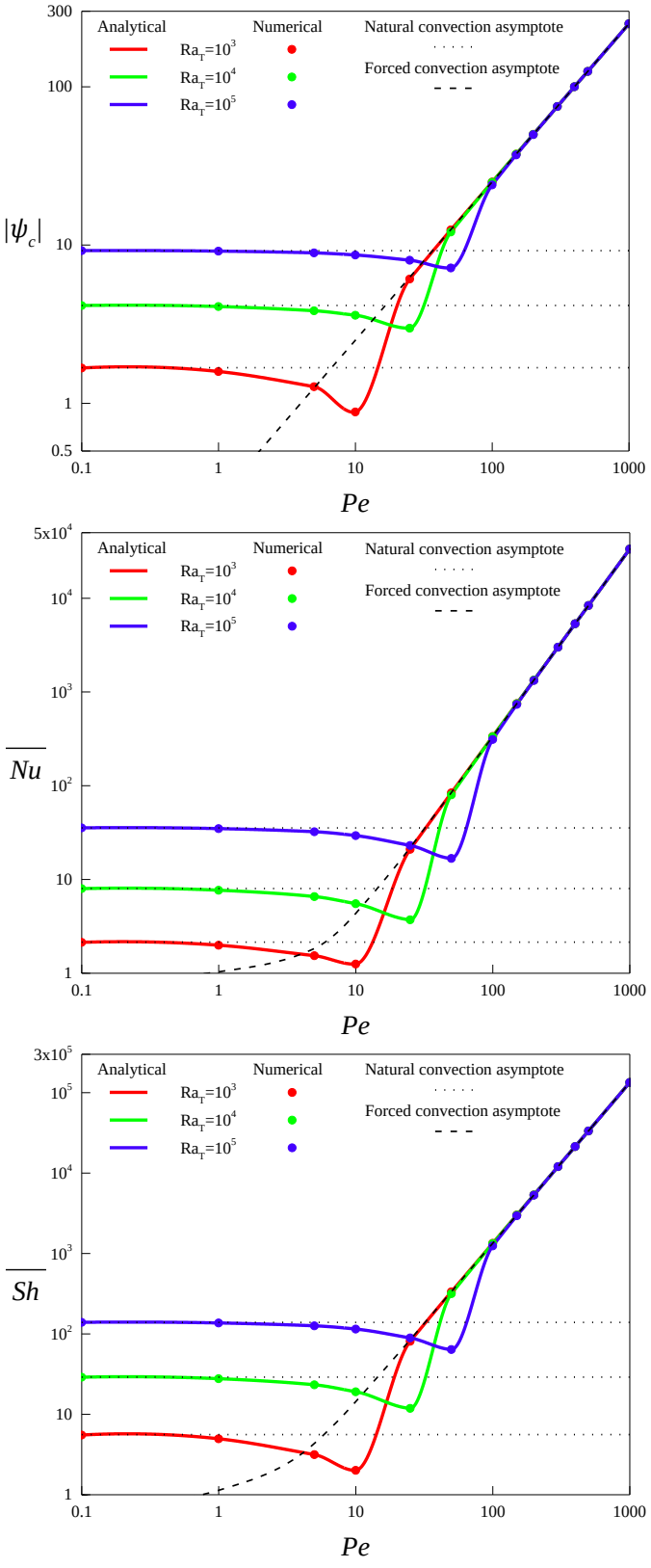


Fig. 3. Evolutions of stream function (top), Nusselt number (middle), and Sherwood number (bottom) with Peclet number for $A = 24, Le = 5, N = 1$, and various values of Ra_T.

thermal Rayleigh number strengthens buoyancy forces; thus, delays the transition from natural regime to mixed one as it requires higher Pe value (higher shear force). Next, after reaching a minimum, $|\psi_c|$, $\overline{Nu}$, and $\overline{Sh}$ begin to rise slowly as shear force intensifies and overthrow buoyancy forces as the main driving force. Finally, for higher values of Peclet number, all quantities increase in a monotonous way as forced convection dominates fluid flow and heat and mass transfer seeing that the results agree with the ones obtained in pure forced convection, illustrated with dashed lines.

VI. Conclusions

The present paper considered the opposing flow case of double-diffusive mixed convection within a shallow rectangular cavity subjected to uniform heat and mass fluxes along the vertical short walls and occupied with a Newtonian fluid. The study establishes and compare two separate solutions, a numerical approach using the finite volume method to solve the governing equations and an analytical approach based on the parallel flow approximation. Both solutions showed a good agreement for the considered range of governing parameters Peclet number Pe, thermal Rayleigh number Ra_T, Lewis number Le, and buoyancy ratio N. Thus, verify the numerical code, the parallel flow assumption, and the choice of large aspect ratio approximation value $A = 24$. As a result, the main conclusions can be listed as follows:

- The mixed convection parameter $\frac{Ra_T}{Pe^{3.0}}$ perfectly outlines the dominance zones of natural, mixed, and forced convective regimes, where the following limits:
$$0.046 < \frac{Ra_T}{Pe^{3.0}} < 3553.367$$
identify the mixed convection regime. Whereas, natural/(forced) convection is dominant for higher/ (lower) values of the mixed convection parameter.
- In dominant natural regime, increasing Ra_T enhances fluid circulation and heat and mass transfer owing to increasing buoyancy forces magnitude.
- In mixed convection regime (natural and forced convection are of comparable scales), increasing Pe hurts fluid circulation and heat and mass transfer instead of promoting it given that buoyancy forces and shear one act in opposite directions.
- Increasing Peclet number to higher values result in a dominant forced convection where fluid circulation and heat and mass transfer enhancement is significant.

References

[1] M. J. H. Munshi, M. A. Alim, M. Ali, and M. S. Alam, "A Numerical Study of Mixed Convection in Square Lid-Driven with Internal Elliptic Body and Constant Flux Heat Source on the Bottom Wall," J. Sci. Res. 9 (2), pp. 145-158, 2017.

[2] T. S. Cheng, "Characteristics of mixed convection heat transfer in a lid-driven square cavity with various Richardson and Prandtl numbers," International Journal of Thermal Sciences 50, pp. 197-205, 2011.

[3] O. Aydm, "Aiding and opposing mechanisms of mixed convection in a shear- and buoyancy-driven cavity," Int. Comm. Heat Mass Transfer 26 (7), pp. 1019-1028, 1999.

[4] I. Nosonov, and M. A. Sheremet, "Conjugate mixed convection in a rectangular cavity with a local heater," International Journal of Mechanical Sciences 136, pp. 243–251, 2018.

[5] M. A. Waheed, "Mixed convective heat transfer in rectangular enclosures driven by a continuously moving horizontal plate," Int J Heat Mass Transfer 52, pp. 5055–63, 2009.

[6] M. M. Abdelkhalek, "Mixed convection in a square cavity by a perturbation technique," Comput Mater Sci 42, pp. 212–9, 2008.

[7] K. M. Gangawane, and H. F. Oztop, "Mixed convection in the semi-circular lid-driven cavity with heated curved wall subjugated to constant heat flux for non-Newtonian power-law fluids," International Communications in Heat and Mass Transfer 114, 2020.

[8] W. Zhou, Y. Yan, X. Liu, H. Chen, and B. Liu, "Lattice Boltzmann simulation of mixed convection of nanofluid with different heat sources in a double lid-driven cavity," International Communications in Heat and Mass Transfer 97, pp. 39-46, 2018.

[9] D. S. Kumar, K. Murugesan, and H. R. Thomas, "Numerical Simulation of Double Diffusive Mixed Convection in a Lid-Driven Square Cavity Using Velocity-Vorticity Formulation," Numerical Heat Transfer, Part A: Applications: An International Journal of Computation and Methodology, 54(9), pp. 837-65, 2008.

[10] A. Ababaei, M. Abbaszadeh, A. Arefmanesh, and A. J. Chamkha, "Numerical simulation of double-diffusive mixed convection and entropy generation in a lid-driven trapezoidal enclosure with a heat source," Numerical Heat Transfer, Part A: Applications, pp. 702-20, 2018.

[11] R. Nath, and M. Krishnan, "Numerical study of double diffusive mixed convection in a backward facing step channel filled with Cu-water nanofluid," International Journal of Mechanical Sciences 153–154, pp. 48–63, 2019.

[12] Y. Tizakast, M. Kaddiri, and M. Lamsaadi, "Double-diffusive mixed convection in rectangular cavities filled with non-Newtonian fluids," International Journal of Mechanical Sciences 208, 2021.

Natural convection of fluids with temperature-dependent viscosity in an enclosure with a localized non-uniform heat source on the left wall

DAGHAB Hamza
Industrial Engineering Laboratory
Sultan Moulay Slimane University
Béni-Mellal, Morocco

LAMSAADI Mohamed
Research Laboratory in Physics and Sciences for Engineers
Sultan Moulay Slimane University
Béni-Mellal, Morocco

ERRITALI Ilham
Industrial Engineering Laboratory
Sultan Moulay Slimane University
Béni-Mellal, Morocco

KADDIRI Mourad
Industrial Engineering Laboratory
Sultan Moulay Slimane University
Béni-Mellal, Morocco

ARROUB Ismail
Research Laboratory in Physics and Sciences for Engineers
Sultan Moulay Slimane University
Béni-Mellal, Morocco

Abstract—**Buoyancy-driven convection of fluids with temperature-dependent viscosity inside a square enclosure with a localized non-uniform heat source mounted on the bottom of the left wall is numerically investigated. The other side is maintained at a relatively cold temperature, while the horizontal walls and the unheated portion of the left wall are insulated. The heat source is assumed to be isothermal with a linearly varying temperature. The governing equations were solved using finite volume method on a staggered grid system. The computational results are presented in the form of isotherm and streamline plots and average Nusselt numbers. The effects of Rayleigh number (Ra), Pearson number (m) and heater length (W) on flow structure and heat transfer are discussed.**

Keywords— finite volume method, heat source, natural convection, non-uniform heating, numerical simulation.

I. INTRODUCTION

Free convection in cavities is significant in a variety of engineering applications, including solar energy use, heat removal from electronic equipment, exchange of heat between building and environment, nuclear reactor design and food processing. An extensive bibliography on free convection in cavities may be found in the review article by Ostrach [1]. majority of natural convection investigations in finite cavities can be broadly classified into two categories: enclosures with differential lateral heating and enclosures heated from below and cooled from above. Many studies concerning the differentially heated vertical walls case have been investigated for both Newtonian [2] and non-Newtonian fluids [3,4], where the momentum and thermal transports initiate inside the enclosure once a finite temperature difference is occurred between the side walls. When the lower wall of the enclosure is maintained at a higher temperature than the top wall while the sidewalls are adiabatic, the classical Rayleigh-Bénard convection is observed within the cavity which starts only when a critical Rayleigh number is exceeded. The classical Rayleigh-Bénard convection has also been studied for both Newtonian [5] and non-Newtonian fluids [6].

The aforementioned studies only concentrated on natural convection in totally heated enclosures. However, it is useful to analyze natural convection in cavities with locally heated walls, which is often encountered in cooling of electronic circuits. In this context, Aydin et al. [7] have numerically investigated natural convection of air confined in a square cavity with localized heating from below, which is considered to be isothermally heated at a constant temperature, and symmetrical cooling from the vertical walls, while the upper one and the rest of the bottom wall are adiabatic. Finite difference procedure has been used to solve the governing equations. It has been found that an increase in Rayleigh number and heat source length enhances the heat transfer inside the enclosure. Corcione and Habib [8] have performed a numerical investigation in a tilted square cavity discretely heated from the left side. They conducted simulations for various values of heater size and location, Rayleigh and Prandtl numbers, and inclination angle. They have reported that the heat transfer got better for the intermediate positions of the heat source and diminished when the heat source was positioned on the upper or lower of the cavity. They have also noted that heat transfer increased with the Rayleigh and Prandtl numbers, and positive inclination angles. Saravanan and Sivaraj [9] numerically analyzed natural convection in an air-filled cavity with a non-

uniform heat source which set up centrally on the bottom wall. It has been found that for Gr = 106 non-uniform heating of the line source enhances the overall heat transfer rate markedly compared to uniform heating of the heat source whereas for Gr = 107 its effect is marginal.

Cooling heat source problem has also been dealt with non-Newtonian fluids because many situations require cooling efficiency improvement, and thus an amelioration of heat transfer is needed, in order to maintain the adequate working conditions and not to surpass certain permitted temperatures. Accordingly, Hassan et al. [10] have numerically investigated natural convection problem in a square cavity whose vertical walls are maintained at a cold temperature and the top wall is adiabatic, while a heating element is located at the lower wall. The cavity is filled with viscoplastic fluids modeled by Bingham fluids. It has been found that the increase in Bingham number causes the convection effect and the fluid circulation to decrease inside the cavity due to the augmentation of flow resistance. This last configuration has also been studied by Yigit et al. [11] by using power-law fluids. It has been noticed that the enclosure performs better with shear-thinning fluids (n < 1) from the perspective of enhancing heat transfer performance.

Majority of the published studies dealing with natural convection cooling of a localized heat source are customarily based on the common assumption that the fluid properties are constant with respect to the temperature except for the fluid density in the buoyancy force term, which obeys the Boussinesq approximation. However, for many real fluids, the viscosity is effectively dependent upon temperature. Consequently, this lack of information has been addressed in the present investigation by numerical simulations for natural convection of fluids with temperature-dependent viscosity in a square enclosure with a partial localized non-uniform heating, which is achieved by a heat source producing a linearly varying temperature and placed at the bottom of the left wall, while the right side is maintained at a relatively cold temperature. Thus, the main purpose of the present work is to examine the effects of Rayleigh number, Pearson number and heat source length on the fluid flow and the resulting heat transfer.

II. MATHEMATICAL FORMULATION

The physical model under consideration is depicted in Figure 1. It consists of a two-dimensional square cavity whose right vertical wall is maintained at a cold temperature T_c'. A heat source producing a linearly varying temperature and placed at the bottom of the left wall, while the horizontal walls the remaining part of left wall are adiabatic, without slipping conditions at all the solid boundaries. The cavity contains a fluid with temperature-dependent viscosity, whose viscosity can be expressed by the following exponential law [11]:

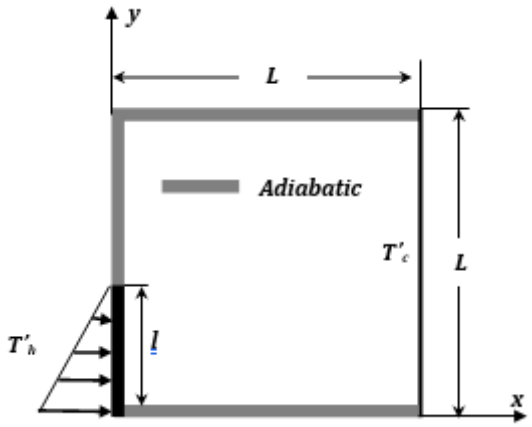


FIGURE 1. Sketch of the Cavity and Boundary Conditions.

$$\mu_a = k\exp\left(-b(T'-T_r')\right) \tag{1}$$

the flow is assumed to be laminar and steady; the fluid density obeys the Boussinesq approximation, and the viscous dissipation effects are negligible.

Under the aforementioned assumptions, and by using the following characteristic scales:

$$X=\frac{x}{L}, Y=\frac{y}{L}, U=\frac{uL}{\alpha}, V=\frac{vL}{\alpha}, T=\frac{T'-T_c'}{\Delta T^*}$$

The dimensionless governing equations written in terms of velocity vector components, (*U*, *V*), pressure, *P*, and temperature, *T*, in Cartesian coordinate system (*X*, *Y*), are:

$$\frac{\partial U}{\partial X}+\frac{\partial V}{\partial Y}=0 \tag{2}$$

$$U\frac{\partial U}{\partial X}+V\frac{\partial U}{\partial Y}=-\frac{\partial P}{\partial X}+Pr\left[\begin{array}{l}\mu_a\left(\frac{\partial^2 U}{\partial X^2}+\frac{\partial^2 U}{\partial Y^2}\right)+\\ 2\frac{\partial \mu_a}{\partial X}\frac{\partial U}{\partial X}+\frac{\partial \mu_a}{\partial Y}\left(\frac{\partial U}{\partial Y}+\frac{\partial V}{\partial X}\right)\end{array}\right] \tag{3}$$

$$U\frac{\partial V}{\partial X}+V\frac{\partial V}{\partial Y}=-\frac{\partial P}{\partial Y}+Pr\left[\begin{array}{l}Ra\,T+\mu_a\left(\frac{\partial^2 V}{\partial X^2}+\frac{\partial^2 V}{\partial Y^2}\right)+\\ 2\frac{\partial \mu_a}{\partial Y}\frac{\partial V}{\partial Y}+\frac{\partial \mu_a}{\partial X}\left(\frac{\partial U}{\partial Y}+\frac{\partial V}{\partial X}\right)\end{array}\right] \tag{4}$$

$$U\frac{\partial T}{\partial X}+V\frac{\partial T}{\partial Y}=\frac{\partial^2 T}{\partial X^2}+\frac{\partial^2 T}{\partial Y^2} \tag{5}$$

The corresponding boundary conditions are:

$$\begin{array}{l}U=V=T-1-\frac{a}{W}(1-2Y)=0 \quad for\\ X=0\ and\ \ 0\le Y\le W\end{array} \tag{6}$$

$$U=V=\frac{\partial T}{\partial Y}=0 \quad for \quad X=0\ and\ \ W<Y\le 1 \tag{7}$$

$$U=V=T=0 \quad for \quad X=1\ and\ \ 0\le Y\le 1 \tag{8}$$

$$U = V = \frac{\partial T}{\partial Y} = 0 \quad for \quad Y = 0 \ \ and \ \ 0 \le X \le 1 \tag{9}$$

$$U = V = \frac{\partial T}{\partial Y} = 0 \quad for \quad Y = 1 \ and \ \ 0 \le X \le 1 \tag{10}$$

The dimensionless parameters appearing now in the present problem are:

$$\Pr = \frac{(k/\rho)}{\alpha} \qquad (\text{Prandtl number}) \tag{11}$$

$$Ra = \frac{g\beta\Delta T H^3}{(k/\rho)\alpha} \qquad (\text{Rayleigh number}) \tag{13}$$

$$m = b\Delta T^* = -\frac{1}{K_T}\frac{dK_T}{T} = \frac{d(ln(K_T/k))}{dT} \quad (\text{Pearson number}) \tag{14}$$

$$W = \frac{l}{L} \qquad (\text{heater dimensionless length}) \tag{15}$$

The steady solution has been used to measure the heat transfer rate of the cavity by calculating the local Nusselt number on the heat source element defined as follow:

$$Nu = \frac{hL}{k} \tag{16}$$

Where *h* is the convection heat transfer coefficient. By using the dimensionless parameters, the equation (16) becomes:

$$Nu = \frac{\partial T}{\partial X} \tag{17}$$

The average Nusselt is obtained by integrating equation (17) along the heat source.

$$\overline{Nu} = \int_0^W Nu \, dY \tag{18}$$

III. NUMERICAL METHOD

All of the governing differential equations (2-5) can be written in the following form of the general transport equation:

$$\frac{\partial}{\partial X}\left(U\Phi - \Gamma\frac{\partial \Phi}{\partial X}\right) + \frac{\partial}{\partial Y}\left(V\Phi - \Gamma\frac{\partial \Phi}{\partial Y}\right) = S_\Phi \tag{19}$$

Where Φ is the working variable which represents U, V or T, Γ is $Pr\mu_a$ for the momentum equations and 1 for the energy equation, and SΦ is the source term.

The equation (19) has to be converted into a linear algebraic equation in order to obtain a numerical solution. Consequently, a finite volume approach is considered for the spatial discretization. The computational field is divided into a set of control volumes ΔV around nodes *P* (Figure 2).

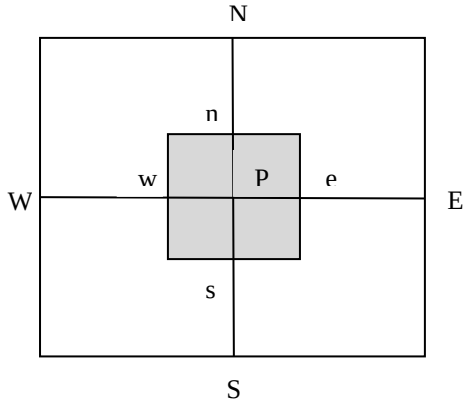


Figure 2. Control Volume

By integrating the equation (19) over control volume and using the divergence theorem, the final form of the discretized equations relating the variable Φ_p to its neighboring grid point values can be written in each control volume as:

$$A_P\Phi_P = A_W\Phi_W + A_E\Phi_E + A_S\Phi_S + A_N\Phi_N + S_\varphi \tag{20}$$

A. Solution method

The obtained discretized equations for every control volume in the computational field consist of a set of linear algebraic equations, which then can be solved iteratively by means of the line by line technique based on the tridiagonal matrix algorithm (TDMA) .

Since the pressure is an unknown in the momentum equations, a derivation of pressure equation is evidently required to solve the discretized equations for obtaining the temperature and kinematic fields. Consequently, Semi-Implicite method for pressure linked equation (SIMPLE) algorithm [49], which allows the transformation of the continuity equation to the pressure equation, is used. In order to avoid physically unrealistic solutions, such as the checkerboard pressure and velocity distribution, the staggered grid system is imposed.

The solution process is reiterated until attaining the convergence criterion:

$$Max\left(\frac{\Phi^{n+1} - \Phi^n}{\Phi^{n+1}}\right) \le 10^{-7} \tag{23}$$

Where Φ = (*U*, *V*, *T*).

