Sio-Iong Ao · Oscar Castillo ·
Hideki Katagiri · Alan Chan ·
Mahyar A. Amouzegar *Editors*

Transactions on Engineering Technologies

International MultiConference of Engineers and Computer Scientists 2021

Springer

Transactions on Engineering Technologies

Sio-Iong Ao · Oscar Castillo ·
Hideki Katagiri · Alan Chan ·
Mahyar A. Amouzegar
Editors

Transactions on Engineering Technologies

International MultiConference of Engineers and Computer Scientists 2021

Springer

Editors
Sio-Iong Ao
IAENG Secretariat
International Association of Engineers
Hong Kong, Hong Kong

Hideki Katagiri
Department of Industrial Engineering
and Management
Kanagawa University
Kanagawa, Japan

Mahyar A. Amouzegar
Department of Economics and Finance
University of New Orleans
New Orleans, LA, USA

Oscar Castillo
Computer Science in the Graduate Division
Tijuana Institute of Technology
Tijuana, Baja California, Mexico

Alan Chan
Department of Advanced Design
and Systems Engineering
City University of Hong Kong
Kowloon Tong, Hong Kong

ISBN 978-981-19-7137-2 ISBN 978-981-19-7138-9 (eBook)
https://doi.org/10.1007/978-981-19-7138-9

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

A large international conference on Advances in Engineering Technologies and Physical Science was held in Hong Kong, October 20–22, 2021, under the International MultiConference of Engineers and Computer Scientists 2021 (IMECS 2021). The IMECS 2021 is organized by the International Association of Engineers (IAENG). IAENG is a non-profit international association for the engineers and the computer scientists, which was founded originally in 1968 and has been undergoing rapid expansions in recent few years. The IMECS conference serves as a good platform for the engineering community to meet with each other and to exchange ideas. The conference has also struck a balance between theoretical and application development. The conference committees have been formed with over three hundred committee members who are mainly research center heads, faculty deans, department heads, professors, and research scientists from over 30 countries with the full committee list available at our conference web site (http://www.iaeng.org/IMECS2021/committee.html). The conference is truly an international meeting with a high level of participation from many countries. The response that we have received for the conference is excellent. There have been more than two hundred manuscript submissions for the IMECS 2021. All submitted papers have gone through the peer-review process, and the overall acceptance rate is 51%.

This volume contains ten revised and extended research articles written by prominent researchers participating in the conference. Topics covered include industrial engineering, electrical engineering, engineering mathematics, and industrial applications. The book offers the state of the art of tremendous advances in engineering technologies and physical science and applications, and also serves as an excellent reference work for researchers and graduate students working with/on engineering technologies and physical science and applications.

Sio-Iong Ao
Oscar Castillo
Hideki Katagiri
Alan Hoi-shou Chan
Mahyar A. Amouzegar

Contents

Secure Distributed Processing of BP with Updatable Decomposition Data

Hirofumi Miyajima[1(✉)], Noritaka Shigei[2], Hiromi Miyajima[2], and Norio Shiratori[3]

[1] Nagasaki University, 1-14 Bunkyomachi, Nagasaki City, Nagasaki, Japan
miyajima@nagasaki-u.ac.jp
[2] Kagoshima University, 1-21-24, Korimoto, Kagoshima, Japan
shigei@ibe.kagoshima-u.ac.jp, k2356323@kadai.jp
[3] Chuo University, 1-13-27, Kasuga, Bunkyoku, Tokyo, Japan
norio.shiratori.e8@tohoku.ac.jp

Abstract. Many studies have been conducted to perform machine learning while maintaining data confidentiality. For example, many studies have been conducted in this field on 1) secure multiparty computation (SMC), 2) quasi-homomorphic encryption, and 3) federated learning (FL), and so on. Methods 1) and 2) are extremely confidential in terms of security. However, their utilization of machine learning is limited. Method 3) is highly utilizable for many machine learning problems owing to the simplicity of the procedure. However, the security level is low compared to Methods 1) and 2). Previous studies have shown that both security and utility are essential for machine learning using confidential data. Therefore, it is desirable to develop a method that strikes a balance between confidentiality and utility. For this reason, in the previous paper, a secure distributed processing of BP with divided data has been proposed. However, this method has room for further improvement in that it fixes the decomposed data during learning. In this chapter, to improve the confidentiality of data, we generalize this method and present a secure distributed processing of BP with updatable decomposition data during learning. The effectiveness of the proposed method is shown in numerical simulations.

Keywords: Federated learning · Back propagation method · Updatable decomposition data · Secure distributed processing · Data confidentiality

1 Introduction

It is desirable to build a super-smart society in which big data is processed using AI to automatically retrieve advanced knowledge. To realize this society, the safety and security of data held by users must be guaranteed [1–3]. Hence, many studies haves been conducted on machine learning while maintaining data confidentiality. For example, many studies have been conducted in this field on

S. Ao et al. (Eds.): IMECS 2021, *Transactions on Engineering Technologies*, pp. 1–15, 2023.
https://doi.org/10.1007/978-981-19-7138-9_1

1) secure multiparty computation (SMC) [4], 2) quasi