IV. RESULTS AND DISCUSSION

Figure.3 exhibits streamlines (on the left) and isotherms (on the right) for different values of Ra, (a) W = 0.25 and (b) W = 0.75 and two values of Pearson number m = 0 (plotted by black solid lines) and m = 5 (plotted by red solid lines). The figure demonstrates that a circulation cell is formed in the clockwise direction for all values of Rayleigh numbers. It is clear from this figure that the increase in Rayleigh number leads to the elongation of streamlines and isotherms along the horizontal direction to form thermal and hydrodynamic boundary layers. Furthermore, the flow intensity and average Nusselt number improve with increasing Rayleigh numbers

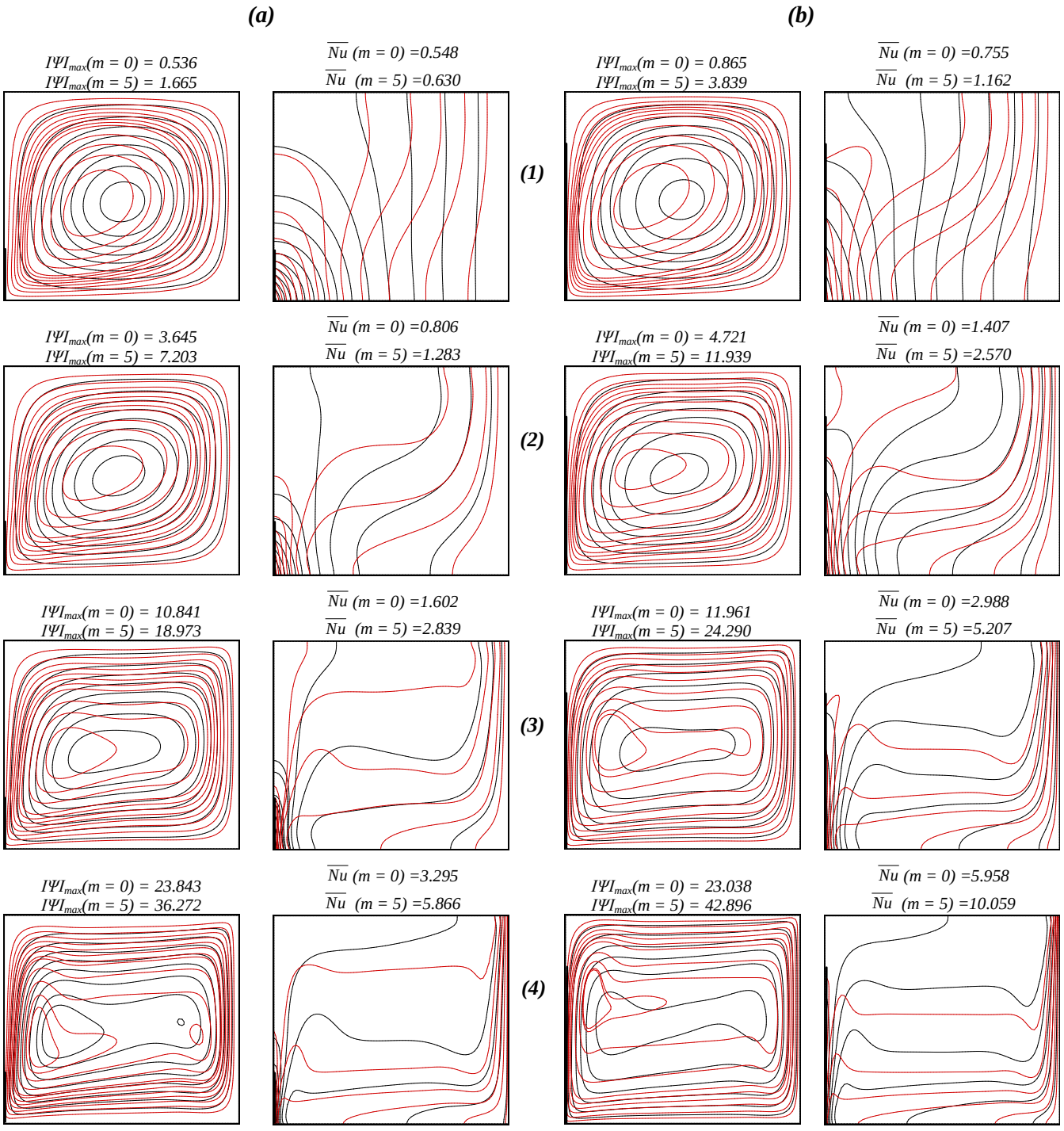


Figure 2. Streamlines (left) and Isotherms (right) for different values of m: m = 0 (plotted by black solid lines), m = 5 (plotted by red solid lines), different values of Ra: (1) Ra = 10^3, (2) Ra = 10^4, (3) Ra = 10^5, (4) Ra = 10^6, and different values of heater length: (a) W = 0.2, (b) W = 0.8

due to the strengthening of the buoyancy forces in comparison to the viscous forces. This means that convective transport enhances inside the cavity with increasing Rayleigh number. Fig.3 also shows the effect of the heater length. It is clearly

seen that heat transfer enhances with heater size as expected due to increasing of heating surface. Moreover, Fig. 3 presents the effect of the viscosity parameter. The flow intensity increases with increasing Pearson numbers due to the decrease in the viscosity as m rises. This is can be shown in equation (1). consequently, the convection ameliorates with increasing m, and then, the average Nusselt is ameliorated with the increase in Pearson number. Furthermore, the flow core moves toward the top part of the heat source, where the higher temperature inside the enclosure is located, due to the diminution of viscosity in this region.

I. CONCLUSION:

In the current investigation, a numerical study has been performed to investigate the steady state, natural convection for temperature-dependent fluids in a square enclosure with a partial localized non-uniform heating, which is achieved by a heat source producing a linearly varying temperature and placed at the bottom of the left wall. The impacts of different control parameters, which are Rayleigh number, Pearson number and heated source length, on heat transfer performance and fluid flow inside the cavity have been analyzed. It has been noticed that the flow intensity and average Nusselt number increase with increasing Rayleigh number due to the strengthening of convective transport. Increasing the Pearson number reduces the apparent viscosity of the fluid, which leads to the enhancement of convection transport due to weakening of resistance to the fluid motion offered by the fluid viscosity. Furthermore, increasing the heat source length enhances the convective transport inside the enclosure.

REFERENCES

[1] S. Ostrach, Natural convection in enclosures, ASME J. Heat Transfer 110 (1988) 1175–1190.

[2] De Vahl Davis, G. (1983). Natural convection of air in a square cavity: a bench mark numerical solution. International Journal for numerical methods in fluids, 3(3), 249-264.

[3] Lamsaadi, M., Naimi, M., Hasnaoui, M., & Mamou, M. (2006). Natural convection in a vertical rectangular cavity filled with a non-Newtonian power law fluid and subjected to a horizontal temperature gradient. Numerical Heat Transfer, Part A: Applications, 49(10), 969-990.

[4] Daghab, H., Kaddiri, M., Raghay, S., Lamsaadi, M., & El Harfi, H. (2021). Numerical study of natural convection for generalized second-grade fluids confined in a square cavity subjected to horizontal heat flux. International Journal of Heat and Technology, 39(2), 345-354.

[5] Ouertatani, N., Cheikh, N. B., Beya, B. B., & Lili, T. (2008). Numerical simulation of two-dimensional Rayleigh–Bénard convection in an enclosure. Comptes Rendus Mécanique, 336(5), 464-470.

[6] Aydin, O., & Yang, W. J. (2000). Natural convection in enclosures with localized heating from below and symmetrical cooling from sides. International Journal of Numerical Methods for Heat & Fluid Flow, 10(5), 518-529.

[7] M. Corcione, and E. Habib, "Buoyant heat transport in fluids across tilted square cavities discretely heated at one side". International Journal of Thermal Sciences 49, 797-808 (2010).

[8] Saravanan, S., & Sivaraj, C. (2011). Natural convection in an enclosure with a localized nonuniform heat source on the bottom wall. International Journal of Heat and Mass Transfer, 54(13-14), 2820-2828.

[9] Hassan, M. A., Pathak, M., & Khan, M. (2013). Natural convection of viscoplastic fluids in a square enclosure. Journal of Heat Transfer, 135(12).

[10] Yigit, S., Battu, M., Turan, O., & Chakraborty, N. (2019). Free convection of power-law fluids in enclosures with partially heating from bottom and symmetrical cooling from sides. International Journal of Heat and Mass Transfer, 145, 118782.

[11] Balmforth, N. J., & Craster, R. V. (2001). Geophysical aspects of non-Newtonian fluid mechanics. Geomorphological Fluid Mechanics, 34-51.

Convection naturelle dans une cavité chauffée latéralement

K.el marrakchi
Laboratoire de Recherche en Physique et Sciences pour l'Ingénieur
Faculté Polydisciplinaire, Université Sultan Moulay Slimane
Béni Mellal, Maroc
elmarrakchikamal@gmail.com

M.lamsaadi
Laboratoire de Recherche en Physique et Sciences pour l'Ingénieur
Faculté Polydisciplinaire, Université Sultan Moulay Slimane
Béni Mellal, Maroc
elmarrakchikamal@gmail.com

***Abstract*—Le présent travail est dédié à l'étude de la convection naturelle dans une cavité rectangulaire bidimensionnelle remplie d'un fluide de Bingham. L'écoulement est initie en imposant aux parois verticales une densité de flux de chaleur uniforme et les parois horizontales de l'enceinte sont supposées adiabatiques et imperméables. La résolution analytique est basée sur l'approximation de l'écoulement parallèle, tandis que l'étude numérique est faite en utilisant la méthode des volumes finis. Les résultats obtenus montrent un bon accord avec les deux approches (analytique et numérique). L'effet de nombre de Bingham sur la structure de l'écoulement et la distribution de la température au sein de la cavité, ainsi que sur les profils de vitesse et de température sera analysé.**

Keywords—fluide de Bingham, convection naturelle, volumes finis, écoulement parallèle.

I. INTRODUCTION

Le transfert de chaleur par convection naturelle est fréquemment rencontré aussi bien dans la nature que dans l'industrie. Les domaines d'application sont nombreux et variés. Pour les fluides avec une limite d'élasticité, c'est-à-dire des matériaux qui se comportent comme des solides rigides pour des contraintes de cisaillement inférieures à une limite d'élasticité critique mais qui s'écoulent pour des contraintes de cisaillement plus élevées. Vola et al. [1] et Turan et al. [2] ont étudié numériquement la convection naturelle bidimensionnelle des fluides à limite d'élasticité suivant le modèle de Bingham dans des enceintes carrées avec des parois verticales chauffées différemment. Turan et al. [2] ont démontré que pour les fluides newtoniens et les écoulements à faible nombre de Bingham, le nombre moyen de Nusselt Nu croît avec l'augmentation de Pr pour une valeur donnée du nombre de Rayleigh. En revanche, le comportement inverse a été observé pour les fluides de Bingham pour de grandes valeurs du nombre de Bingham. Vola et al. [1] ont examiné l'impact des effets de la limite d'élasticité sur les schémas d'écoulement et les champs de température. Dans Turan et al. [3] les effets des conditions aux limites thermiques des parois latérales ont été traités pour des enceintes carrées soumises aux conditions aux limites CWHF.

Toutefois les travaux antérieurs, portant sur l'étude ce phénomène, ont visés essentiellement le cas où la cavité est carrée. En revanche, très peu de travaux ont été consacrés à la convection naturelle dans les fluides de Bingham dans une cavité rectangulaire horizontale.

L'objectif de ce travail est d'étudier l'effet de nombre de Bingham, sur la structure de l'écoulement et la distribution de la température au sein de la cavité, ainsi que sur les profils de vitesse et de température en mode de convection naturelle lors de l'écoulement d'un fluide de Bingham dans une cavité rectangulaire horizontale.

II. EQUATIONS DE LA CONVECTION ET CONDITIONS AUX LIMITES ASSOCIEE

Dans la présenté étude, on s'intéresse à l'étude de la convection naturelle dans un fluide de Bingham confiné dans une cavité rectangulaire de hauteur H ′ et de longueur L′. Les axes de coordonnées x′ et y′ sont respectivement orientés suivant les directions horizontale et verticale. Les parois verticales de la cavité sont soumises à un flux de chaleur q′ uniforme tandis que les parois horizontales sont considérées comme adiabatiques et imperméables.

Les équations gouvernant la convection, exprimées en formulation de vitesses, pression et température, et les conditions aux limites correspondantes se présentent sous forme adimensionnelles comme suit :

A. Equations de conservation

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0 \qquad (1)$$

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}=-\frac{\partial P}{\partial x}+\mathrm{Pr}\left[\mu_a\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right)+2\frac{\partial \mu_a}{\partial x}\frac{\partial u}{\partial x}+\frac{\partial \mu_a}{\partial y}\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)\right] \qquad (2)$$

$$\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}=-\frac{\partial P}{\partial y}+\mathrm{Pr}\left[\mu_a\left(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}\right)+2\frac{\partial \mu_a}{\partial y}\frac{\partial v}{\partial y}+\frac{\partial \mu_a}{\partial x}\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)+RaT\right] \quad (3)$$

$$\frac{\partial T}{\partial t}+u\frac{\partial T}{\partial x}+v\frac{\partial T}{\partial y}=\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2} \quad (4)$$

L'équation constitutive du fluide de Bingham, proposée par Papanastasiou [4], est jugée satisfaisante par plusieurs auteurs [5,6] et représentatifs d'un fluide de Bingham idéal.

$$\mu_a=1+\frac{Bn}{\dot{\gamma}}\left(1-e^{-M\dot{\gamma}}\right) \quad (5)$$

M étant un paramètre adimensionnel qui représente la croissance exponentielle. Des chercheurs [5-7] recommandent de prendre M= 1000.

B. Conditions aux limites

$$u=v=\frac{\partial T}{\partial x}+1=0 \text{ pour x=0 et A} \quad (6)$$

$$v=\frac{\partial T}{\partial y}=0 \text{ pour y=0 et y=1} \quad (7)$$

$$u=0 \text{ pour y=0 et y=1} \quad (8)$$

III. MÉTHODES DE RÉSOLUTION

A. Approche numérique

Dans le présent travail, les équations de conservation, présentées dans le paragraphe 2, sont résolues par le biais de la méthode des volumes finis et l'algorithme SIMPLER dans un système de grille uniforme [8]. Le choix d'une telle technique, réside dans les avantages qu'elle offre en termes de stabilité numérique, vis-à-vis des paramètres gouvernants, de convergence et de conservation des flux sur chaque volume élémentaire.

Le choix de la grille de maillage dépend du rapport d'aspect A. Des essais numériques ont été nécessaires pour optimiser le temps et la précision des calculs. Ainsi la grille uniforme de 321x81 est trouvée suffisante pour modéliser avec précision les champs d'écoulement de température dans une cavité ayant A = 24. Cette configuration d'écoulement et de température, obtenu numériquement pour $Ra=10^3$ et des conditions aux limites hydrodynamiques et thermiques appliquées sur les parois sont présentées sur la figure 1. Comme on peut le voir, l'écoulement présente un aspect parallèle et une stratification thermique dans la région centrale de la cavité, ce qui y prouve l'existence d'une solution analytique.

B. Approximation de l'écoulement parallèle

Sur la base des observations effectuées, juste ci-dessus, les simplifications suivantes s'imposent :

$$u(x,y)=u(y)=\frac{\partial \psi}{\partial y}, v(x,y)=-\frac{\partial \psi}{\partial x}=0 \quad (9)$$

$$T(x,y)=C_T(x-A/2)+\theta_T(y) \quad (10)$$

La constante C_T représente le gradient horizontal de température. Elle est déterminée en intégrant l'équation (4), sur le volume de contrôle choisi arbitrairement dans la figure 1, en prenant en compte les conditions aux limites (6) en raccordant avec la région de l'écoulement parallèle. Soit :

$$C_T+1=\int_0^1 u(y)\,\theta_T(y)\,dy \quad (11)$$

Moyennant une telle approximation, le système d'équations (1)-(8) devient :

$$\frac{d^2}{\partial y^2}\left[\mu_a\frac{du}{dy}\right]=RaC \quad (12)$$

$$\frac{d^2\theta_T}{dy^2}=C_T u \quad (13)$$

$$\frac{d\theta_T}{dy}=0 \text{ en } y=0 \text{ et 1} \quad (14)$$

A ces équations, s'ajoute les conditions (6-8) et celle d'écoulement de retour, traduisant la conservation du débit :

$$\int_0^1 u(y)\,dy=0 \quad (15)$$

Et celle exprimant la nullité de la température adimensionnelle moyenne

$$\int_0^1 \theta_T(y)\,dy=0 \quad (16)$$

L'intégration des équations (12-13) associées aux conditions (14-16), conduit aux expressions de la vitesse et de la température.

IV. RÉSULTATS ET DISCUSSION

L'examen de figure 2, illustrant les lignes de courant et les isothermes obtenues pour A = 24, $Ra=10^3$ et différentes valeurs de Bn, permet d'affirmer l'existence d'une solution analytique, pour le présent problème, étant donne le parallélisme des lignes de courant et la stratification des champs de température. Ces figures aussi montrant que l'écoulement est parallèle dans la bonne partie centrale de la cavité et monocellulaire et rendent bien compte de l'excellent accord entre les deux types de solutions. La figure 3 montre la forme des profils de vitesse établis d'un écoulement dans une cavité rectangulaire pour différentes valeurs de la contrainte seuil Bn. la vitesse maximale diminue avec l'augmentation du nombre Bingham Bn, ce qui confirme que la force de convection s'affaiblit avec une valeur croissante du nombre Bingham Bn en raison de l'affaiblissement de la force de flottabilité par rapport à aux

forces visqueuses pour une valeur donnée du nombre de Rayleigh nominal Ra. Afin de démontrer les effets du nombre de Bingham Bn sur la convection naturelle dans les fluides de Bingham, la distribution de température au centre de la cavité est illustrée à la figure 4 pour différentes valeurs de Bn à un nombre de Rayleigh nominal Ra = 10^3. De plus, la non-linéarité de la variation de température diminue avec l'augmentation de la valeur de Bn pour une donnée de valeurs du nombre de Rayleigh nominal Ra. Cette diminution suggère essentiellement que les effets de la convection s'affaiblissent à mesure que la valeur du nombre de Bingham Bn augmente.

Les figures 3 et 4, dans lesquelles sont reportés les profils de vitesse et de température respectivement, pour Ra = 10^3 et diverses valeurs de nombre de Bingham, confirme l'excellent accord entre les résultats des deux approches avec une erreur relative ne dépassant pas 1**%**.

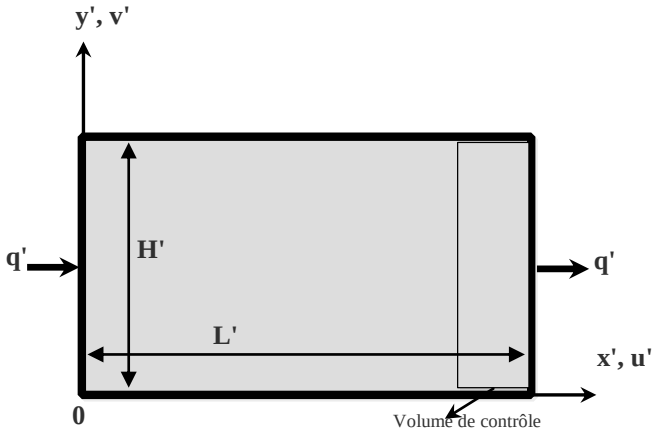


Fig. 1. Configuration étudiée.

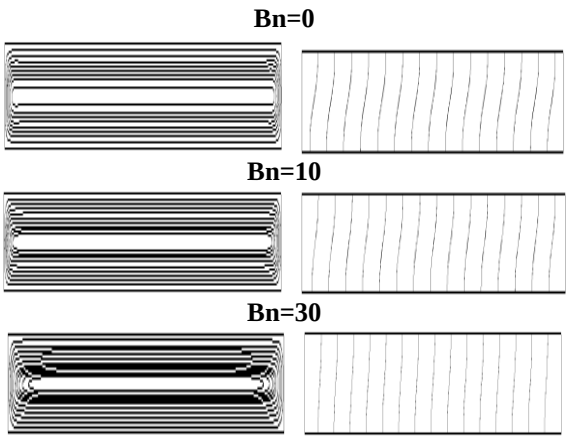


Fig. 2. Lignes de courant (gauche) et isothermes (droite) pour Ra = 103,A=24 et différentes valeurs de Bn.

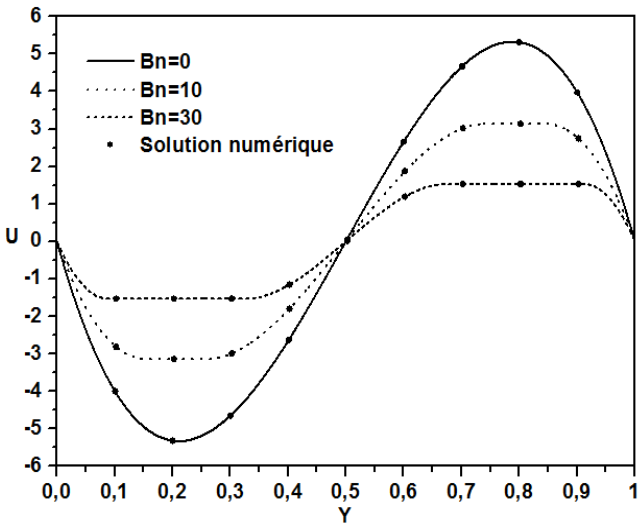


Fig. 3. Profils de la vitesse horizontale pour Ra=10^3 et différentes valeurs de Bn.

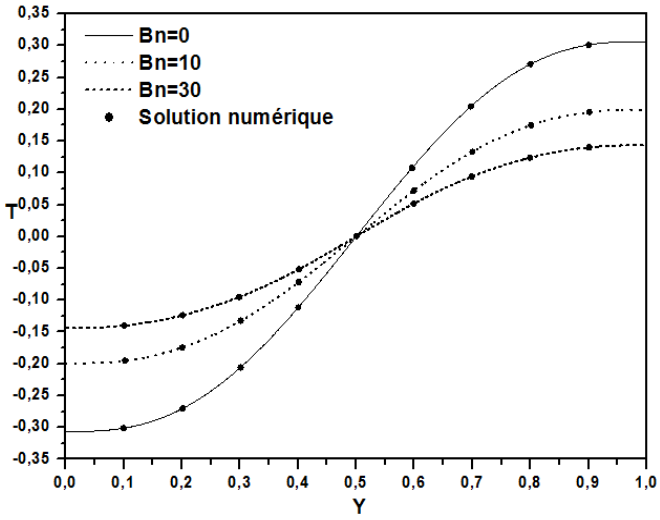


Fig. 4. Profils de la vitesse horizontale pour Ra=10^3 et différentes valeurs de Bn.

V. Conclusion

Dans ce travail, nous avons réalisé une étude numérique et analytique de la convection naturelle dans une cavité rectangulaire allongée, remplie par un fluide de Bingham. L'approximation de Boussinesq est utilisée pour établir le modèle mathématique décrivant le phénomène. Dans le but de résoudre numériquement les équations adimensionnelles régissantes, la méthode des volumes finis, a été élaboré. Une solution analytique, basée sur l'approximation de l'écoulement parallèle a été déterminée. Il a été observé que l'écoulement présenté un aspect parallèle et une stratification thermique un peu loin des bords latéraux de la cavité lorsque le rapport de forme de la cavité est assez large, ce qui prouve l'existence d'une solution analytique pour le problème étudie. On constate que la structure de l'écoulement et la distribution de la température au sein de la cavité, ainsi que sur les profils de

vitesse et de température diminuent avec l'augmentation du nombre de Bingham.

REFERENCES

[1] D. Vola, L. Boscardin, and J.C. Latche, "Laminar unsteady flows of Bingham fluids: a numerical strategy and some benchmark results," J. Comput. Phys, 187, 441–456, 2003.

[2] O. Turan, N. Chakraborty, and R.J. Poole, "Laminar natural convection of Bingham fluids in a square enclosure with differentially heated side walls," J. Non-Newtonian Fluid Mech, 165, 903–913, 2010.

[3] O. Turan, A. Sachdeva, R.J. Poole, and N. Chakraborty, "Laminar natural convection of Bingham fluids in a square enclosure with vertical walls subjected to constant heat flux," Numer. Heat Trans, A 60, 381–409, 2011.

[4] T. C. Papanastasiou, "Flows of materials with yield," J. of Rheology, 31, 385-404, 1987.

[5] E. Mitsoulis and Th. Zisis, "Flow of Bingham plastics in a lid-driven square cavity," J. of Non- Newtonian Fluid Mechanics, 101, 173-180, 2001.

[6] T. Min, H. G. Choi, J. Y. Yoo, and H. Choi, "Laminar convective heat transfer of a Bingham plastic in a circular pipe-II. Numerical approach-hydrodynamically developing flow and simultaneously developing flow," Int. J. Heat and Mass Transfer, 40, 3689-3701, 1997.

[7] E. Mitsoulis, "On creeping drag flow of a viscoplastic fluid past a circular cylinder: wall effects," Chem. Eng. Sci, 59, 789-800, 2004.

[8] S. Patankar, Numerical heat transfer and fluid flow. Hemisphere, Washington DC, 1980.

A review of materials used in extrusion based 3D Concrete Printing

Asmae NAJM-EDDINE, Ismail ARROUB, Ahmed BAHLAOU, Mohamed ABOUELMAJD, Ismail CHIGUER, Youssef NAJM-EDDINE, Soufiane BELHOUIDEG*

Research Laboratory in Physics and Sciences for Engineers (LRPSI), Polydisciplinary Faculty, Béni-Mellal, Morocco;

**Email: najmeddine.asmae.fpb21@usms.ac.ma*

Abstract— The reported number of harmful site accidents, large time, and material wastages in present labor-intensive techniques have relentlessly spurred the need to automate the civil infrastructure development process. Additive 3D printing could promote the construction field. Concrete 3D printing has a number of advantages, including shape freedom, a safer working environment, and decreased construction time and cost. The most difficult aspect of concrete printing is determining and maintaining the parameters of the mixture that are acceptable for printing and layering. In this study, different fresh cementitious mixtures proposed by different researchers were analyzed based on the rheological and the hardened requirements of printable concrete.

Keywords—: 3D Concrete printing, yield stress, raw materials, strength.

I. INTRODUCTION

Additive manufacturing is defined as ''a process of joining materials to make objects from 3D model data, usually layer upon layer, as opposed to subtractive manufacturing methodologies [1]. In other words, components are developed as volumetric objects utilizing 3D modeling software in the 3D concrete printing process. They're divided into two-dimensional layers. The data is sent layer by layer to a printing machine, which uses regulated extrusion of materials to create structural elements. The most known additive manufacturing processes are D-shape, contour printing, and concrete crafting, and they are shown in Fig. 1.

D-Shape is a particle-bed 3D printing adopted by Enrico Dini. The printing process consists of two repetitive work steps: the first is the application of a layer of dry particles and the second is the selective deposition of a fluid phase onto the particle packing by means of a print head or nozzle in order to bind the particles. Finally, the non-bonded particles are removed in a de-powdering process [2].

Concrete Printing, like CC, is an extrusion-based technology for layered cementitious structures. Although this print technique has the ability to integrate functional holes into the structure Concrete Printing, the technology has been developed to maintain 3-dimensional flexibility and a finer deposition resolution [3].

Khoshnevis created contour crafting (CC), an extrusion-based technique that integrates material supply and installation into a single system. CC is a layered construction technique that automates the creation of a whole house's structural elements, such as reinforcement, electricity, plumbing, drywall, and insulation [3].

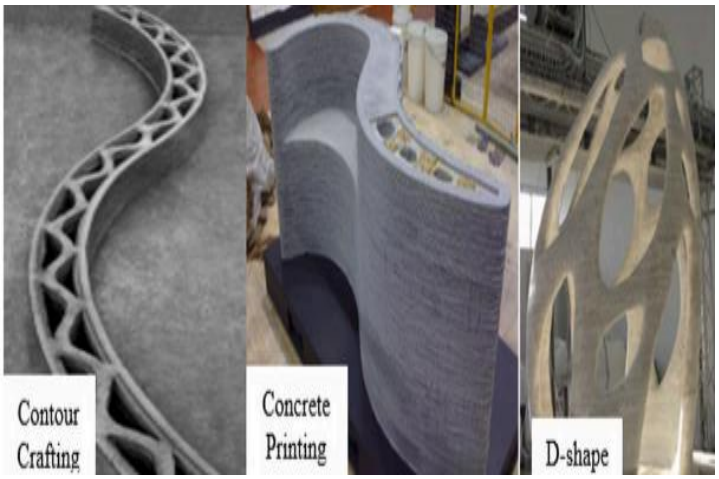


Fig. 1: Large-scale components made with different 3D concrete printing processes [4].

II. MATERIAL REQUIREMENTS OF PRINTABLE MIXTURES

A. Fresh state requirements

The new technologies pose specific requirements on 3D printable concrete, especially on its rheological properties in the fresh state. The selection of materials and the mix proportions to meet the pumpability, extrudability and buildability, which are key process-related material characteristics required for a successful design, is a fundamental issue in 3D concrete printing (3DCP) [6].

- Pumpability: The ease with which material is delivered through the pumping system;
- Extrudability: The simplicity and consistency with which material may be deposited throw the nozzle;

- Shape retention: The aptitude of the deposited concrete layer to keep its cross-section identical to that of the nozzle.
- Buildability: a property of a cementitious mixture to be printed in layers up to a certain height without considerable displacement or crash;
- Open time: The time after combining water, cement, aggregate, and other additives when concrete may be extruded through the printing head.

III. COMPOSITIONS AND THEIR IMPACT ON PROPERTIES OF 3D PRINTABLE CONCRETE:

Compressive, and flexural strengths requirements of 3DPC are similar to conventional concrete. Researchers were able to create printed concretes that were of minimal strength., e.g. 8 MPa [7], to very high strengths, e.g., 100 MPa [8].

The materials used in the extrusion based 3D concrete printing process are chosen to meet the high requirements of fresh and hardened properties of 3DPC structures. Additive manufactured concrete uses high content of cement than ordinary concrete which lead to high cost and high shrinkage so different supplement cementitious materials (SCMs), admixtures, fibers aggregates are applied. SCMs like fly ash, silica fume, limestone filler, and blast furnace slag, are used to partially replace cement [9]. In order to offer extrusion and printability criteria, materials used in 3DPC mixes should exhibit four essential features. These include; reducing autogenous and drying shrinkage, reducing the heat of hydration, decreasing the setting time value, and increasing early age strength [10]. superplasticizers and other chemical additives are employed to satisfy the requirements. In this study the proportions of different materials used in different researches and their impact on printed concrete behavior are displayed.

- Jo et al (2020) [11]

To discover the best mix ratios printable cementitious materials, three different types of mix designs were employed for functional testing: (1) cement and water, (2) cement, water, and sand, and (3) cement, water, sand, and PVA fibers.

During the first experiment different proportions Water/Cement were tested, the mix design with Water/Cement ratios of 0.32 and 0.30, respectively, were continuously dispensed and successfully layered the materials. The printed wall using the materials was stable despite some defects caused by shrinkage cracking. Sand (0.7mm and 0.8mm) was added to the paste during the second experiment to enhance the workability. Since shrinkage crack occurs in the previous experiments PVA fibers were added to the paste. See table1.

Table1: Optimum mixture and results found [11].

Water/Cement ratio	0.30 à 0.32
Sand/ binder	0.4 (0.7 mm)
PVA fibers	0.1%
Compressive strength	60.4 à 62 MPa

- Chen et al (2021) [12]

Y. Chen et al uses ordinary portland cement (OPC 42.5), Local river sand (maximum diameter 4.75 mm) as the fine aggregate, the basalt stones (size 5–10 mm''case1'') and 10–20 mm''case2'') coarse aggregate, and chemical admixtures: Polycarboxylate-based superplasticizer as water reducing agent (WRA), and hydroxypropyl methylcellulose ether as thickening agent (TA). Different mixtures were tested, the optimum one that was extrudable and buildable and present the highest strength is shown in table 2. The materials proportions are in volume ratio.

Table2: Mix designs by Y.Chen et al. [12]

Water	0.16
Cement	0.21
Sand	0.27
Basalt/ binder	0.241 (case1) 0.162 (case2)
WRA	0.5%
TA	0.05%
Compressive strength	64.6 MPa
Flexural strength	6.7 MPa

In this study, the compressive and the flexural strength increase with the decrease of cement to aggregate ratio, but the specimens show more direction-dependent behavior with the rise of aggregates in the paste.

- Liu et al (2019) [13]

In his research, Z. Liu et al utilizes ordinary portland cement (OPC type I 42.5), fly ash (diameter less than 10 µm), silica fume (class F, diameter less than 1 um), sand (sieved sand, less than 2.5 mm), and water, he tested 20 mixtures by varying proportion of each component, and measuring fresh parameters: static yield stress and dynamic yield stress. The optimum design in volume fraction is summarized in table 3.

Table3: optimum mixture by Z. Liu et al. [13]

Cement	0.15
Sand	0.22
Fly ash	0.26
Silica fume	0.04
Water	0.33
Superplasticizer	6g
Static yield stress	4880 Pa
Dynamic yield stress	201 Pa

- Hambach et al (2019) [14]

For his mix, M. Hambach et al utilize type I 52.5 R Portland cement, silica fume, water reducing agent, hydration inhibitor, three types of short reinforcing fibers (carbon, glass, and basalt fibers, 3-6-6 mm length). The optimum mix design is shown in table 4.

Table 4: material proportions by volume fraction [14].

Cement	0.62
Silica fume	0.21
Water	0.15
WRA	0.025%
Hydration inhibitor	0.03%
Carbon fibers	1%
Compressive strength	63 MPa
Flexural strength	30 MPa

- Zhang et al (2019) [15]

The cement paste in this study was made of Portland cement (P. II. 42.5 R) and one of three types of rheological modifiers – silica fume (diameter less than 3µm), which were employed to perform the buildability of 3DPC. The 3DPC components were separated into two categories: cement paste and aggregate. It was determined whether there was a link between paste flowability and aggregate (0/4,75mm, and 0/1,18mm fine sand) content in 3DPC combinations. One of the mixtures proposed in the study is shown in table5, the proportions are calculated to binder ratio except for the superplasticizer which is given by volume percentage.

Table 5: Optimized material adopted. [15]

Cement	0.32
Sand (1.18-4.75mm)	0.55
Silica fume	0.03
Water	0.25
superplasticizer	0.8%
Static yield stress	4250 Pa
Dynamic yield stress	250 Pa

- Chen et al (2021) [16]

In this study, calcium sulphoaluminate (CSA) cement (type R 425), bentonite, ultra-fine quartz sand, Polypropylène (PP) and polyvinyl alcohol (PVA) fibers (6mm), and admixtures: water reducing agent (WRA) hydroxypropyl methyl cellulose, and tartaric acid (TA). All raw materials are mass percentage of the CSA cement, except the fiber content (volume percentage). The mix that presents higher strengths is shown in table 6.

Table 6: mix proportions by M.Chen et al [16].

CSA cement	100
HPMC	0.4
Bentonite	1
UQS	1
WRA	0.3
TA	0.2
PVA fibers	1
Water	35
Compressive strength	70 MPa
Flexural strength	12.72 MPa

- Nerella et al (2019) [17]

Table 7 illustrates the composition of mortar adopted in this study, the contents are quantified by volume percentage, the amount of water is given in water to cement ratio.

Table 7: mixture proportion by V. N. Nerella et al [17].

Cement type (CEM I 52.5 R)	19
Fly ash	8
Micro Silica	8
Sand (0-2mm)	56
Water	23
Superplasticizer	0.5
Compressive strength	81 MPa
Flexural strength	5.9 MPa

- Huang et al (2022) [18]

Materials used as binders include Type III Portland cement, Class C fly ash (FAC), and silica fume (SF), Three types of sands are used: river sand (RS), masonry sand (MS), and lightweight sand (LWS). At 2% by volume, micro steel fibers (diameter: 0.2 mm, length: 13 mm) were used. The air-detraining admixture (ADA) was utilized in conjunction with the polycarboxylate-based high-range water-reducer (HRWR). The optimum mix design with higher compressive and flexural strength is in table 8.

Table 8: Mixtures proportions by volume proportions [18].

Cement	5 %
Water	20 %
Fly ash	37 %
SF	4 %

MS	27 %
RS	64 - 32 %
LWS	0 – 22 %
Steel fibers	2 %
HRWR	0.40 %
ADA	0.40 %
Compressive strength	125 - 144 MPa
Flexural strength	18.5 - 21.25 MPa

- Le et al (2012) [19]

In this research Le et al uses Portland cement type I 52,5, sand (diameter less than 1,6mm), silica fume (diameter less than 0,1mm), fly ash (diameter less than 0,6mm), superplasticizer, retarder, and 12/0.18 mm length/diameter polypropylene fibres. The raw materials proportions are in table9. The values are content to binder ratios except for superplasticizer and retarder which are total mass ratio.

Table 9: printable mix proportions [19].

cement	0.7
Sand	1.5
Silica fume	0.1
Fly ash	0.2
Water	0.26
superplasticizer	1%
Retarder	0.50%
PP fibers	1.2 kg/m3
Compressive strength	110 MPa

- Yang et al (2020) [20]

This article studies mixes of: P.O. 52.5 cement, silica fume (diameter less than 1mm), and fly ash (diameter less than 2mm), medium sand (Particle size: 0.212 mm–0.355 mm), and fine sand (Particle size: 0.106 mm–0.180 mm), NCC (nano calcium carbonate), cellulose (WRA), PS (polycarboxylate superplasticizer), and steel fibre (diameter 0,12mm, length 6-10mm). see table 10. All materials are calculated to binder ratio, except for the steel fibre, which is in volume fraction.

Table 10: raw materials and results [20].

cement	0.69
water	0.14
SF	0.15
FA	0.15
MS	0.33
FS	0.5
NCC	0.02
WRA	0.001
PS	0.009
Steel fibre	1%
Compressive strength	151 MPa
Flexural strength	45 MPa

Some authors study the buildability of 3D concrete printing from the yielding behaviors and thixotropic performance of cement composites. And others focus on the hardened state properties (compressive and flexural strength. Different mineral and chemical additives are used to enhance the properties of printable concrete. From the former researches [11-20], the main raw materials used in extrusion based 3d concrete are summarized in table11.

Table 11: average mass percentage of supplement cementitious materials and other additives and their utility.

Material	Mass proportion	Characteristics
Cement	11 to 70 %	Ordinary Portland cement is largely used in 3D concrete printing for its high early strength, but without additives, structures are exposed to shrinkage.
Water	15 to 35 %	The high water to cement ratio induces lower static yield stress (buildability) and lower dynamic yield stress (flowability).
Fly ash	8 to 37 %	The use of fly ash increases the workability of fresh mixtures. But it increases the reinforcement corrosion due to its high content of chlorides [21].
Silica fume	4 to 21 %	Increased silica fume concentration generates greater static yield stress, and reduces the workability of the paste. It's a corrosion catalyst [21].
Fibers	0.1 to 2 %	The insertion of fibers to concrete binds fractures and slows their spread. But high amount of fibers is inconvenient to workability and extrudability. different types of fibers are used including steel fibers (seems to be the most performant) [24], carbon fibers, PE and PVA fibers.
Aggregates	22 to 64 %	The use of aggregates decreases the shrinkage and increase the buildability and the workability of the mix, but coarse aggregates lead to weak interface joints and clog the nozzle printer.
Superplasticizer	0.1 to 1 %	Superplasticizers decrease viscosity (better flowability) without adding more water.
Other additives	Less than 1%	Additives such as: attapulgite nano-clay are used to prevent segregation and increase viscosity, nano calcium carbonate (NCC) is added to enhance the compactness of

		concrete materials.

IV. Conclusion

To conclude, this study provides a review of the most recent studies on extrusion based 3D concrete printing from the perspective of concrete materials, mix designs and their effect on fresh concrete behavior and hardened strengths of printed specimens. To make a comparison between the different studies and mixes, a conventional geometrical design with specified layers' geometry, amount of layers, interlayer time, nozzle angle and distance, printing speed, test measurements, and printing machines should be normalized.

References

[1] ASTM F2792-12a (withdrawn), American Society of Testing and Materials, 2012.

[2] D. Lowke, D.Enrico, A.Perraut, D.Weger, C.gehlen, B.Dillenburger Particle-bed 3D printing in concrete construction – Possibilities and challenges, August 2018, Cement and Concrete Research.

[3] A. Kazemian, X. Yuan, E. Cochran, B. Khoshnevis, Cementitious materials for construction-scale 3D printing: laboratory testing of fresh printing mixture, 2017, Constr. Build. Mater, 145, 639–647.

[4] S. Lim, R.A. Buswell, T.T. Le, S.A. Austin, A.G. Gibb, T. Thorpe, Developments in construction-scale additive manufacturing processes, 2012, Autom. Constr, 21, 262–268.

[5] P. Wu, J. Wang, X. Wang, A critical review of the use of 3-D printing in the construction industry (2016), Autom. Constr, 68, 21–31.

[6] C. Zhang, V. N. Nerella, A. Krishna, S. Wang, Y. Zhang, V. Mechtcherine, N. Banthia, Mix design concepts for 3D printable concrete: A review, (2021), Cement and Concrete Composites, 122.

[7] V. Markin, V.N. Nerella, C. Schrofl, G. Guseynova, V. Mechtcherine, Material design and performance evaluation of foam concrete for digital fabrication, 2019, Materials, 12, 15.

[8] V.N. Nerella, S. Hempel, V. Mechtcherine, Effects of layer-interface properties on mechanical performance of concrete elements produced by extrusion-based 3D- printing, 2019, Construct. Build. Mater, 205, 586–601.

[9] S. Hou, Z. Duan, J. Xiao, J. Ye, A review of 3D printed concrete: Performance requirements, testing measurements and mix design, 2020, Construction and Building Materials, 121745.

[10] H. G. Şahin, A. Mardani-Aghabaglou, Assessment of materials, design parameters and some properties of 3D printing concrete mixtures; a state-of-the-art review, 2022, Construction and Building Materials 316, 125865.

[11] J. H. Jo, B. W. Jo, W. Cho, and J-H. Kim, Development of a 3D Printer for Concrete Structures: Laboratory Testing of Cementitious Materials, International Journal of Concrete, 2020, Structures and Materials.

[12] Y. Chen, Y. Zhang, B. Pang, Z. Liu, G. Liu, Extrusion-based 3D printing concrete with coarse aggregate: Printability and direction-dependent mechanical performance, 2021, Construction and Building Materials 296 123624.

[13] Z. Liu, M. Li, Y. Weng, T. N. Wong, M. J. Tan, Mixture Design Approach to optimize the rheological properties of the material used in 3D cementitious material printing, 2019, Construction and Building Materials,198, 245–255.

[14] M. Hambach, D. Volkmer, Properties of 3D-printed fiber-reinforced Portland cement paste, 2019, Cement and Concrete Composites 79, 62-70.

[15] C. Zhang, Z. Houa, C. Chena, Y. Zhang, V. Mechtcherine, Z. Suna, Design of 3D printable concrete based on the relationship between flowability of cement paste and optimum aggregate content, 2019, Cement and Concrete Composites, 104, 103406.

[16] M. Chen, L. Yang, Y. Zheng, L. Li, S. Wang, Y. Huang P. Zhao, L. Lu, X. Cheng, Rheological behaviors and structure build-up of 3D printed polypropylene and polyvinyl alcohol fiber-reinforced calcium sulphoaluminate cement composites, 2021, Journal of Materials Research and Technology, 10, 1402-1414.

[17] V. N. Nerella, V. Mechtcherine, Studying the Printability of Fresh Concrete for Formwork-Free Concrete Onsite 3D Printing Technology (CONPrint3D), 2019, 3D Concrete Printing Technology.

[18] H. Huang, L. Teng, X. Gao, K H. Khayat d, F. Wang, Z. Liu, Use of saturated lightweight sand to improve the mechanical and microstructural properties of UHPC with fiber alignment, 2022, Cement and Concrete Composites 129, 104513.

[19] T. T. Le, S. A. Austin, S. Lim, R. A. Buswell, A. G. F. Gibb, T. Thorpe, Mix design and fresh properties for high-performance printing concrete, 2012, Materials and Structures, 45, 1221–1232.

[20] Y. Yang, C. Wu, Z. Liu, H. Wang, Q. Ren, Mechanical anisotropy of ultra-high performance fibre-reinforced concrete for 3D printing, 2022, Cement and Concrete Composites 125, 104310.

[21] M. Angel, B. Zamora, D. M Bastidas, G. Santiago-Hurtado [3], J. M. Mendoza-Rangel, C. Gaona-Tiburcio, J. M Bastidas, F. Almeraya-Calderón, Effect of Silica Fume and Fly Ash Admixtures on the Corrosion Behavior of AISI 304 Embedded in Concrete Exposed in 3.5% NaCl Solution, 2019, Materials (Basel) Dec 3, 12(23), 4007.

Length and diameter effects on earth-to-air heat exchanger performance in in hot temperate climate

Youssef NAJM-EDDINE, Ismail ARROUB, Ahmed BAHLAOUI, Mohamed ABOUELMAJD, Ismail CHIGUER, Asmae NAJM-EDDINE, Soufiane BELHOUIDEG

Research Laboratory in Physics and Sciences for Engineers (LRPSI), Polydisciplinary Faculty, Béni Mellal, Morocco

youssef.najmeddine@gmail.com

ABSRACT – *In this article, we study the effects of the length on the performance of a horizontal earth-to-air heat exchanger (EHAE) in a hot temperate climate of BENI MELLAL for two diameters (40 & 50 mm). Our approach aims to determine the length from which the thermal comfort temperature is obtained for both cases of diameters. TRNSYS software is used to do dynamic simulations of performance of an earth-air heat exchanger (EAHE) during the summer. The optimal length for the two aforementioned diameters from which the thermal comfort temperature is obtained at the heat exchanger's outlet is determined by the current study. Indeed, on the hottest day of July (reserved for this study), the differences of about 17 °C is reached when the ambient temperature is 44.6 °C and the outlet temperatures (EAHE) are 27.85 °C and 27.76 °C for the diameters 40 and 50 mm respectively.*

Keywords: *Energy efficiency, earth-air heat exchanger, length and diameter, passive cooling, TRNSYS*

I. INTRODUCTION

With the world's growing population and economies, as well as concerns about energy security and pollution reduction, researchers around the world have been pondering new systems—basically, a transformation in the way energy is produced, delivered, and consumed, ushering in a sustainable energy future.

Space cooling and heating account for almost a third of total energy demand (International Energy Agency, 2013). Space cooling and heating account for almost a third of total energy demand (International Energy Agency, 2013). By 2050, it is predicted that the number of building air-conditioning systems (ACs) would have increased from 1.6 billion to 5.6 billion, nearly tripling the power demand for building ACs [1]. Using a traditional heating, ventilation, and air conditioning (HVAC) system has a number of drawbacks, including high power consumption, increased peak load, environmental impact (ozone depletion and global warming), and lowered indoor air quality [2]. Aside from that, it raises the average cost of electricity because more captive power plants are required to fulfill peak electricity demand. Because of all of these problems with traditional HVAC systems, there has been a lot of interest in researching energy-efficient and environmentally friendly building heating and cooling systems. As a result, passive heating/cooling systems may play a critical role in averting an energy catastrophe.

The Earth-air heat exchanger (EAHE) is one of the most important passive cooling and heating systems, which is a subterranean passive system produces heating and cooling effects in the winter and summer. It makes use of earth as a heat source or sink, as well as air as a heat transfer medium. Outdoor air is routed through the subterranean pipes of a conventional EAHE system, where it exchanges heat with the subsoil.

As a result, throughout the summer / winter, the ground temperature is lower / (higher) than the ambient air temperature (winter) at adequate depth. The ground's constant temperature can be used for cooling and heating purposes [3]. When outside air is drawn through the EAHE system, the pipes are inhumed

in the soil. In the winter, the air can be heated, and in the summer, it can be cooled. As a result, the EAHE can reduce the amount of energy used for cooling or heating, as well as improve thermal comfort in buildings [4-6].

Several studies have looked into the effect of pipe diameter and length on the performance of the EAHE system and discovered that these factors have an impact [7]. The difference in air temperature between the outlet and inlet parts of the pipe is improved by increasing the pipe length in the May June July August September EAHE system [8]. However, when the pipe diameter increased, it deteriorated.

The purpose of this paper is to find the ideal length for two diameters of 40 mm and 50 mm on the hottest day of the year.

II. METHODOLOGY

A. Location and weather data

Morocco, is a country situated in the north of Africa. The position of city Beni Mellal (32.36°N, -6.4°E) is indicated in the map as shown in Figure 1. The meteorological data used in this study were taken from a typical year weather file.

The atmospheric temperature is presented in figure 2, for the period between 1st June and 31th August 2020. The ambient temperature ranges between 11.1 and 44.55 degrees Celsius, with the lowest and maximum temperatures occurring on the 5th of June and the 19th of July, respectively. The temperate climate is characterized by a significant amplitude of temperature fluctuations.

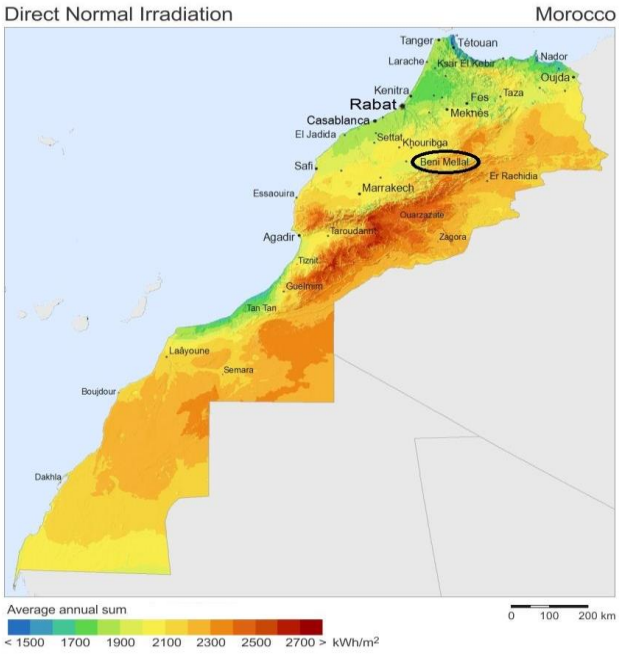


FIGURE 1. Variation of solar irradiation (direct normal irradiation) in Morocoo and location of Beni Mellal city.

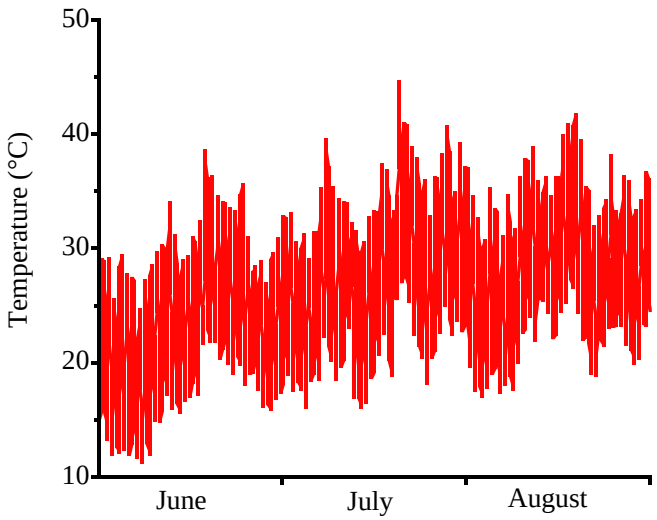


FIGURE 2. Atmospheric temperature (°C) for the period from 1st June to 31st August

B. Dynamic simulation models

The dynamic thermal simulation was carried out through TRNSYS which designates Transient System Simulation software to simulate the building. The systems are simulated using components called "types" that are interconnected through time-dependent inputs and outputs. The Earth-Air Heat Exchanger (EAHE) was simulated using Type 556 [9]. A time step of 1h was used to reduce computation time. A simplified scheme of EAHE is presented in the figure 3. The EAHE system routes outdoor air through subterranean pipes, where it exchanges heat with the subsoil and produces cooling effects in summer. The study is starts on 1st June and ends on 31th of August for the year of 2020.

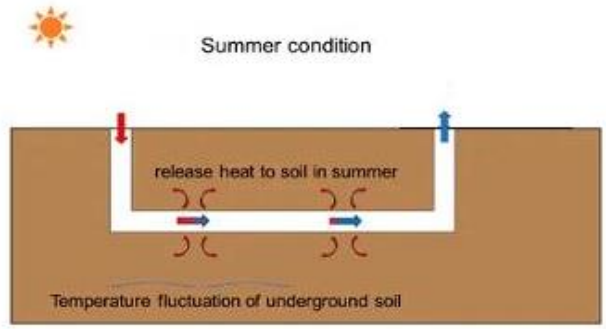


FIGURE 3. Simplified schematic of EAHE as a passive cooling system in summer

The EAHE is formed by only 1 pipe of PVC (polyvinyl chloride). The pipe is assumed horizontal, and inhumed at the mean depth of 3 m. The thermo-physical characteristics of the PVC pipe, soil and air at average ambient temperature (20 °C) are reported in Table 1.

TABLE 1. Physical properties [10].

	Thermal conductivity (W/m.K)	Specific heat (kJ/kg.K)	Density (kg/m^3)
Soil	1.4	1.3	1400
Air	0.025	1.01	1.16
PVC	0.17	1.3	1400

III. RESULTS AND DISCUSSION

In this part, the essential results obtained by means of the dynamic thermal simulations carried out in this study are plotted and discussed in detail. The results concern the variations of the temperature in the outlet of EAHE in continuous operation over the period between June and August for of the year 2020, as mentioned above.

Figure 4 represents a curve variation of the outlet temperature of the EAHE according to the pipe length for the two diameters 40 mm and 50 mm. Globally, it can be observed from this figure that the increase in the length of the pipe implies a decrease of the outlet temperature of the EAHE. And therefore a reduction in the range of variation of the EAHE outlet temperature, this is due to a thermal inertia of the soil and the increment in heat transfer area between the exchanger and the ground. The thermal comfort temperature is reached for a length of 40 m / 45m for a diameter of 40 mm / 50 mm. It is noted that the range of variation of thermal comfort temperature is [18 °C - 28 °C].

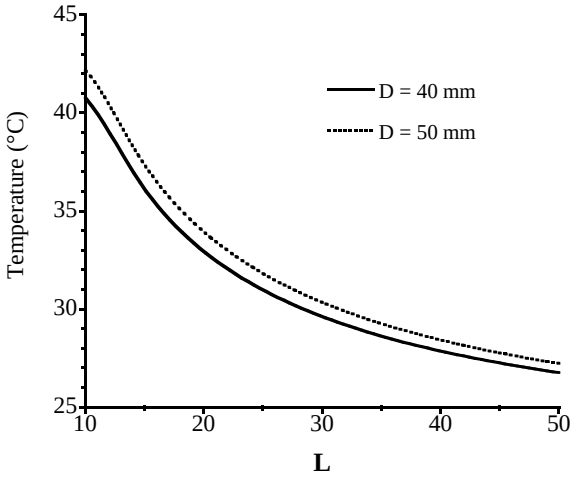
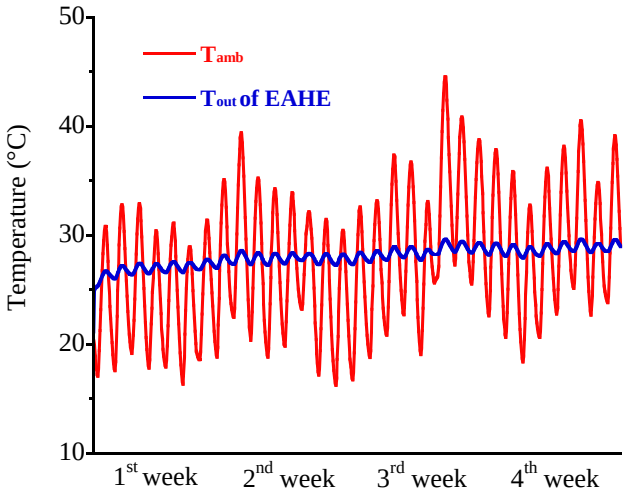


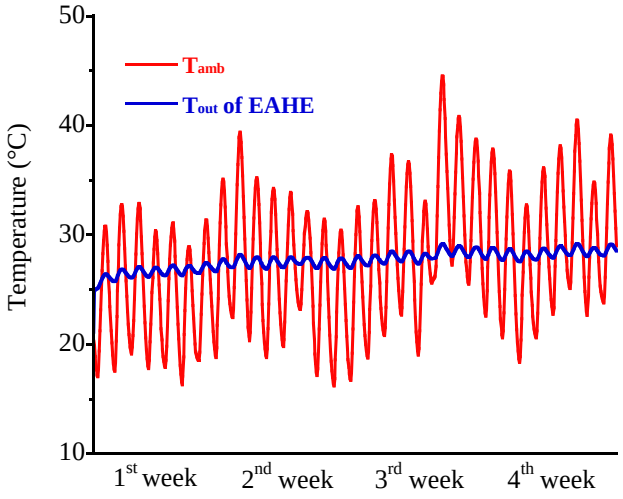
FIGURE 4. Variations, with length (L), of the outlet temperature, T_{out}, in the EAHE for two values of D

Figure 5 illustrates the temperature fluctuations, comparing the thermal behavior of the simulated EAHE (dynamically) and the atmospheric air temperature during the month of July for two cases (Case 1: D = 40 mm and $L_{exchanger}$ = 40m and Case 2: D = 50 mm and $L_{exchanger}$ = 45m).

In case 1, the graph shows that the ambient air temperature oscillates between 16 °C and 44.6 °C, while the air temperature in the outlet of the EAHE varies from 25°C to 27.85 °C. But in case 2, the air temperature in the outlet of the EAHE varies from 25°C to 27.76 °C. Results clearly demonstrate that the use of EAHE can significantly improve passive temperature stabilization and temperature fluctuations; with an increase / (decrease) of about 9 °C / (17 °C) compared to the minimum / (maximum) ambient temperature for both cases.



a) .Case 1



b) .Case 2

FIGURE 5. Temperature fluctuations of T_{amb} and $T_{out\ of\ EAHE}$ for the month of July in two cases, a) case 1 and b) case 2

As a complement, to assess the influence of the diameter on the outlet temperature, figure 6 represents the variation of the ambient temperature and the outlet temperature of EAHE during the hottest day of the month of July for both cases. The results show that the thermal comfort temperature is reached for case 1 (D = 40 mm) with the length of 40 m and for case 2 (D = 50 mm) of length equal to 45 m.

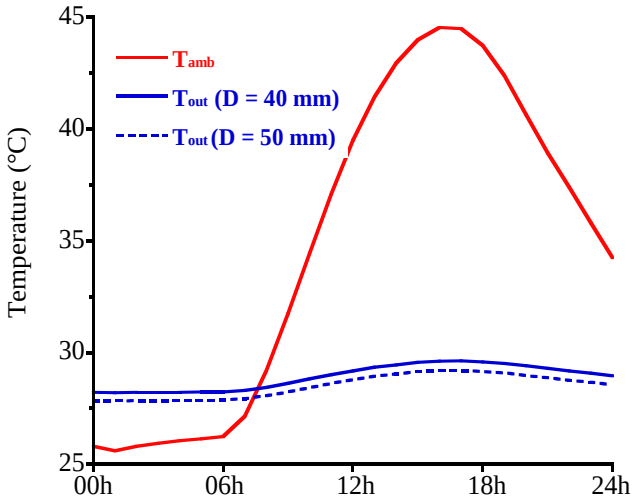


FIGURE 6. Variations, with time, of the temperature, T_{amb} , $T_{out\ (D = 40mm)}$ and $T_{out\ (D = 50mm)}$ in the hottest day of July

IV. CONCLUSION

In this study we conducted the simulation with the TRNSYS program to find the ideal length of the heat exchanger for the two diameters D = 40 mm and D = 50 mm in this study.

The results show that the EAHE with L= 50 m may reduce air temperature from 44.6 °C by up 26.76 °C and 27.23 °C for D = 40 mm and D = 50 mm respectively. The thermal comfort temperature is reached for a length of 40 m / 45m for a diameter of 40 mm / 50 mm. Moreover, the relatively soil can reduce the electrical costs for cooling the air in a temperate area as the Beni Mellal (Morocco). The results show that the EAHE can be an effective thermal system for air refreshing in buildings, as it maintains a near-constant air temperature of around 27 degrees Celsius. In the conclusion, we propose that the Moroccan thermal code encourages the employment of passive cooling strategies in temperate climates in order to increase the energy efficiency.

REFERENCES

[1] F. Birol, 2018. The Future of Cooling: Opportunities for energy efficient air conditioning,

[2] I. Attar, N. Naili, N. Khalifa, M. Hazami, A. Farhat, 2013. Parametric and numerical study of a solar system for heating a greenhouse equipped with a buried exchanger. Energy Convers. Manag. 70, 163–173.

[3] M. Khabbaz, B. Benhamou, K. Limam, P. Hollmuller, H. Hamdi and A. Bennouna, Experimental and numerical study of an earth-to-air heat exchanger for air cooling in a residential building in hot semi-arid climate, Energy and Buildings, 125, 109-121, (2016).

[4] S. Menhoudj, A. M. Mokhtari, M. H. Benzaama, C. Maalouf, M. Lachi and M. Makhlouf, Study of the energy performance of an earth-air heat exchanger for refreshing buildings in Algeria, Energy and Buildings, 158, 1602-1612, (2018).

[5] M. Grosso and G. Chiesa, Horizontal earth-to-air heat exchanger in Imola, Italy. A 30-Month- long monitoring campaign, Energy Procedia, 78, 73-78, (2015).

[6] K. K. Agrawal, R. Misra, G. D. Agrawal, M. Bhardwaj and D. K. Jamuwa, Effect of different design aspects of pipe for earth air tunnel heat exchanger system: a state of art, International Journal of Green Energy, 0, 1-17. (2019).

[7] K. K. Agrawal, G. D. Agrawal, R. Misra, M. Bhardwaj and D. K. Jamuwa, A review on effect of geometrical,

[8] S. A. Klein and W.A. Beckman, TRNSYS 18: A Transient System Simulation Program, Solar Energy Laboratory, University of Wisconsin: Madison, WI, USA, (2017).

[9] TESS library mathematical reference. In: TESS Libs 17: component Libraries for the TRNSYS simulation environment, vol. 4. Madison, Wisconsin: Thermal Energy System Specialists, LLC; (2013).

[10] I. Arroub, A. Bahlaoui, S. Belhouideg and A. Kaddour, Dynamic modelling of an earth-to-air heat exchanger for air cooling on the building in hot temperate climate of Beni Mellal Morocco, International Journal of Energetica (IJECA), 6 (1), (2021).

Comparison between the Analytical and Numerical Solutions of th Thermosolutal Mixed Convection in a Vertical Cavity

M. Rahmoun [a], M. Lamsaadi [a], T. Makayssi [b], S. Ziane [a], B. EL Hadoui [b], H. El Harfi [a]

[a] *Laboratory of Research in Physics and Engineering Sciences (LRPSI). Department of Physics, Polydisciplinary Faculty (FPBM), Sultan Moulay Slimane University (USMS), Beni-Mellal, Morocco*

[b] *Sultan Moulay Slimane University, Faculty of Sciences and Technologies, LGI, B.P. 523, Beni-Mellal, Morocco*

Gmail: rahmoun.mohamed.fpb21@usms.ac.ma; lamsaadima@yahoo.fr; tmkayssi@gmail.com; ziane.sihame.fpb21@usms.ac.ma; bilal.elhadoui@usms.ac.ma; elharfi.hassan@yahou.fr

***Abstract*—A comparison is made analytically and numerically on the thermosolutal mixed convection in a vertical rectangular cavity, filled with a Newtonian fluid of Prandtl number, Pr = 7, whose vertical walls are mobile and subjected to constant heat and mass fluxes. The horizontal walls are assumed to be impermeable and adiabatic. The mathematical model is based on the Navier-Stokes equations as well as the conservation of energy and concentration. The analytical resolution, which is based on the parallel flow approximation, has been developed in the case of long cavities (A>>1). Moreover, the equations governing the phenomenon have been discretized by the finite difference method. The numerical approach is based on the implicit method with alternating directions (A.D.I) and the successive over-relaxation method (P.S.O.R).**

Keywords: **Rectangular cavity, Thermosolutal mixed convection, Heat and mass transfer, Parallel flow approximation.**

I. INTRODUTION

The study of heat and mass transfer by thermosolutal mixed convection in rectangular cavities has been of primary interest. This interest is dictated by the role played in different industrial systems such as cooling of electronic components, thermal power plants and heat exchangers etc.

The phenomenon of mixed convection in fluids confined in cavities with mobile walls has been studied by several authors. H. F. Oztop and I. Dagtekin [1], studied two-dimensional steady-state mixed convection problem in a vertical two-sided square cavity differentially heated by a lid is numerically investigated. The left and right moving walls are kept at different constant temperatures while the top and bottom walls are thermally insulated. It has been performed for three different cases characterized by the direction of motion of the vertical walls. The governing parameter is the Richardson number, which characterizes the heat transfer regime in mixed convection. Similarly without considering the thermosolutal convection, Elharfi et al [2] investigated the lid-driven convection in a rectangular cavity and filled with nanofluid, copper-water type, and subjected to a uniform heat flux along the vertical side walls, and the horizontal walls are adiabatic and impermeable. They studied analytically based on parallel flow approximation and numerically by finite volume method. The parameters of the study are the cavity aspect ratio, A = 8, the Prandtl number, Pr = 7 , the Reynolds number, Re, characterizing the forced convection, and the Richardson number, Ri, measuring the relative importance of the natural and forced convections, in addition to the solid volume fraction (Φ), translating the nanoparticle concentration.

Another numerical study of the double diffusive mixed convection in a rectangular cavity has been done by M. A. Teamah, and Wael M. El-Maghlany [3] but this time in a rectangular cavity with movable upper lid. Both upper and lower walls are adiabatic and impermeable. Different temperatures and constant concentrations are imposed along the vertical cavity walls, the steady laminar regime is considered. Numerical results are presented for the effect of the Richardson number, the Lewis number and the buoyancy ratio on the streamline, temperature and concentration contours. In addition, the predicted results for local and average Nusselt and Sherwood numbers are presented and discussed for various parametric conditions. This study was performed for $0.1 \leq Le \leq 50$. Throughout the study, the Grashof number and the aspect ratio, are held constant at 10^4 and 2 respectively and $-10 \leq N \leq 10$, while the Richardson number was varied from 0.01 to 10 to simulate flow dominated by forced convection, mixed convection and natural convection. A. Louaraychi, et al. [4] studied analytically and numerically the mixed convection in a horizontal rectangular cavity filled with Newtonian fluid, where the horizontal walls are adiabatic, with the upper moving one, and a uniform heat flux is applied on the vertical walls. Simulations were performed for the Reynolds and Richardson numbers in the ranges $0.1 \leq Re \leq 10$ and $0.1 \leq Ri \leq 10^5$, respectively. The obtained results show that the increase of Ri and Re serves for a growth of the average Nusselt number and an increase to the heat transfer inside the cavity. Similarly, the analytical and numerical study of mixed convection in horizontal rectangular cavities with single and double lids, filled with a Newtonian fluid, and subjected to a uniform heat flow along their vertical short sides, and the two horizontal walls are isolated and move (mobile) in opposite directions, was

carried out by Louaraychi et al. [5]. The finite volume method with SIMPLER algorithm is used to solve the complete governing equations for which the Boussinesq approximation is adopted while the analytical approach is based on the parallel flow assumption. They examined the effects of the governing parameters on the flow and heat transfer characteristics are analyzed for both types of cavities. The zones characterizing the dominance of natural and forced convection as well as the moment when the two phenomena compete (mixed convection) are delineated. It is found that the transition from one dominated regime to another depends on the Ra_T/Pe^3 ratio. Similarly studied by Youssef Tizakast et al [6] but both thermosolutal mixed convection in horizontal rectangular cavity filled with Newtonian fluid, both numerically using finite volume method to solve the governing equations and analytically based on the parallel flow approximation developed for the case of shallow enclosures A≫1. Uniform heat and mass fluxes are applied to the short vertical walls, while the horizontal walls are insulated and impermeable, with the top wall sliding from left to right. The results show good agreement between the two solutions for a wide range of control parameters: Peclet number, Pe, Lewis number, Le, buoyancy ratio, N, and thermal Rayleigh number, Ra_T. In order to highlight how convection regimes influence the effect of control parameters on flow and heat and mass transfer characteristics, the Ra_T/Pe^3 parameter is established to delineate areas where natural, mixed, and forced convection dominate heat and mass transfer. The effects of the control parameters on the flow intensity and heat and mass transfer rates are illustrated and discussed in terms of the flow function, Ψ, mean Nusselt number, $\overline{Nu}$, mean Sherwood number, $\overline{Sh}$, for the three separate convective regimes.

II. MATHEMATICAL FORMULATION

The physical problem considered is schematically shown in figure (1). It is a rectangular cavity of length L′ and height H′, filled with a Newtonian fluid, the vertical walls are movable, transversely heated and subjected to constant flows of heat (q') and mass(J'), the other walls are assumed impermeable and adiabatic and are immovable. The axes of coordinates (x) and (y) are respectively oriented along the vertical and horizontal directions.

$$\left.\begin{array}{ll} (x,y)=\frac{(x',y')}{L'}\ , & (u,v)=\frac{(u',v')L'}{\alpha} \\ t=\frac{t'\alpha}{L'^2}\ , & p=\frac{p'}{\rho_0\left(\frac{\alpha}{L'}\right)^2} \\ T=\frac{(T'-T_0')}{\Delta T^*}\ , & C=\frac{(C'-C_0')}{\Delta C^*} \\ \Delta T^*=\frac{q'L'}{\lambda}\ , & \Delta C^*=\frac{J'L'}{D} \end{array}\right\} \quad (1)$$

The dimensionless variables are used:
Where ΔT^* *and* ΔC^* are the characteristic differences in temperature and concentration, respectively.

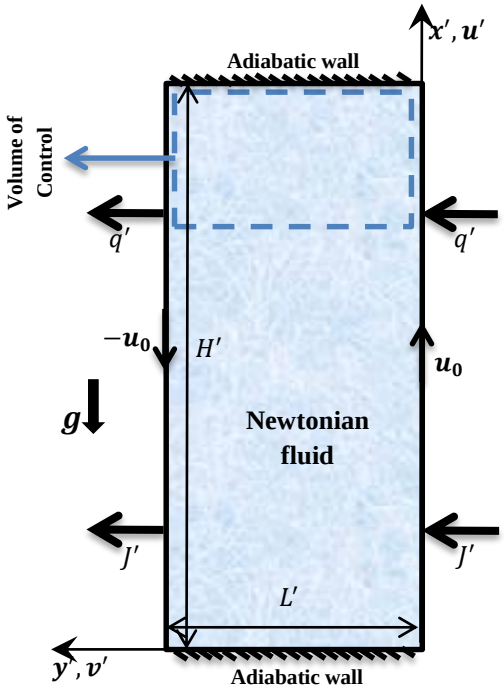


Figure 1: Schematic representation of the physical problem considered.

The dimensionless governing equations are written as follows :

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0 \quad (2)$$

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y} = -\frac{\partial p}{\partial x}+Pr\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right) + PrRa_T(T+NC) \quad (3)$$

$$\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}=-\frac{\partial p}{\partial y}+Pr\left(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}\right) \quad (4)$$

$$\frac{\partial T}{\partial t}+u\frac{\partial T}{\partial x}+v\frac{\partial T}{\partial y}=\left(\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}\right) \quad (5)$$

$$\frac{\partial C}{\partial t}+u\frac{\partial C}{\partial x}+v\frac{\partial C}{\partial y}=\frac{1}{L_e}\left(\frac{\partial^2 C}{\partial x^2}+\frac{\partial^2 C}{\partial y^2}\right) \quad (6)$$

The dimensionless boundary conditions for the physical system considered are as the following:

$$\left\{\begin{array}{c} pour\ \ x=0\ \ et\ \ x=A\ ,\forall y\ \ \Rightarrow \\ \mathrm{u=v=0\ et}\ \frac{\partial T}{\partial x}=0\ et\ \frac{\partial C}{\partial x}=0 \\ pour\ \ \ y=0\ ,\forall x\ \ \Rightarrow \\ u-Pe=v=0\ \ et\ \ \frac{\partial T}{\partial y}=-1\ et\ \frac{\partial C}{\partial y}=-1 \\ pour\ \ y=1\ ,\forall x\ \ \Rightarrow \\ u+Pe=v=0\ \ et\ \ \frac{\partial T}{\partial y}=-1\ et\ \frac{\partial C}{\partial y}=-1 \end{array}\right. \quad (7)$$

During our study, dimensionless numbers are used to characterize the flow and transfers are defined as:

$$A = \frac{H'}{L'} \ , \ Pe = \frac{u'L'}{\alpha} \ , \ Pr = \frac{\nu}{\alpha} \ , Le = \frac{\alpha}{D}$$
$$Ra_T = \frac{g\beta_T q' {L'}^4}{\nu\alpha\lambda} \ , \ N = \frac{\beta_C \Delta C^*}{\beta_T \Delta T^*} \qquad (8)$$

Heat and mass transfer

We can quantify the rate of heat and mass transfer transported within the fluid by mixed thermosolutal convection through the active walls of the vertical cavity, by dimensionless numbers that are respectively the Nusselt number "Nu" and Sherwood number "Sh" defined by the following relations:
For a given position of x we find:

$$Nu(x) = \frac{q'L'}{\lambda \Delta T'} = \frac{1}{\Delta T(x)} \qquad (9)$$

$$Sh(x) = \frac{J'L'}{D\Delta C'} = \frac{1}{\Delta C(x)} \qquad (10)$$

Where $\Delta T(x) = T(x,0) - T(x,1)$ and $\Delta C(x) = C(x,0) - C(x,1)$ represent the local dimensionless temperature and concentration difference between the two horizontal walls of equations y= 0 and y= 1.

III. NUMERICAL APPROACH

In the present study we will use the finite difference method as a discretization method. The latter consists in replacing the derivatives appearing in the equation system governing our studied phenomenon by divided differences or combinations of the point values of the function in a finite number of discrete points or nodes of the mesh. This method was introduced by Euler. The resolution of the system of governing equations using the ADI method (Altenating Direction Implicit). Within our system of elements, using Taylor series, we can determine the fields of all the variables of the considered problem.

The steady state is reached when the following criterion is verified for all the independent variables of the problem:

$$\frac{\left| f^{iter+1}{}_{i,j} - f^{iter}{}_{i,j} \right|}{\left| f^{iter+1}{}_{i,j} \right|} < \varepsilon_P \qquad (11)$$

Where f represents the values of the four fields of the system (Ω , ψ , T *and* C) taken at each point in the domain at time $n\Delta$t. ε_P is a constant (between 10^{-3} and 10^{-4}).

IV. ANALYTICAL APPROACH

The analytical solution is only possible in certain limit cases for which the equations are considerably simplified. Thus, for the case of elongated cavities (A >> 1) with active walls exposed to constant heat and mass fluxes, an approximate analytical solution is possible, using the concept of parallel flow.

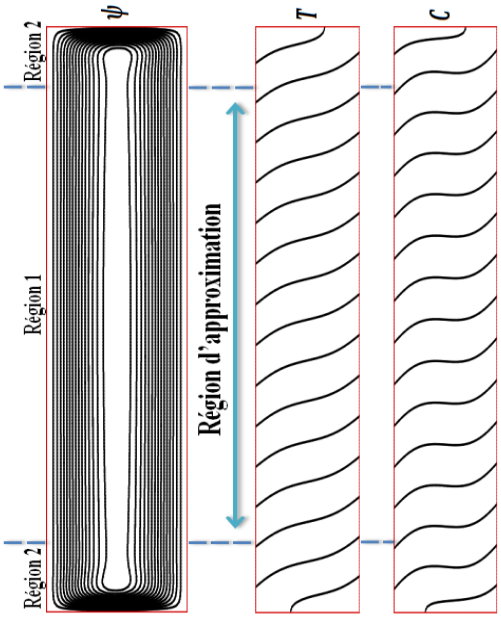


Figure 2: Approximation of the parallel flow

$$u(x,y) = u(y) \quad ; \quad v(x,y) = 0$$
$$\psi(x,y) = \psi(y)$$
$$\begin{cases} T(x,y) = C_T(x - A/2) + \theta_T(y) & (a) \\ C(x,y) = C_C(x - A/2) + \theta_C(y) & (b) \end{cases} \qquad (12)$$

Where : C_T and C_C are respectively the temperature and concentration gradients along the x-axis.

Approximated governing equations

As we have seen previously, the current function depends only on y, the temperature is a function of x and y. By substituting the approximations (12) in the basic equations, (2), (3), (4), (5), and (6), we obtain the following ordinary differential equation:

$$\frac{d^3u}{dy^3} + Ra_T\left(\frac{d\theta_T}{dy} + N\frac{d\theta_C}{dy}\right) = 0 \qquad (13)$$

$$\frac{d\theta_T}{dy} = C_T\psi - 1 \qquad (14)$$

$$\frac{d\theta_C}{dy} = LeC_C\psi - 1 \qquad (15)$$

By inserting equations (14) and (15) into equation (13), we obtain the following equation:

$$\frac{d^4\psi}{dy^4} + Ra_T((C_T\psi - 1) + N(LeC_C\psi - 1)) = 0 \qquad (16)$$

To these can be added the return flow condition (conservation of matter):

$$\int_0^1 u(y)dy = 0 \ , \ \int_0^1 \theta_T(y)d = 0$$
$$et \ \int_0^1 \theta_C(y)dy = 0 \qquad (17)$$

The constants C_T and C_C represent the temperature and concentration gradients, respectively. After the calculation we find the expressions of C_T and C_C in the following form:

$$C_T = \frac{\int_0^1 \psi(y)\,dy}{1+\int_0^1 \psi^2(y)\,dy}$$
$$C_C = \frac{Le\int_0^1 \psi(y)\,dy}{1+Le^2\int_0^1 \psi^2(y)\,dy} \quad (18)$$

The general form of the differential equation (16) can be put in the form:

$$\frac{d^4\psi}{dy^4} + K\,\psi = S \quad (19)$$

With $K = Ra_T(C_T + NLeC_C)$ et $S = Ra_T(1+N)$

From a physical point of view, the constant K is in general positive when the volume forces are cooperating, N > 0

The form of the solution of equation (19) is given by:

$$\psi(y) = \left[(\mathrm{A}-\mathrm{B}(1+\gamma))\cosh\left(\delta\left(y-\tfrac{1}{2}\right)\right)\cos\left(\delta\left(y-\tfrac{1}{2}\right)\right) - (\mathrm{C}+\mathrm{B}(1-\gamma))\sinh\left(\delta\left(y-\tfrac{1}{2}\right)\right)\sin\left(\delta\left(y-\tfrac{1}{2}\right)\right)\right] + \mathrm{G} \quad (20)$$

$$\theta_T(y) = \frac{C_T}{2\delta}\left[(A-C-2B)\cosh\left(\delta\left(y-\tfrac{1}{2}\right)\right)\sin\left(\delta\left(y-\tfrac{1}{2}\right)\right) + (A+C-2B\gamma)\sinh\left(\delta\left(y-\tfrac{1}{2}\right)\right)\cos\left(\delta\left(y-\tfrac{1}{2}\right)\right)\right] - (1-GC_T)\left(y-\tfrac{1}{2}\right) \quad (21)$$

$$\theta_C(y) = \frac{LeC_C}{2\delta}\left[(A-C-2B)\,cosh\left(\delta\left(y-\tfrac{1}{2}\right)\right)sin\left(\delta\left(y-\tfrac{1}{2}\right)\right) + (A+C-2B\gamma)sinh\left(\delta\left(y-\tfrac{1}{2}\right)\right)cos\left(\delta\left(y-\tfrac{1}{2}\right)\right)\right] - (1-GLeC_C)\left(y-\tfrac{1}{2}\right) \quad (22)$$

The constants B, γ, A, C and G are given by :

$$\begin{cases} \delta = \left(\dfrac{Ra_T(C_T+N\mathrm{L_e}C_C)}{4}\right)^{\frac{1}{4}} \quad ; \quad \delta_0 = \dfrac{\delta}{2} \\ \gamma = \tanh(\delta_0)\,cotan(\delta_0) \quad ; \quad G = \dfrac{Ra_T(1+N)}{4\delta^4} \\ A = \dfrac{Pe*tanh(\delta_0)}{\delta\left((1+\gamma)\cosh(\delta_0)\cos(\delta_0)+(1-\gamma)sinh(\delta_0)\sin(\delta_0)\right)} \\ B = \dfrac{G}{(1+\gamma)\cosh(\delta_0)\cos(\delta_0)+(1-\gamma)sinh(\delta_0)\sin(\delta_0)} \\ C = \dfrac{Pe*\mathrm{cotan}(\delta_0)}{\delta\left((1+\gamma)\cosh(\delta_0)\cos(\delta_0)+(1-\gamma)sinh(\delta_0)\sin(\delta_0)\right)} \end{cases}$$

We can also, from relation (20) obtain the following velocity distribution y :

$$u(y) = \delta\left[(\mathrm{A}-\mathrm{C}-2\mathrm{B})sinh\left(\delta\left(y-\tfrac{1}{2}\right)\right)\cos\left(\delta\left(y-\tfrac{1}{2}\right)\right) + (2\mathrm{B}\gamma-(\mathrm{A}+\mathrm{C}))\cosh\left(\delta\left(y-\tfrac{1}{2}\right)\right)\sin\left(\delta\left(y-\tfrac{1}{2}\right)\right)\right] \quad (23)$$

Finally obtaining these equations that allow us to calculate the temperature and concentration gradients, C_T and C_C :

$$\begin{cases} C_T = C_T I - (1-GC_T)(J-G) \\ C_C = L_e^2 C_C I - Le(1-LeGC_C)(J-G) \end{cases} \quad (24)$$

The coefficients *I* and *J* are given by the following relations :

$$I = \frac{1}{4\delta}\left[\left(\frac{A^2}{2}-\frac{C^2}{2}-AC-2B(A-C\gamma)-B^2(\gamma^2-2\gamma-1)\right)cosh(\delta)\,sin(\delta) + \left(\frac{A^2}{2}-\frac{C^2}{2}+AC-2B(A\gamma+C)+B^2(\gamma^2+2\gamma-1)\right)sinh(\delta)\,cos(\delta) + \delta(-A^2+C^2+2AB(1+\gamma)+2BC(1-\gamma)-4\gamma B^2)\right]$$

$$Et\ J = \left[\frac{1}{\delta}(2B+C-A)cosh(\delta_0)sin(\delta_0) + \frac{1}{\delta}(2B\gamma-A-C)sinh(\delta_0)cos(\delta_0) + A\,cosh(\delta_0)cos(\delta_0) - C\,sinh(\delta_0)sin(\delta_0)\right]$$

$$G = B\big((1+\gamma)cosh(\delta_0)cos(\delta_0) + (1-\gamma)sinh(\delta_0)sin(\delta_0)\big)$$

The numerical solution of the previous system (24) using, for example, the Newton-Raphson method, the temperature and concentration gradients, C_T and C_C, can be calculated. Once C_T and C_C are determined, the current function field can be calculated, as well as the temperature and concentration profiles.

The heat and mass transfer rates, expressed in Nusselt and Sherwood numbers, are obtained by substituting equations (9), (10), and (20), we get:

$$Nu = \frac{1}{(1-C_TG)+C_TJ-A\,cosh(\delta_0)cos(\delta_0)+C\,sinh(\delta_0)sin(\delta_0)}$$
$$Sh = \frac{1}{(1-GLeC_C)+LeC_CJ-A\,cosh(\delta_0)cos(\delta_0)+C\,sinh(\delta_0)sin(\delta_0)} \quad (25)$$

V. RESULTS AND DISCUSSIONS

V.1 Quantitative validation of the numerical code

The numerical program that we have developed to solve the basic equations has been validated by taking as reference the results of G. de Vahl Davis and I. P. Jones [7] , Makeyssi [8], and Y. Tizakast et al [6] (Table 1).

Table 1. Comparison of $|\Psi_{min}|, |\Psi_{max}|, Nu$ and Sh.

Ra	Vahl davis et al [7]		Present study	
	$\lvert\Psi_{min}\rvert$	Nu	$\lvert\Psi_{min}\rvert$	Nu
10^4	5.071	2.243	5.071	2.242
Ra	T. Makeyssi [8]		Present study	
	Nu	Sh	Nu	Sh
10^4	30.459	2850	30.901	2912.797
Pe	Y. Tizakast et al [6]		Présent étude	
	Nu		Nu	
5	33.343		33.314	

V.2. Effect of mesh size

The effect of mesh size on the results of the numerical code is presented in Table 2 for the case of a vertical cavity, for different mesh sizes by setting the following parameters:

$Ra = 10^3, Pe = 5, N = 1, Le = 2\ and\ A = 16.$ The tables below represent the evolution of the current function at the center of the cavity Ψ_C of the Nusselt Nu and Sherwood Sh numbers for different meshes.

Table 2. The effect of mesh size on numerical results, for $Ra = 10^3, Pe = 5, N = 1, Le = 2\ and\ A = 16.$

	Numerical results for different grids				
M*N	60x140	60x180	60x200	80x160	80x180
Ψ_C	2.064	2.056	2.060	2.093	2.105
Nu	2.151	2.152	2.150	2.138	2.131
Sh	3.409	3.439	3.439	3.376	3.366

These results show that a mesh size of 60×180 is sufficient to ensure a good solution of the studied problem and that the results converge to a constant solution. Increasing the mesh size to 80×180 the results will be almost stable and do not change for, Ψc with an acceptable error that does not exceed 2%. So, we chose the 60×180 mesh which provides a good compromise between the accuracy of the results and the computation time.

V.3 The aspect ratio effect

To estimate the minimum value of the aspect ratio A for which the flow is parallel in a vertical layer, numerical tests were performed. Typical results are reported in Figure 3, for different values of the aspect ratio A, where the numerical results are indicated by black symbols and the approximate analytical solution is indicated by lines, valid for the case of an infinite vertical cavity (A>>1). From this figure, we can estimate the minimum value of the aspect ratio A, above which the flow in the cavity can be considered parallel. Increasing the aspect ratio A for ($Ra = 10^4, Pe = 5$), leads to an asymptotic situation for which the solutions are independent of the aspect ratio A. Following several tests, we can therefore conclude that the numerical results can be considered independent of the cavity aspect ratio when A ≥ 12. For this reason the majority of the results reported in this work are obtained for a value of the aspect ratio A=16.

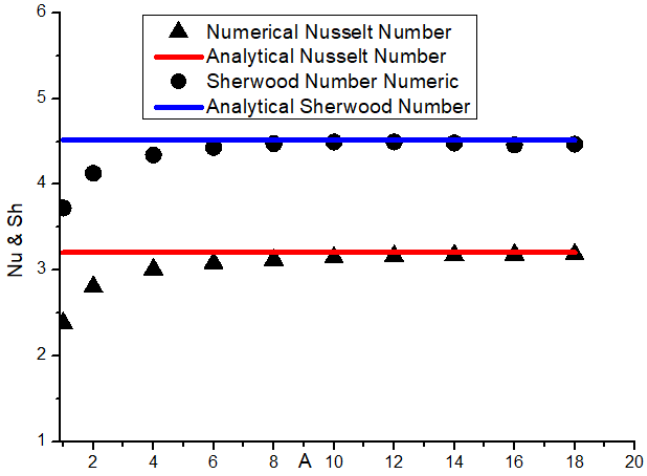


Figure 3: Influence of cavity aspect ratio on heat and mass transfer rates, for $Ra_T = 10^4, Le = 2, Pe = 5, N = 1\ and\ Pr = 7.$

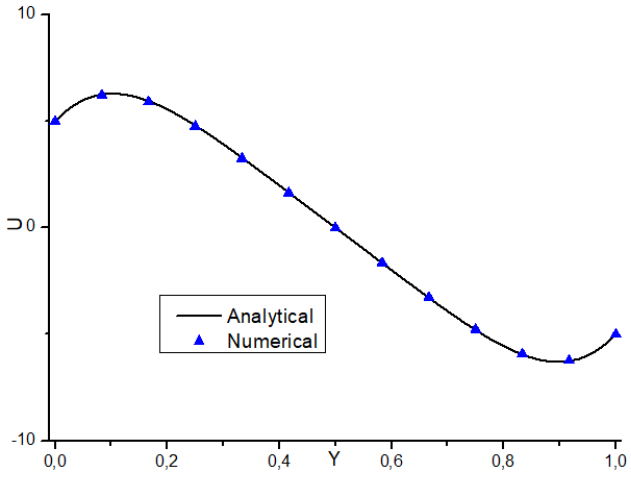
V.4 Comparison of analytical and numerical results.

In this section, typical results are shown in Figure 4, The profiles of velocity, current function, temperature and concentration in the plane (x=A/2) present a good agreement between numerical and analytical results.

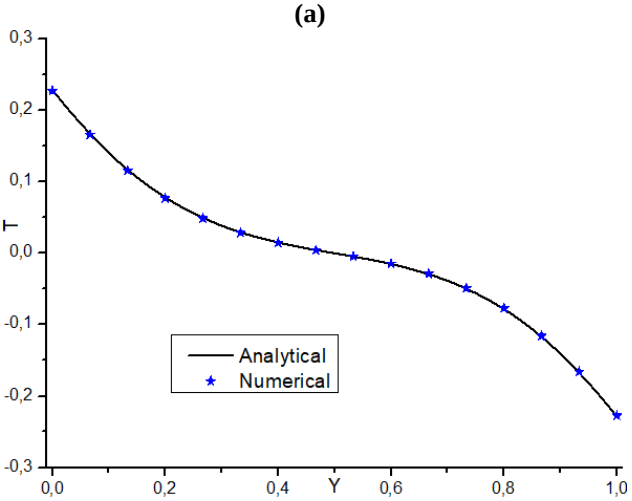
Table 3. The comparison between the analytical and numerical numerical results, for $Ra = 10^3, Pe = 5, N = 1, Le = 2\ et\ A = 16.$

Numerical	Analytical	Error
Ψc = 2.0528	Ψc = 2.0671	0.69%
Nu = 2.1544	Nu = 2.1840	1.37%
Sh = 3.4460	Sh = 3.4976	1.49%

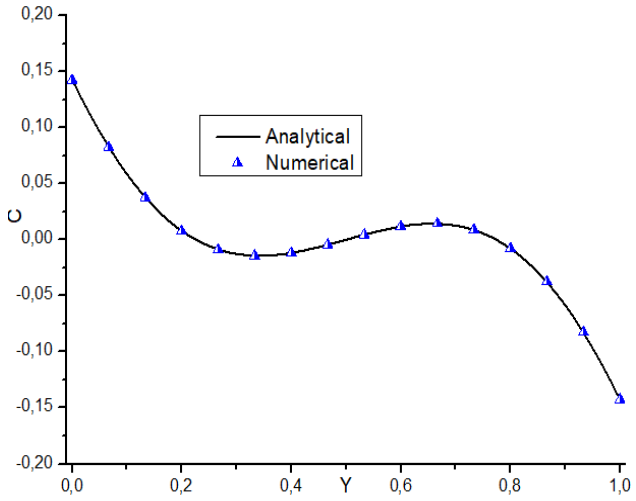
There is an acceptable error of no more than 2% in the Nusselt, Sherwood numbers and the current function at the center of the cavity Ψc.



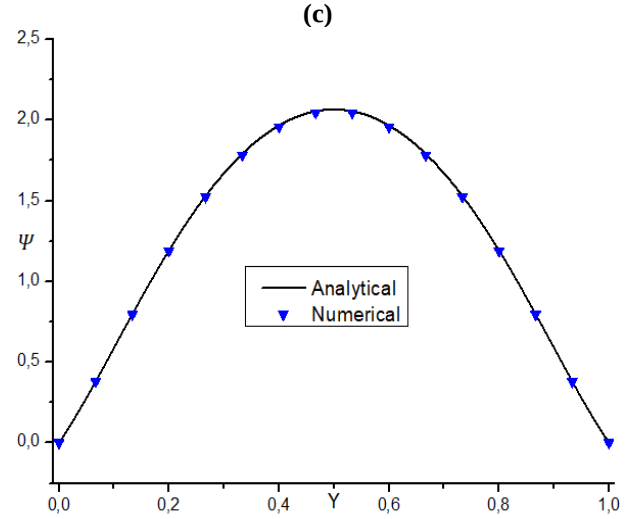
(a)



(b)



(c)



(d)

Figure 4 : Numerical and analytical results for $Ra_T = 10^3, Pe = 5, Le = 2$ *and* $N = 1$; distribution along the horizontal x=A/2 axis of (a) velocity u, (b) Temperature (c) concentration (d), and current function Ψ.

The profiles of velocity U, temperature T, concentration C, and current function Ψ, in the midplane of the cavity are shown in Fig. 4 (a, b, c, d) for N = 1, Pr = 7, A = 16, and Le = 2, respectively. This figure clearly shows that the numerical and analytical results are in excellent agreement. We note that the profiles of Ψ, U, T *and* C, exhibit symmetry with respect to the center of the cavity.

We notice that the profiles of U, T *and* C start to decrease moving away from the hot vertical wall towards the left vertical wall which is colder, this is due to the fact that the counterclockwise fluid flow allows to transport heat from the hot vertical wall located on the right and carry the cold fluid towards the part on the left after the heat removal through the upper horizontal wall.

VI. CONCLUSION

In this work we have studied theoretically and numerically heat and mass transfer by mixed thermosolutal convection in a fluid medium confined in a two-dimensional vertical rectangular cavity.

The long walls are mobile and subjected to constant heat and mass fluxes, while the short walls are assumed to be adiabatic. The boundary conditions adopted are of Newman type. The system of equations governing the problem is solved numerically by a computational code based on the finite difference method. Moreover, an analytical solution, based on the parallel flow approximation, has been obtained for the limiting case of an elongated cavity (A >> 1).

The analytical and numerical results obtained are the following:

— The validity of the use of the parallel flow approximation for (A >> 1), the obtained results showed an excellent agreement between the analytical solution and the numerical simulation.

NOMENCLATURE

Latin Alphabet

A : Cavity aspect ratio, $A = L'/H'$
H' : Height of the cavity, (m)
L' : Length of the cavity, (m)
C_T : Horizontal adimential temperature gradient, $\left(\frac{\partial T}{\partial x}\right)_{x=A/2}$
C_C : Adimentiel horizontal concentration gradient, $\left(\frac{\partial C}{\partial x}\right)_{x=A/2}$
J' : Constant mass flux per unit area (kg/(m^2.s))
q' : Constant flux of heat per unit area (W/m^2)
D : Mass diffusity of the fluid (m^2/s)
g : Gravitational acceleration (m/s^2)
P^* Characteristic pressure
t^* Characteristic time
u^* Characteristic velocity
T_0 Reference temperature (constant)
C_0 Reference concentration (constant)
T Dimensionless temperature
C Dimensionless concentration
N Ratio of thermal and solutal volume forces
(u, v) Dimensionless velocity components
(u′, v′) Dimensional velocity components
(x, y) Dimensionless coordinates
t Dimensional time
$\Delta T'$ Characteristic temperature difference
$\Delta C'$ Characteristic difference of the concentration
ΔT^* Temperature characteristic deviation $\left(= \frac{q'L'}{\lambda}\right)$
ΔC^* Concentration characteristic difference $\left(= \frac{J'L'}{D}\right)$

Dimensionless numbers

Ra Thermal Rayleigh number
Le Lewis number
Pr Pranndtl number
Nu Nusselt number
Sh Sherwood number

Greek letters

α Thermal diffusity of the fluid (m^2/s)
v Kinematic viscosity of the fluid (m^2/s)
μ Dynamic viscosity of the fluid (Pa.s)

λ Thermal conductivity of the fluid $(w/(m.k))$
ρ Mass of the fluid (Kg/m^3)
ρ_0 Density at reference temperature and concentration (constant).
β_T Coefficient of thermal expansion of the fluid (K^{-1})
β_c Coefficient of solutal expansion of the fluid (m^2/kg)
ψ Current function

Exponent

$'$ Dimensional variables
* Characteristic variable

Indices

° State of referance
i Index of a point of the following mesh x
j Index of a point of the following mesh y
c Value relative to the center of the cavity (x, y) = (A/2, 1/2)
(a,b,c,d) the ranking of the figures

Abbreviations

ADI Implicit method with alternating directions
PSOR Point Successive Over Relaxion

REFERENCES

[1] H. F. Oztop I. Dagtekin, "Mixed convection in two-sided lid-driven differentially heated square cavity", International Journal of Heat and Mass Transfer 47 (2004) 1761–1769

[2] A. Louaraychi, M. Lamsaadi, M. Naïmi, H. El Harfi et M. Kaddiri, "Étude de la convection mixte dans une cavité rectangulaire horizontale chauffée latéralement", Nature & Technology Journal. Vol. A: Fundamentals and Engineering Sciences, (2017).

[3] M. Touiker, S. Guerbaai, A. Omara, "Simulation numérique de la convection mixte double-diffusive dans une cavité rectangulaire partiellement chauffé", ICEMAEP2016, October 30-31, 2016, Constantine, Algeria.

[4] A. Louaraychi, M. Lamsaadi, M. Naïmi, H. El Harfi, M. Kaddiri, A. Raji, and M. Hasnaoui, "Mixed convection heat transfer correlations in shallow rectangular cavities with single and double-lid driven boundaries", International Journal of Heat and Mass Transfer, 132, 394–406 (2009).

[5] Mohamed A. Teamah, Wael M. El-Maghlany, "Numerical simulation of double-diffusive mixed convective flow in rectangular enclosure with insulated moving lid", International Journal of Thermal Sciences 49 (2010) 1625e1638

[6] Youssef Tizakast, Mourad Kaddiri, Mohamed Lamsaadi, "Thermosolutal Mixed Convection in Shallow Rectangular Cavity with Imposed Uniform Heat and Mass Fluxes and Filled with Newtonian Power-Law Fluid", International Journal of Heat and Technology, vol.39, NO. 3, June, 2021, pp. 669-680.

[7] G. De Vahl Davis and I.P. Jones, "Natural Convection of Air in a Square Cavity: A Compariso Exercise", International Journal for Numerical Methods in Fluids, Vol. 3, N°3, pp. 227-248, 1983.

[8] Taoufik Makayssi, "Etude analytique et numérique de la convection doublement diffusive dans des géometries rectangulaires confinant des fluids non-netoniens", These Présentée à la faculté pour obtenir le Diplôme de Doctorat-ès National, Option : Mécanique des Fluides et Energétique (2013).

Machine Learning approach based on meteorological parameters for long term Solar Radiation forecasting

Hasna Hissou*†[1], Said Benkirane[2], Abderrahim Beni- Hssane[1], and Azidine Guezzaz[2]

[1]Chouaib Doukkali university – Avenue of Faculties El Jadida 24000, Morocco, Morocco
[2]Higher School of Technology, Cadi Ayyad University – Km 9, Road to Agadir, , Essaouira Aljadida PB. 383, Ghazoua 44000, Morocco

Abstract

the sun is one of the main sources of energy on the entire globe. It is clean, free, and sustainable, It offers a lot of potentials to serve the world's energy needs and It's expected to play an important role soon. Solar radiation (Rs) is the amount of energy attaining the Earth's floor, it is used for a variety of applications and models. Nowadays, many countries are introducing solar energy into their power grids. But they face many difficulties to maintain a perfect balance between the production and consumption of electricity as the Rs is sporadic, unexpected, not steady, and not controllable. measuring instruments are not available at all locations around the world due to the considerable costs and the difficulty of installation, and other factors such as equipment failure or water and dirt accumulating in the sensors that might cause the data to be missing or out of the predicted range. To address these issues, several methods were used like empirical models, statistical models, satellite-based approaches, time series-based models, and so on. Recently, Machine Learning (ML) models (neuron-based: artificial neural networks (ANNs) and Multilayer perceptron (MLP) networks, kernel-based and tree-based models, Fuzzy logic techniques and ensemble learning methods, etc.) are thought to be a viable approach for forecasting Rs. In this paper, we present the progress report of our works including the state of the art of ML approach used in solar radiation forecasting, we discuss the background and we propose a ML model to minimize Rs forecasting's limitations.

*Speaker
†Corresponding author: hhissou@gmail.com

An Enhanced Intrusion Detection Approach For Cloud Environments

Hanaa Attou*†[1], Azidine Guezzaz[1], and Said Benkirane[1]

[1]higher school of technology,cadi ayyad university – essaouira, Morocco

Abstract

Due to its flexibility and pay-per-use services, cloud computing (CC) is currently the most popular alternative for IT organizations. Nonetheless, due of its distributed and open designs, which are prone to breaches, privacy and security concerns are significant barriers to CC's successful implementation. The open nature of the cloud is enticing potential cybercriminals. Traditional intrusion detection solutions are frequently useless because to CC's openness. In order to provide network security and protect network resources and infrastructure, an Intrusion Detection System (IDS) is a critical component. We apply an IDS based on ensemble learning approaches to improve the level of accuracy (ACC) of cloud security, in this study. Experiments on NSL-KDD dataset validate the effectiveness and practicality of the suggested IDS model. The suggested model obtains 99.99% ACC using just a few number of features from NSL-KDD dataset. As a result, when compared to earlier works, this one provides consistent results.

*Speaker

†Corresponding author: hanaa.attou@ced.uca.ma

Online learning on e-learning platforms during and after COVID-19 containment: perceptions of students at SMS University

Radoine Hamzaoui*†[1], Boutalline Boutalline‡[2], Hassan Faouzi§[2], and Belaid Boukhalene[1]

[1]Innovation Laboratory in Mathematics, Applications Information Technologies – FP-BM, Morocco
[2]Systems Engineering Laboratory – FST-BM, Morocco

Abstract

This study serves to better understand students' perceptions of the learning platforms used in e-learning during and after covid-19 confinement. Distance learning has helped Moroccan universities to ensure pedagogical continuity in general and especially in the Sultan Moulay Slimane University (USMS) . This survey aimed to investigate three factors: (1) Students' perceptions of the e-learning using platforms(2) Richness of online content and its relevance and how the students contribute to use these contents (3) Preference of type education.
A total of 293 students completed a closed-ended questionnaire following the survey method which was distributed online using the Google forms. The responses were evaluated and analyzed by SPSS software.
The results of the descriptive statistics of this study show that the e-learning program of the USMS, using online platforms was more important. they highlight the need to focus on the development of distance learning to improve its effectiveness, especially in health crises such as covid19.

*Speaker
†Corresponding author: hamzaoui.radoine@gmail.com
‡Corresponding author: boutalline@gmail.com
§Corresponding author: faouzi.hassan.mi@gmail.com

Binary Segmentation of Microscopic Bacterial Images based on Convolutional Neural Networks (CNNs)

Abderrahim El Hatimy*[1], Said Safi†[2], and Ahmed Boumezzough‡[3]

[1]abderrahim el hatimy – sidi moumen casablanca, Morocco
[2]SAID SAFI – Beni mellal, Morocco
[3]AHMED BOUMEZZOUGH – BENI MELLAL, Morocco

Abstract

Cell counting is a time consuming especially because of the large amount of data generated every day. Therefor, many computer vision based approach tried to automatically predict the number of cells in a microscopic images. Nevertheless, these methods lack of precision when it comes to very condensed bacterial image. This study aimed to use a type of convolutional neural network (CNN) called generative adversarial networks (GANs) for a binary segmentation. The dataset used for this purpose obtained from the internet, is manually croped and preprocessed by using OpenCv libraries.

*Speaker
†Corresponding author: safi.said@gmail.com
‡Corresponding author: ahmed.boumezzough@gmail.com

Browser security and privacy: challenges and issues

Abdelhadi Zineddine*†[1] and Yassine Sadqi‡

[1]Laboratory LIMATI – FPBM, USMS University, Beni Mellal, Morocco

Abstract

The privacy and security of web users represent a primary concern to governments, practitioners and cybersecurity researchers. Indeed, web browsers must be more secure as they are the first interfaces where users enter their personal information, make online purchases or exchange their various data. The aim of this paper is to examine the performance of the latest desktop browser versions (Google Chrome, Mozilla Firefox, Microsoft Edge, Opera, and Internet Explorer) with respect to privacy and security requirements, including the type of user information that an attacker can collect through potential browser vulnerabilities. According to our technical analysis, the default browser versions provided systematically leak users' confidential data. For this reason, users are advised to be more aware and adapt to the best practices required for safe browsing, as well as to add some security mechanisms to the default settings provided by their web browsers to fully protect their privacy.

*Speaker
†Corresponding author: abdelhadi.zineeddine@usms.ac.ma
‡Corresponding author: yassine.sadqi@gmail.com

A NIDS based on a Genetic Algorithm and ensemble learning for IIoT Edge Computing

Mouaad Mohy-Eddine*†[1], Said Benkirane[1], and Azidine Guezzaz[2]

[1]Higher School of Technology, Cadi Ayyad University – Km 9, Route d'Agadir, Essaouira Aljadida BP. 383, Ghazoua 44000, Morocco

[2]Higher School of Technology, Cadi Ayyad University – Km 9, Route d'Agadir, Essaouira Aljadida BP. 383, Ghazoua 44000, Morocco

Abstract

The Internet of Things (IoT) is the field that interconnects billions of objects to serve a meaningful purpose. The industrial IoT (IIoT) is the use of the IoT in the industrial world. Its purpose is to engage embedded devices in the industrial sectors to improve their activities. It is always vulnerable to various menaces, thus, IIoT security represents a big concern in the research field. The intrusion detection system (IDS) is one of the mechanisms that was introduced to mitigate IoT and IIoT vulnerabilities. IDS is implemented to monitor and detect normal and intrusion in hosts, networks, and systems. IDS accuracy and detection rate are enhanced using machine learning (ML) models. In this paper, we design and validate a network IDS (NIDS) using AdaBoost to improve the performance of our model. Moreover, we will focus on the genetic algorithm as a feature selection method to generate a more efficient dataset. The proposed model showed 99.99% and 99.89% accuracy, and 96,63% and 89.86% of AUC-Score on the Bot-IoT and NSL-KDD datasets, respectively. By detailing our proposition, we have established a classification model and validated it to assure the security task.

*Speaker

†Corresponding author: mouaadmohyeddine@gmail.com

Contribution to the Coverage Optimization in Wireless Sensor Networks

Anasse Baali*[1], Hicham Ouchitachen*†[1], and Abderrazak Farchane‡[1]

[1]Laboratoire d'Innovation en Mathématiques et Applications et Technologies de l'Information (LIMATI) – Faculté Polydisciplinaire, Béni Mellal, Morocco

Abstract

Wireless sensor networks are predestined to become an integral part of our daily life. These networks have been the subject of intensive research activity aimed at producing responses to the limitations and particularities associated with them. The different constraints related to these networks imply several contradictory objectives which must be optimized simultaneously. The sense of contradiction makes optimizing these goals a difficult task and arguably a hot topic of research for theorists and engineers. This is justified by the fact that some goals need to be minimized while some others require maximization. It is from this perspective that the research carried out in this paper comes from. Indeed, the aim of our study is to address the problem of target coverage optimization in wireless sensor networks. Thus, given a set of targets to cover and a set of sensors, we are looking for an optimization algorithm to cover a maximum number of targets using a minimum number of sensors.

*Speaker
†Corresponding author: h.ouchitachen@gmail.com
‡Corresponding author: a.farchane@gmail.com

Radar de vision par ordinateur

Hassan Facoiti*†[1], Ahmed Boumezzough[2], and Said Safi*[3]

[1]Équipe Physique Appliquée et Nouvelles Technologies (EPANT) – Université Sultan Moulay Slimane, Beni Mellal, Morocco

[2]Équipe Physique Appliquée et Nouvelles Technologies (EPANT) – Université Sultan Moulay Slimane, Beni Mellal, Morocco

[3]Laboratoire LIMATI – Université Sultan Moulay Slimane, Beni Mellal, Morocco

Abstract

Le développement des systèmes d'aide à la conduite (ADAS : Advanced Driver Assistance Systems) a connu un bond énorme au cours de la dernière décennie. Ces systèmes font partie des options offertes par la majorité des constructeurs automobiles, soit en termes de réduction de coût pour d'équiper plus de catégories de voitures intelligentes avec ces systèmes ou d'amélioration des performances face aux exigences croissantes de la sécurité routière.
Le radar automobile est au cœur de tous les systèmes d'aide à la conduite, sa fonction principale est de prendre en compte l'environnement de la voiture. Cette fonction est réalisée en détectant des cibles pertinentes autour de la voiture, à des distances variables. Des algorithmes avancés de traitement du signal sont développés par les OEM (Original Equipment Manufacturer). La qualité du signal utilisé par ces algorithmes est critique pour le succès de leur implémentation.
Ce type de radar a été consommé face aux défis des contraintes de la conduite autonome moderne. Pour cette raison, nous avons étudié la conception d'un radar de vision par ordinateur pour reproduire le comportement humain, avec des approches de détection de lignes de voies routière basées sur quatre algorithmes qui donnent de bonnes estimations en temps réel.

*Speaker
†Corresponding author: hassan.facoiti@gmail.com

Algorithmes de sécurité pour les réseaux véhiculaires Ad hoc

Khalfaoui Hafida, Farchane Abderrazak, and Safi Said

LIMATI Laboratory, Department of Mathematics and Computer Science,
Polydisciplinary Faculty, Sultan Moulay Slimane University
Beni Mellal, Morocco
hafida.khalfaoui@usms.ac.ma

Abstract. Dans ce colloque, nous présentons l'état d'avancement de notre thèse dans laquelle nous discutons les problèmes de sécurité dans les réseaux véhiculaires Ad hoc (VANET). En raison aux caractéristiques spéciales de ces réseaux, comme la topologie dynamique et la mobilité des noeuds, il existe de nombreuses vulnérabilités lorsque les attaquants souhaitent effondrer les réseaux partiels ou entiers. Par la suite, nous présentons une application de ces réseaux dans le domaine du transport et la logistique. Son objectif est de créer un écosystème digital intelligent et décisionnel (EcoDiD) efficace dans ce domaine en élaborant une solution qui intègre diverses technologies telles que Edge Computing, Blockchain, Intelligence artificielle, le data-Science et le Big data. Enfin, nous donnons nos contributions de sécurisation qui peuvent améliorer la sécurité existante et dont proposés dautres solutions plus efficaces afin de garantir la sécurité dans le système et éviter les risques de coupure dans les chaînes de transport et de distribution générant des surcoûts, des retards de distribution et finalement la dissatisfaction du client.

Keywords: VANET, attaques, sécurité, intelligence artificielle.

Effect of Semiconductor Geometric Parameters on Thermoelectric Generator Properties: Iterative Computational Study

Lahcen Ouhmad*[1] and Abdessamad Malaoui[2]

[1]Lahcen Ouhmad – Beni Mellal, Morocco
[2]Abdessamad Malaoui – Beni Mellal, Morocco

Abstract

The world's energy needs are increasing every year, which is why we need new sources of renewable energy, one of the most important being thermoelectricity. The principle of thermoelectricity is the direct conversion of heat into electricity, based on a coupling of thermal (heat transport) and electrical (circulation of electric current) phenomena. A thermoelectric module is made of a series of two or more semiconductor materials, one is P-type and the others are different N-type. In this context, through this scientific work, we study the effect of certain geometric parameters, in particular the length, of the semiconductors of the thermoelectric cells on their characteristics according to other external physical parameters, namely the temperature. The method used to solve the equations is based on an iterative calculation and discretization by finite elements, which avoids the creation of complex models. This technique allowed us to easily study and simulate the performance of these generators by taking into account the heat transfer mechanisms through the outer layers.

*Speaker

Study and Simulation of a Thermoelectric Generator Feeding Intelligent Agriculture Systems

Lahcen Ouhmad*[1] and Abdessamad Malaoui[2]

[1]Lahcen Ouhmad – Beni Mellal, Morocco
[2]Abdessamad Malaoui – Beni Mellal, Morocco

Abstract

Intelligent agriculture has become one of the priorities that have recently emerged and will be emerging in the agricultural sector in the future due to the availability of programmable electronic devices that measure and monitor the necessary parameters on the ground.
In this context, we seek to study and solve the problem of feeding local electronic devices placed in agricultural fields through this scientific work, which detects, measures and monitors plant growth by measuring plant development and indicators related to growth.
These devices are implanted in each tree or locally in each plant, and of course require continuous and permanent electrical energy to power the electrical sensors.
It should be noted that the power supply of these devices is considered one of the technical challenges for the following reasons:
Multiple electrical wiring when using the conventional electrical power source.
The use of photovoltaic energy for each measuring device/sensor/transmitter is also costly and impractical because it requires the use of photovoltaic cells and a storage battery for each of these devices.
Using a dry battery is a reliable solution, but it remains financially costly and requires intervention to replace and control these batteries.
In this scientific work, we propose a new way of producing electricity to power devices from the thermal energy present in the ground and according to the temperature difference between those of the ground and those measured in the environment. exterior, using semiconductors based on the Seebeck effect. In this work, we use special software such as "ANSYS" and "TeGDS" to simulate and design an ideal miniature electric generator of "16cm x 16cm x16cm" geometry, capable of producing electric power to power the electronic devices used. The results found show the importance of this work in the field of smart agriculture.

*Speaker

Theoretical investigation of structural, optoelectronic, elastic and thermoelectric properties of halide cubic perovskite $CsInZrX_6$ (X = Cl, Br and I)

A. Harbi[a]*, Y. naimi[a], M. Tace[a], M. Moutaabbid[a]

[a] University of Casablanca, Laboratory of Chemistry and Physics of Materials LCPM, Faculty of Sciences, Department of Chemistry, Casablanca, Morocco.

*Corresponding author.

Email: amine.harbi-etu@etu.univh2c.ma

Abstract

Perovskite materials have received wide interest from researchers in recent years because of their potential applications in the field of renewable energy. In the present paper, the halide perovskites $CsInZrX_6$ (X = Cl, Br and I) have been investigated in terms of its structural, electronic, optical, elastic and thermoelectric properties using density-functional theory (DFT). The computed lattices constants were found to be in agreement with the available experimental and theoretical results. The electronic profile exhibits a direct semiconducting band gap in the range of 1.2 eV and 3.6 eV. The optical properties show a high optical absorption coefficient on the order of 10^6 cm^{-1} and a low reflectively in visible and ultra violet regions, making these materials more suitable for solar cell applications. The study of elastic constants and their derived parameters revealed that these materials are brittle, mechanically and dynamically stable and have a covalent bond. Finally, thermoelectric properties are computed using the BoltzTraP package. The calculate values of figure of merit (ZT) for $CsInZrBr_6$, $CsInZrCl_6$ and $CsInZrI_6$ are 0.76, 0.73 and 0.725, which clearly suggests that these materials are also used as potential candidate for thermoelectric devices.

Keys works: halide perovskite; optoelectronic; thermoelectric; solar cell applications

Contribution to increasing the energy efficiency of photovoltaic cells with formamidinium perovskite layers: study and simulation by SCAPS-1D

Abdessamad MALAOUI
Equipe de Recherche en Energies Renouvelables et Innovation Technologique,
Laboratoire de Recherche en Physique et Sciences pour l'Ingénieur (LRPSI),
Université Sultan Moulay Slimane, Béni Mellal, Maroc
a.malaoui@usms.ma

Abstract

This work presents the study of new 3rd generation photovoltaic cells using SCAPS-1D simulation. The aim is to study the performance of three types of Formamidinium perovskite layers. Indeed, we have studied the effect of Cesium cations and Bromine halogens on the performance of a FAPbI3 reference cell, as well as the effect of bismuth metal by introducing an interfacial layer. The first results show the possibility of reducing the lead content and improving the extraction of the holes, as well as the performance of the cell. In addition, we have been able to show that the addition of small amounts of Cesium and Bromine increases the stability of the optically active perovskite phase, which allows a significant increase in the energy conversion efficiency of about 275%. This work also presents the comparison of the simulation results with the SCAPS-1D tool with real experimental measurements, something that will surely be useful for improving the energy efficiency of future photovoltaic cells.

Keywords: Energy efficiency, Photovoltaic cell, Perovskites, SCAPS-1D, FAPbI3.

Amélioration de la qualité des signaux électriques injectés au réseau électrique

Jamal Khyat*[1], Naima Taifi†[2], and Abdessamad Malaoui‡[2]

[1]Laboratoire de Recherche en Physique et Sciences pour l'Ingénieur (LRPSI) – Université Sultane Moulay Slimane-BENI MELLAL, Morocco

[2]Laboratoire de Recherche en Physique et Sciences pour l'Ingénieur (LRPSI) – Université Sultan Moulay Slimane-BENI MELLAL, Morocco

Abstract

L'injection d'énergie électrique produit par les sources d'énergies renouvelables avec une bonne qualité est l'une des contraintes présentées au niveau des réseaux électriques.
Dans ce travail on va voir les types des architectures d'onduleurs et leurs stratégies de commande et avec les nouvelles architectures, on va faire un traitement de signal pour produire des signaux électriques de bonne qualité. Cette amélioration peut-être faire par plusieurs étapes basée sur la gestion d'énergie et le contrôle électronique quand le passage par les appareils de chaîne d'injection.
On s'intéresse dans notre travail de faire une étude comparative au niveau de la qualité des signaux entre les onduleurs les plus connues et les onduleurs multiniveaux.

*Speaker
†Corresponding author: Taifinaima@yahoo.fr
‡Corresponding author: a.malaoui@usms.ma

Evidence based solar energy virtual learning: Evaluating students experimentation using BigBlueButton

Amine Moulay Taj*[1] and Abdessamad Malaoui[2]

[1]Moulay Taj – Sultane Moulay Slimane University of Beni Mellal, Morocco
[2]Malaoui – Sultane Moulay Slimane University of Beni Mellal, Morocco

Abstract

This paper aims to measure the using effect of BigBlueButton (BBB) as a pedagogic tool for interaction between students and teachers during the course of the Practical Works (PW). A prototype of simulation, which describes the behavior of a solar panel under varied climatic conditions, was developed and presented to two groups of university students (26 students for each group). For the first group, the presentation of the practical works uses BBB web conference system, while the presentation for the second group was given directly in the (PW) classroom by the traditional method. Questions were developed with the purpose of collecting data about the quality of the produced prototype and the use of the BBB tool in this situation using a survey functionality proposed by the BBB platform. Students' opinions in the laboratory were gathered toward the end of each practical work, and the results were analyzed by the survey software Sphinx. These results show the advantages and the limits of this new pedagogical method on the teaching of the remote Practical Works.

*Speaker

Conception d'une commande MPPT basé sur l'algorithme d'optimisation des loups gris en combinaison avec le filtre de KALMAN

Ilyass Belaalia*[1], Naima Taifi*[1], and Abdessamad Malaoui*†[1]

[1]Laboratoire de Recherche en Physique et Sciences pour l'Ingénieur (LRPSI), – Université Sultan Moulay Slimane, Béni Mellal, Maroc, Morocco

Abstract

L'Energie photovoltaïque est devenue la plus utilisé des énergies renouvelables, pour son abondance, son critère non polluant et sa maintenance à bas prix. Cependant, le rendement de conversion des cellules PV actuelle reste très faible, en plus de ça, si ces dernières ne fonctionnent pas dans un point appelé : Point de Puissance Maximal (MPP), alors son efficacité sera considérablement faible, également le changement de position de MPP en cas de changement de condition environnementale nous demande de le suivre. Dans la littérature plusieurs techniques on était adopté pour suivre ce point (MPPT). Dans cet article, nous proposons une optimisation de la commande MPPT à base de l'algorithme des loups gris " GWO " amélioré en intégrant le filtre de KALMAN pour minimiser les fluctuations autour du MPP. La technique proposée sera évaluée dans un premier temps par simulation sous MATLAB/Simulink et ensuite expérimentalement.

*Speaker
†Corresponding author: a.malaoui@usms.ma

First-principles calculations of the structural, electronic, magnetic, thermodynamic and optical properties of the cubic Sr1-xMnxTe ternary alloys

Youness Chaib*[1], Said Belaaouad†[1], Said Benmokhtar‡[1], and Mohammed Moutaabid§[1]

[1]Université Hassan 2, Casablanca – Morocco

Abstract

This article presents the structural, electronic, magnetic, thermodynamic and optical properties of the Sr1-xMnxTe (x = 0.25, 0.50 and 0.75) compounds. The systematic simulations have been studied using DFT method under the context of the full potential linearized augmented plane wave (FP-LAPW). The structural parameters of SrTe are in good agreement with that available theoretical and experimental values. The results reveal that Mn doping reduces the

*Speaker
†Corresponding author: sbelaaouad@yahoo.fr
‡Corresponding author: sbenmokhtar@gmail.com
§Corresponding author: mohammedmoutaabid@yahoo.fr

Étude théorique de la diffusion élastique d'un neutrino muonique par un électron en présence d'un champ laser intense

Sabrine El Asri*[1], Said Mouslih[†] , Mohamed Jakha[‡] , Souad Taj[§] , and Manaut Bouzid[¶]

[1]sabrine – sabrine1996.se@gmail.com, Morocco

Abstract

Compte tenu de la grande contribution de la diffusion de neutrino-électron à la compréhension profonde des interactions électrofaibles, nous nous concentrons dans ce travail sur l'étude de la diffusion élastique d'un neutrino muonique par un électron (e- $\mu \rightarrow$ e- μ) en présence d'un champ électromagnétique polarisé circulairement. Nous effectuons notre calcul théorique dans le cadre de la théorie de Fermi en utilisant les fonctions d'onde exactes des particules chargées dans un champ électromagnétique. Cette étude, ajoutée aux précédentes, va considérablement enrichir nos connaissances en physique fondamentale. Voir l'article complet publié : Phys. Rev. D 104, 113001 (2021).

*Speaker
[†]Corresponding author: moslih1992.sa@gmail.com
[‡]Corresponding author: jakhatal@gmail.com
[§]Corresponding author: s.taj@usms.ma
[¶]Corresponding author: bmanaut@gmail.com

Étude de l'effet de la direction du champ laser sur le processus de la désintégration du pion

Said Mouslih*[1], Mohamed Jakha[†] , Sabrine El Asri[‡] , Souad Taj[§] , Manaut Bouzid[¶] , and Siher El Arbi[‖]

[1]Université Sultan Moulay Slimane – Av Med V, BP 591, Béni Mellal, Maroc, Morocco

Abstract

Le choix d'une direction spécifique pour la propagation des ondes du champ laser représente souvent un problème pour les chercheurs qui étudient les processus quantiques ultrarapides assistés par laser. Ils sont confrontés à la question de savoir pourquoi exactement cette direction et pas une autre. ce travail résout cette question en ce qui concerne les processus de désintégration. Par conséquent, nous étudions théoriquement le processus de désintégration du pion en présence d'un champ laser à polarisation circulaire se propageant dans une direction générale arbitraire. En utilisant la première approximation de Bron et les états de Dirac-Volkov pour les particules chargées, nous dérivons une expression analytique pour le taux de désintégration. On a constaté que la direction du champ laser n'a pas d'effet significatif sur la nature du résultat obtenu. Cette étude généralise les résultats trouvés pour un champ avec un vecteur d'onde le long de l'axe z dans un article récent (Phys Rev D 102:073006, 2020). ce travail servira à justifier le choix d'une direction spécifique pour le champ laser dans les processus de désintégration assistée par laser. L'effet du champ laser sur le taux de désintégration total a également été rapporté et discuté.

*Speaker
[†]Corresponding author: jakhatal@gmail.com
[‡]Corresponding author: sabrine1996.se@gmail.com
[§]Corresponding author: s.taj@usms.ma
[¶]Corresponding author: bmanaut@gmail.com
[‖]Corresponding author: siher@fstbm.ac.ma

First principle calculation of thermoelectrical behavior of Simple Perovskite SrGeO3

Abderrahmane Waqdim* , Abderrahman Abbassi , Mohamed Agouri , Bouzid Manaut[1], Souad Taj , Moha El Idrissi , Aymane El Haji , and Salah Bennajare

[1]Department of physic – fst beni mellal, Morocco

Abstract

In this work, thermoelectrical properties of the simple perovskite SrGeO3 were investigated with several approximations, GGA-Pbe, unmbj, mbj, nmbj using the Full Potential-Linearized Augmented Plane Wave method which are implemented in the Wien2K code. The method used is based on the Density Functional Theory approach. Thanks to these method, the thermoelectrical properties which calculated with Bolztrap package show the reliability of these composite oxides to be exploited in many applications especially in solar cells.

*Speaker

Laser-Assisted High Energy Process

M. Ouhammou,[1] M. Ouali,[1] S. Taj,[2] and B. Manaut[1,*]
[1] *Polydisciplinary Faculty, Laboratory of Research in Physics and Engineering Sciences, Team of Modern and Applied Physics, Sultan Moulay Slimane University, Beni Mellal, 23000, Morocco.*

May 17, 2022

Abstract

In this work, we studied theoritically in the centre of mass frame, the Higgs boson production in the presence of an intense laser field via e^+e^- annihilation ($e^+e^- \to ZH$) in the framework of the electroweak standard model. By comparing our results with those obtained by DJOUADI for laser-free process, we show that the circularly polarized laser field affects significantly the s-channel Higgs boson production. In the centre of mass frame and for a given number of exchanged photons, laser field strength and frequency, the total cross section decreases by several order of magnitude. These effect of laser field on cross section is found to be consistent with what was found for muon pair production via QED process in the presence of a circulary polarized field .

Keywords: Standard model and beyond, Electroweak interaction, Laser-assisted processes.

*Corresponding author, E-mail: mh.ouhammou@gmail.com

Laser effect on unstable particle decay

Mohamed Jakha*[1], Said Mouslih[1], Souad Taj[1], and Bouzid Manaut[1]

[1]Université Sultan Moulay Slimane – Av Med V, BP 591, Béni Mellal, Maroc, Morocco

Abstract

Through this poster, I will try to popularize my area of research, which revolves around the effect of lasers on the decay processes of unstable particles, and present some of the results we have obtained.

*Speaker

Numerical simulation of InAS/GaAs quantum dots solar cells

Abdellatif Hmairrou*[1], El Hadi Chahid , Said Belaaouad , Mohamed Azza , and Abdessamad Malaoui

[1]Laboratory of Chemistry-Physics of Materials – Faculty of Sciences Ben M'Sik B.P. 7955, University Hassan II of Casablanca, Morocco

Abstract

This study will focus on the simulation and modeling of quantum dot solar cells based on InAs/GaAs, using the theoretical model based on the resolution of the Schrodinger equation :
(1)

The optimization of the solar cell structure consists to study the effect of the different parameters characterizing the InAs quantum dots such as the QD concentration, the surface recombination rate and their size.

The analytical resolution of equation (1) gives the expression for the effective energy band gap of the material with quantum dot :

(2)

Where , represents the bound exciton energy, and a is the radius of the quantum dot, and is the reduced mass.

The obtained results show that the effect of the QDs sizes has an influence on the solar cell efficiency. The efficiency in the QDs sizes from 20 nm to 5 nm, produce an increase the short- circuit density however the open circuit voltage has decreased.

*Speaker

Adsorption and photocatalytic reduction investigation of methylene blue on TiO2-bentonite surface in aqueous solution.

Youssef GHERRABY[1], Rachid CHEROUAKI[1], Said Belaaouad[1], Jamal NAJA[2]

[1] *Laboratory of Chemistry-Physics of Materials, Faculty of Sciences Ben M'Sik B. P. 7955, University Hassan II of Casablanca, Morocco.*

[2]*Applied Chemistry and Environment Laboratory. University of Science and Technology Settat 577. Morocco*

Abstract

This work investigated the adsorption and Photocatalytic degradation performance of dyes (methylene blue (MB)) on the surface of Raw bentonite (Raw-B) and doped bentonite with titanium oxide (TiO_2-B). The bentonite was selected from the Nador region in northern Morocco, it was doped with various TiO_2 mass ratios. The structural, mineralogical, and morphological composition of raw-B and TiO_2-B was characterized using several techniques including inductively coupled plasma atomic emission spectroscopy, X-ray diffraction, Fourier transformed infrared, transmission electron microscopy, and X-ray energy dispersive analysis. The adsorption process of MB was affected according to the Langmuir and Freundlich isotherms. The maximum removal capacity of MB was around 60.97 mg/g. MB degression ratios achieved 100 % under a lamp irradiation at 245 nm. The results obtained show that the application of titanium oxide-doped bentonite has a very high performance in wastewater treatment.

Keywords: Adsorption; Photocatalytic degradation; TiO_2-bentonite; Methylene blue.

Performance of perovskite solar cell using Mg-doped ZnO as electron transport layer (ETL)

Marouane Archi*[1], Omar Bajjou†[2], Lhouceine Moulaoui‡[1], Abdelhafid Najim§[3], and Khalid Rahmani¶[1]

[1]LRPSI, Research Laboratory in Physics and Sciences for Engineers, Poly-disciplinary Faculty, Sultan Moulay Slimane University – benimellal, Morocco

[2]Energy and materials engineering laboratory, Faculty of Sciences and Technics, Sultan Moulay Slimane University, BP 523, 23000 Beni Mellal, Morocco. – Benimellal, Morocco

[3]LDD, Faculty of sciences and Tehncs, Sultan Moulay SlimaneUniversity – B.P 523, 23000, Beni Mellal, Morocco, Morocco

Abstract

Perovskite solar cell is one of the best efficient devices with a low-cost fabrication, and researchers are trying to improve the performance of this solar cell device, while its power conversion efficiency has developed from 3.8% in 2009 to 25.5% in 2020.
In this study, SCAPS-1D software was used in the simulation of FTO/MZO/CH3NH3PbI3 device, where we utilized Mg-doped ZnO as an electron transport layer(ETL) to improve the performance of perovskite solar cell, while CH3NH3PbI3 is the p-layer of this structure, so it is a free HTL free perovskite solar cell.

Diverse factors that affect the solar cell performance have been investigated, such Mg-doped ZnO concentration and thickness of CH3NH3PbI3 absorber layer and the working temperature achieve a high optimized PSCs (perovskite solar cell) performance. Compared to other studies, in our simulation, the conversion efficiency of 17.19% is obtained and a fill factor of 79,58%.

*Speaker
†Corresponding author: bajjou.omar@gmail.com
‡Corresponding author: moulaoui1987@gmail.com
§Corresponding author: najim.phymo@gmail.com
¶Corresponding author: kh.rahmani@hotmail.fr

CARACTÉRISATION DES ARGILES ACTIVÉES THERMIQUEMENT (BENSLIMANE MAROC)

Khalid Azizi*†1, Maisara Jabran*‡1, Youssef Naimi*§1, and Said Belaaouad*¶2

[1]Laboratory of Chemistry and Physics of Materials LCPM – Faculty of Sciences, Department of Chemistry Avenue Cdt,Driss el harti B.P: 7955 -Ben M'Sik CP : 20800 - 20800 Casablanca, Morocco, Morocco

[2]Laboratory of Chemistry and Physics of Materials LCPM – Faculty of Sciences, Department of Chemistry Avenue Cdt,Driss el harti B.P: 7955 -Ben M'Sik CP : 20800 - 20800, Morocco

Abstract

L'objectif de mon projet de recherche est l'étude et la caractérisation de l'état d'armorphisation des argiles thermiquement activées. Les argiles sont des Aluminosilicates largement utilisé comme matériaux de construction. La cuisson de ces nanoparticules permet de créer des matériaux qui peuvent être utilisé dans l'industrie cimentière, par exemple : le Métakaolin issu de cette activation thermique ; il devient une pouzzolane de synthèse utilisée comme un ajout minérale dans les liants hydrauliques. Donc notre bute est la caractérisation de cette transformation thermique (Déshydroxylation), par différentes méthodes spectroscopiques, thermiques et la Calorimétrie de la dissolution dans l'acide fluorhydrique.
Mettre en évidence les propriétés Physico-Chimiques, ainsi une description des phénomènes minéralogiques, et de déterminées les conditions optimales de calcination permettant d'obtenir des pouzzolanes réactives.
L'objectif d'intérêt général à long terme est : trouver les conditions optimales pour avoir des pouzzolanes réactives.

*Speaker
†Corresponding author: khalidazizimaphar@gmail.com
‡Corresponding author: jabrane.maissara@gmail.com
§Corresponding author: youssef.naimi@univh2.ma
¶Corresponding author: belaaouad.s@gmail.com

Etude analytique et numérique de la convection mixte dans une cavité verticale

Sihame Ziane*[1], Mohamed Lamsaadi*[1], Hassan El Harfi*[1], and Mohamed Rahmoun*[1]

[1]laboratoire de recherche en physique et sciences de l'ingénieur.Département de physique – Faculté Polydisciplinaire (FPBM) , Université sultan Moulay Slimane (USMS), Béni-Mellal, Morocco

Abstract

Dans ce travail, on a effectué une étude analytique et numérique de la convection mixte dans une cavité rectangulaire verticale chauffée longitudinalement à l'aide d'un flux de chaleur uniforme, la cavité confine un fluide newtonien ayant Pr=7, ses parois verticales sont considérées imperméables à la chaleur. Une solution analytique basée sur l'approximation de l'écoulement parallèle a été obtenue et qui sert de référence pour la comparaison avec celle issue de la résolution numérique. L'accent est mis sur le comportement engendré par le changement des valeurs des paramètres déterminants (nombre de Rayleigh et Peclet), les résultats obtenus montrent l'accord entre les résultats analytiques et numériques.

*Speaker

Etude de quelques propriétés mécaniques du Béton Autoplaçant.

Mustapha Kajja*†[1], Hassan Bita[2], Naima Taifi[1,3], and Abdessamad Malaoui[1]

[1]Polydisciplinary Faculty - Sultan Moulay Slimane University - Maroc – Polydisciplinary Faculty - Sultan Moulay Slimane University - Maroc, Morocco

[2]Ibn Zohr University, Laayoune Technology School, Laayoune, Metrology and Information Processing Laboratory, Agadir, Morocco – Ecole Supérieure de Technologie Quartier 25 Mars BP 3007 - Laâyoune – Maroc, Morocco

[3]Laboratoire de Recherche en Physique et Sciences pour l'Ingénieur (LRPSI), – Université Sultan Moulay Slimane, Béni Mellal, Maroc, Morocco

Abstract

Le béton auto- plaçant (BAP, ou SCC, en anglais, self compacting concrete) a été développé par des chercheurs de l'université de Tokyo au japon en 1980. La première application réelle de ce type du béton a commencé en 1988. Chez les constructeurs, le choix de ce type de béton réside dans sa fluidité, qui facilite la mise en place et qu'il est aussi très utile pour les Structures avec des distributions d'armatures complexes. Dans ce travail, on présente l'efficacité de ce type de béton et son influence sur la qualité mécanique d'une construction en particulier la résistance à la compression, la résistance à la traction et la résistance à la flexion.

*Speaker

†Corresponding author: mus.enea@gmail.com

Effect of Electrode Modification Substrates on Fuel Cell Microbial Performance: Review

Siham Elmazouzi*†[1], Hassan Mabrak*[1], Salim Ayoub*[2], Fatiha Saouti*[3], Ilham Zerdani*[4], and Youssef Naimi*[1]

[1]Laboratoire Chimie Physique Des Matériaux (LCPM) – Faculté des Sciences Ben M'Sick Av Driss El Harti Sidi Othmane Casablanca, Morocco

[2]Laboratoire Chimie Physique Des Matériaux (LCPM) – Faculté des Sciences Ben M'Sick Av Driss El Harti Sidi Othmane Casablanca, Morocco

[3]Laboratoire Chimie Physique Des Matériaux (LCPM) – Faculté des Sciences Ben M'Sick Av Driss El Harti Sidi Othmane Casablanca, Morocco

[4]Laboratoire d'Ecologie et Environnement (LEE) – Faculté des Sciences Ben M'Sick Av Driss El Harti Sidi Othmane Casablanca, Morocco

Abstract

The development of new energy sources is a key challenge at the global level. The need is growing day by day for various reasons, the best known of which is the increase in the cost of oil, but also because of the growing energy needs of populations that do not have access to conventional energy resources. More and more we are looking to develop clean and renewable energy production sectors that will provide energy while reducing the environmental impact. In the recent decade, microbial fuel cells (MFC) have attracted a lot of attention because of their ability to convert organic waste into electrical current through the use of microorganisms. A microbial fuel cell is a future source of renewable energy that will produce electricity and clean effluent and will not create any type of pollution in the environment.

MFCs are similar to conventional fuel cells; where microorganisms adhered to the anode catalyze the oxidation reactions of organic matter. The diversity of anodic reactions that can be catalyzed by microorganisms opens a vast field of fuels. Concerning the performance of the anode, the electrodes used can be of high electrical conductivity, high surface and porosity, chemical and physical stability, biocompatibility and low cost. Several materials have been proposed and used, including carbon, graphite, stainless steel, platinum, glassy carbon reticule. They can be in the form of wire, grid, and plate, sponge, bar felt, tube or granules.

However, up to now the use of this system at industry level has not been foreseen and production is still low.

The researchers worldwide showed significant attention and considerable developments on MFC system for bioenergy production, based on this, many bibliographic research efforts have been made to overcome the problems and improve the performance of the MFC. The

*Speaker

†Corresponding author: siham.elmazouzi16@gmail.com

main scope of this work is to present the importance of MFC for power generation and wastewater treatment, and also the different substrates used for power generation in microbial fuel cells and the main applications. This work focuses on improving the performance of the MFC by modifying the electrode surfaces.

Keywords:
Renewable energy, microbial fuel cells, microorganisms, environment, wastewater treatment.

Comparative study of water electrolysis technologies

Hassan Mabrak*[1], Siham Elmazouzi*[1], Fatiha Saouti*[1], Driss Takky*[1], and Youssef Naimi*[1]

[1]Laboratoire Chimie Physique des Matériaux – Faculté des Sciences Ben M'Sick Av Driss El Harti Sidi Othmane Casablanca B.P 7955., Morocco

Abstract

The increase in the world's population and the evolution of living standards are leading to an increase in global energy consumption. Hydrogen is presented as a prodigious and universal solution to energy problems, as it could potentially significantly reduce dependence on (and in some cases replace) fossil fuels. Hydrogen has become very popular in recent years for several reasons: it is perceived as a clean fuel; it can be produced from a variety of energy sources, including renewables; and it can be used to produce electricity.The electrolysis of water or steam can be considered among the best green methods of producing hydrogen while using renewable energy sources.
This review presents a comparative study of water electrolysis technologies: alkaline (AWE), proton exchange membrane (PEM) and solid oxide (SOEC). The most influencing parameters on the production of hydrogen are noted as well as the possible ways of their improvements.
Keywords: Hydrogen; Water electrolysis; Green hydrogen production; Electrochemistry; Renewable energy

*Speaker

Hydrogen storage in solid materials using DFT studies

A. Benaddi[1], A. Hasnaoui[1]
[1] *Sultan Moulay Slimane University, Polydisciplinary Faculty, Laboratory of Materials Science, Mathematics and Environment, Khouribga, 25000, Morocco.*

May 20, 2022

Abstract

Development of eco-friendly transportation technologies with less carbon emissions is becoming a major necessity to protect the environment. Hydrogen presents a cleaner solution for futuristic renewable energies to overcome the drastic global warming escalation, However the set back of hydrogen as an alternative fuel is due to its difficulties in feasible storage and safety concerns. Solid-state hydrogen storage materials are used to store hydrogen efficiently and effectively. In this review, an attempt has been developed to provide a comprehensive overview of the recent advances in hydrogen storage materials(MOFs, 2D-surfaces...) in terms of capacity, content, efficiency, and mechanism of storage.

Cubic $PbCeO_3$ perovskite oxide: A compound with striking electronic, thermoelectric and optical properties explored using DFT studies

M. Agouri[1], A. Awqdim[1], A. Abbassi[1], S. Taj[1], B. Manaut[1], M. Driouich[1]
[1] *Sultan Moulay Slimane University, Polydisciplinary Faculty,*
Laboratory of Research in Physics and Engineering Sciences, Team of Modern and Applied Physics,
Beni Mellal, 23000, Morocco.

May 20, 2022

Abstract

In this study, we mainly interested on the calculation and investigation of various properties of cubic perovskite oxide $PbCeO_3$ by using the density functional theory (DFT) within Wien2k code. We treated the structural, electronic, optical and thermoelectric properties of this coumpound. The structural optimization have been calculated using PBEsol functional which showed a significant results that are in good agreement with the experiment. The electronic, thermoelectric and optical properties are treated by nmBJ approximation included in Wien2K Code. The results found show the occurrence of a gap energy which is equal to 2.68 eV and a good behavior of thermoelectric and optical aspects which make it a good candidate in different applications like sensor manufacturing and PV technologies.

Evaluation of mechanical and chemical pretreatments in the anaerobic degradation of organic matter of mixture of fruit and vegetable waste and mowed lawn waste

Sara Rhandouriate*†[1], Youssef Naimi*‡[2], Mohamed Tahiri*§[3], and Mohammed Saghir*¶[2]

[1]Laboratoire Physico-Chimie des Matériaux – Faculté des Sciences Ben M'Sick Av Driss El Harti Sidi Othmane Casablanca B.P 7955., Morocco

[2]Laboratoire Physico-Chimie des Matériaux – Faculté des Sciences Ben M'Sick Av Driss El Harti Sidi Othmane Casablanca B.P 7955., Morocco

[3]Laboratoire Synthèse organique Extraction ; Valorisation – Faculté des sciences Aïn Chock de Casablanca, Km 8 Route d'El Jadida, B.P 5366 Maarif Casablanca 20100 Maroc, Morocco

Abstract

In recent decades, waste production in Morocco has increased considerably with the country's population growth and economic and social development. This situation has made the collection and disposal of organic waste costly. This waste is often disposed of in landfills, waterways and randomly in free spaces without any control or treatment, which has impacts on public health, the environment and contributes widely to Green House Gases emission in atmosphere.
The city of Casablanca is the largest city (_˜ 6 million inhabitants) in Morocco. It has huge potential for organic waste (1.1 million tons/year). In this work, we focused on the study of anaerobic fermentation of fruits and vegetable; and mowed lawn wastes and the comparison between the mechanical, chemical and the association of the two pre-treatment of the mixture. Without pre-treatment, the co-susbtrates give rise to production of biogas at yields around 60% in methane.

The organic waste from the neighbourhood markets is heterogeneous, so it is necessary to carry out a mechanical pre-treatment with a shredder in order to homogenise the substrates and obtain a representative sample with the same characteristics of the organic waste from the neighbourhood markets that will be the subject of our study.

The chemical pre-treatment of the mixture with hydrogen peroxide (H2O2) process in anaerobic digestion, which was developed to improve the biodegradation yield of fruit and vegetable and grass waste, and can be used to increase the methane yield to 76% during biogas treatment.

*Speaker
†Corresponding author: sararhandouriate@gmail.com
‡Corresponding author: youssefnaimi@outlook.com
§Corresponding author: mohtahiri@yahoo.fr
¶Corresponding author: energiessaghir@gmail.com

The combination of the two applied pre-treatment allows the C/N ratio to be balanced, which promotes the anaerobic digestion process in mesophilic mode to be improved.
The analyses of the biogas composition in the three cases, the kinetic study of the metabolism and the determination of the respective methane content will be presented and interpreted.

Quantum size effect in Cadmium Selenium quantum dots

Nabila Jarmouni[1,2*], Marco Tomaiuolo[2], Alessio Gabbani[2], Francesco Pineider[2],Rajaa Bassam[1], Saad Belaaouad[1], Said Benmokhtar[1]

[1] *Laboratory of Physical Chemistry of Materials, Faculty of Sciences, University Hassan II of Casablanca, 7955 Casablanca, Morocco*

[2] *Department of Chemistry and Industrial Chemistry, University of Pisa, 56124 Pisa, Italy*

Email: nabilajarmouni@gmail.com

CdSe quantum dots (QDs) are semiconductor nanocrystals that exhibit excellent optical and electronic properties, including their bright luminescence, a broad excitation profile, narrow emission peaks, and exceptional photostability. In addition, their optical and electrical properties are highly size-dependent due to the quantum size effect. In this present work, we investigate the quantum size effect in CdSe quantum dots, by tuning the temperature of the reaction. The synthesis of cadmium selenide quantum dots was performed via hot injection method, using the oleic acid as the primary ligand and octadecene as the noncoordinating solvent with the addition of trioctylphosphine oxide like a secondary ligand, this approach leads to improve the size distribution of CdSe quantum dots. The CdSe QDs were synthesized at different temperatures 270 °C; 280 °C; and 315 °C with a reaction time of 50 min. The obtained CdSe QDs at different temperatures exhibited a blueshift emission wavelength ranging from 610 to 590 nm and an average diameter from 5.1 nm to 4.2 nm with increasing the temperature of the reaction.

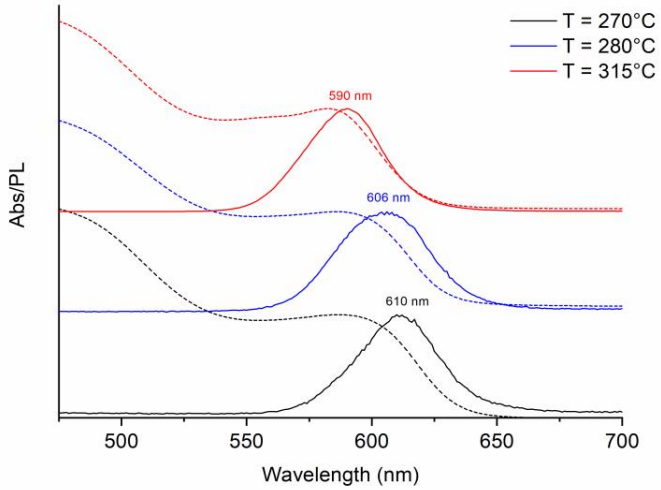


***Figure 1**: Absorption and photoluminescence (PL) spectra of CdSe QDs synthesis at different temperatures*

Key words: CdSe, quantum dots, hot injection method, different temperatures, quantum size effect.

Fenoprofen as a copper corrosion inhibitor in sulfuric acid solution

Said El Harrari* and Said Elharrari*†[1]

[1]said el harrari – Casablanca, Maroc

Résumé

Unused, defective and outdated medicines are disposed of properly according to local or international requirements, they can pollute in different sectors especially the environment. For this reason, most researchers in the field of electrochemistry and corrosion are still trying to find a more effective corrosion inhibitor that is environmentally friendly. Therefore, there is a possibility to use the award-winning pharmaceutical compounds or in the manufacturing defect phase as corrosion inhibitors. In the local market, anti-inflammatory drugs such as Fenoprofen C15H14O3, has been evaluated as a potential copper corrosion inhibitor in a commercial H2SO4 sulphuric acid solution. Following an investigation was carried out in this area by applying electrochemical and weight loss measurements for copper. The results obtained by the electrochemical techniques revealed the ability of Fenoprofen to protect copper from corrosion with a high efficiency compared to pharmaceutical products used by other researchers in the field. It is observed that the inhibition efficiency of Fenoprofen increases with increasing concentration.

*Intervenant

†Auteur correspondant: said.elharrari@gmail.com

Study of the inhibiting efficiency of the corrosion inhibitor (prop-2-yn-1-ol, methyloxirane) used in the chemical pickling of mild steels in HCl medium.

Ayoub Salim*[1], Siham Elmazouzi* , Youssef Naimi* , and Said Elharrari*

[1]Laboratoire chimie physique des matériaux – faculté des sciences ben msik driss elharti casablanca, Morocco

Abstract

The study of the inhibitory efficacy of prop-2-yn-1-ol, methyloxirane for corrosion control of mild steel in concentrated hydrochloric acid at temperature 80 ∘C (chemical pickling conditions of steel) was investigated using electrochemical and gravimetric methods for different concentrations. The morphological structure was examined by scanning electron microscopy, which clearly shows this inhibitor's good protection, which delays the acid attack of bare steel with a uniform and bright pickling appearance. The inhibitory efficiencies obtained by electrochemical and gravimetric methods are in agreement. We found the optimum inhibitor concentration to be 0.6 g/L. Keywords: corrosion inhibitor; hydrochloric acid; chemical pickling; mild steel

*Speaker

Elaboration d'une cellule photovoltaïque organique à couche photo-active de polyaniline et des nanotubes de carbone

Fatiha Saouti*[1], Soufiyan Hamham[†] , Siham Elmazouzi[†] , Hassan Mabrak[†] , Said Belaaouad[†] , and Youssef Naimi[†]

[1]Laboratoire Chimie Physique des Matériaux – Faculté des Sciences Ben M'Sick Av Driss El Harti Sidi Othmane Casablanca B.P 7955., Morocco

Abstract

Les caractéristiques électriques et optiques des cellules photovoltaïques organiques à hétérojonction dépendent intimement des conditions de préparation, de la nature des matériaux utilisés, de la taille et de la morphologie de la couche photo-active ainsi que des différentes techniques utilisées pour la caractérisation et la mise au point des différentes structures de la cellule.
En ce qui concerne les dispositifs contenant des nanotubes de carbone, leurs caractéristiques sont fortement corrélées à l'orientation des nanotubes ainsi que leurs longueurs, les propriétés des mélanges des matériaux ont aussi une influence sur les performances des cellules.

En premier lieu nous allons décrire la procédure de la mise en œuvre d'une cellule photovoltaïque à base de polyaniline et de nanotubes de carbone, deuxièmement nous allons décrire les échantillons préalablement préparés et caractérisés. Cela nous permettra donc de déceler l'influence des nanotubes de carbone ainsi que les impuretés de synthèse éventuellement infiltrées dans la préparation sur les propriétés et les performances de la cellule.
Les différentes techniques de caractérisation permettront en effet la détermination des liens entre la structure et les propriétés physico-chimiques ainsi que la morphologie et la taille de la couche active et les propriétés opto-électriques de la cellule.

*Corresponding author: saoutifatiha2@gmail.com
[†]Speaker

Efficiency of clay beads for heavy metals removal from aqueous environment

Rajaa Bassam*[1], Marouane Elalouani , Jabrane Maissara , Nabila Jarmouni , Younes Rachdi , Mohammed El Mahi Chbihi , and Said Belaaouad

[1]Laboratory of physical chemistry of materials – Laboratory of physical chemistry of materials LCPM, Faculty of Sciences Ben M'Sik, Hassan II University of Casablanca, B.P.7955, Bd Cdt Driss El Harti, Morocco., Morocco

Abstract

The optimization of the performance of the removal of heavy metal ions using clay minerals are the subject of several scientific research. This study reports on the application of clay beads for the recovery of heavy metals (Cd (II), Cr (VI), and As (III)). The process of adsorption was influenced by various effects including the pH of the solution, the initial concentrations of metals, the mass of clay beads, the contact time, and the temperature. To more understand the mechanism of Cd (II), Cr (VI) and As (III) ions elimination using clay beads, a study of characterization before and after the adsorption was carried out by the X-ray diffraction (XRD), Fourier transform infrared (FTIR), scanning electron microscopy (SEM), and atomic force microscopy (AFM). According to the results, a maximum value of the adsorption capacity of heavy metals was decreed by the Langmuir model, suggesting that the removal of Cd (II), Cr (VI), and As (III) ions were monolayer and active sites on the surface of the clay beads were homogeneous. The kinetic study showed that the adsorption of Cd (II), Cr (VI), and As (III) ions was followed the pseudo-second-order model, indicating that the elimination of these metals was dominated by chemical adsorption. The thermodynamic parameters revealed that the adsorption of heavy metals was exothermic, favorable, and spontaneous in nature. The adsorption/desorption cycles of Cd (II), Cr (VI), and As (III) were also investigated using several acids. The clay beads showed a high removal efficiency of heavy metals and accessible conditions in the treatment of groundwater pollutants.

*Speaker

Les catalyseurs ORR en métaux non précieux pour l'application des piles à combustible H2/O2

Hind Achakour*[1], Amal Antar*†[1], and Youssef Naimi*‡[1]

[1]Laboratoire physique chimie des matériaux LCPM – Faculté des Sciences Ben M'Sik ;Avenue Cdt,Driss el harti B.P: 7955 -Ben M'Sik CP : 20800 - 20800, Morocco

Abstract

Ces dernières années, une grande attention a été accordée au développement de sources d'énergie alternatives qui sont propres et efficaces ; parmi celles-ci, l'énergie électrochimique recèle un potentiel pour sa compatibilité avec l'énergie solaire et éolienne, ainsi que leurs applications dans les piles à combustible.
La pile à combustible à hydrogène d'aujourd'hui est peut-être le moyen le moins polluant de produire de l'énergie électrique, mais leur fabrication reste très coûteuse et les matériaux qui les composent s'abîment très vite alors que la pile à combustible est pensée pour durer plusieurs milliers d'heures. Cependant, le coût élevé des matières premières des catalyseurs PGM (Les métaux du groupe du platine) est devenu un problème important.

Les métaux du groupe du platine (PMG) tels que le platine, le rhodium ou le palladium sont des catalyseurs les plus courants pour la réaction de réduction de l'oxygène (ORR) en raison de leurs excellentes performances catalytiques. Mais notre étude focalisera sur l'utilisation de métaux de transition non nobles pour synthétiser des catalyseurs cathodiques, qui sont des matériaux prometteurs à faible coût pour la réaction de réduction de l'oxygène dans les piles à combustible à membrane échangeuse de protons (PEMFC).
Cette étude analyse les meilleures performances des cellules PEMFC obtenues jusqu'à présent ces dernières années lorsque des catalyseurs sans PGM sont utilisés pour l'ORR côté cathode, pour les PEMFC alimentés en hydrogène, en particulier, la méthode de synthèse et leurs propriétés structurelles sont liées à leurs capacités électrochimiques, ainsi qu'à leur comportement lorsqu'ils sont introduits dans d'autres matériaux catalytiques.

*Speaker
†Corresponding author: amal.antar@gmail.com
‡Corresponding author: youssef.naimi@univh2.ma

Évaluation des risques de sécurité dans réseau IOT et application dans smart agriculture

Maryam BOUZIDI*,
Mr.Belaid BOUIKHALENE and Mr . Said SAFI
LIMATI Laboratory, Department of Mathematics and Computer Science
Polydisciplinary Faculty, Sultan Moulay Slimane University
Beni Mellal, Morocco
maryambouzidi17@gmail.com

Résumé— Dans cet papier nous avons présenté une état d'avancement de notre mémoire dans lequel on a discuter les différentes technologies de communications sans fil utilisées dans le cadre de l'internet des objets ,Nous nous sommes focalisés principalement sur leurs caractéristiques de performances en termes de débit, portée, utilisation des bandes ISM sans licence et le coût., nous avons discuté les avantages des réseaux LPWANs pour s'accommoder aux exigences de l'IoT. Nous avons ensuite comparé les différentes technologies LPWAN en discutant leurs caractéristiques de chacune. nous nous sommes focalisés principalement sur la technologie LoRa qui offre plus d'avantages que les autres réseaux LPWAN . Toutefois, malgré les réponses aux exigences IoT,. Et aussi on a discuter les protocoles de IOT et on compare le mieux pour la sécurité et pour adaptation avec cette technologie LoRa, et finalement application dans le domaine de smart agriculture et comment on a évaluer les risques que veut attaquer .

***Index Terms*— IoT (internet of thing) ; LoRa (long Range) ; LPWAN ((Low Power Wide Area Network) ; smart agriculture**

Charged Higgs pair production inside an external field in Inert Higgs Doublet Model (IHDM)

M. Ouali,[1] M. Ouhammou,[1] S. Taj,[1] R. Benbrik,[2] and B. Manaut[1,*]

[1] *Sultan Moulay Slimane University, Polydisciplinary Faculty,*

Research Team in Theoretical Physics and Materials (RTTPM), Beni Mellal, 23000, Morocco.

[2] *High Energy Physics and Astrophysics Laboratory, FSSM, UCAM, Morocco.*

May 30, 2022

Abstract

The production of a pair of charged Higgs bosons from e^+e^- annihilation in the presence of a circularly polarized laser field is investigated in Inert Higgs doublet Model (IHDM) at e^+e^- colliders. We have derived the analytical expression for the laser-assisted differential cross section by using the Dirac-Volkov formalism at leading order including the Z and γ diagrams. Since the laser-free cross section of this process depends, in the lowest order, only on the mass of the charged Higgs boson and the energy of the center of mass, we have analyzed the dependence of the total laser-assisted cross section on these parameters and also on the laser parameters such as the number of photons exchanged, the laser field intensity and its frequency. The results found indicate that the electromagnetic field decreases the order of magnitude of the total cross section as much as the laser field intensity increases or by decreasing its frequency. Moreover, it becomes high for low charged Higgs masses and low center of mass energies.

Keywords: Standard model and beyond, Electroweak interaction, Laser-assisted processes, Cross section, IHDM.

*Corresponding author, E-mail: b.manaut@usms.ma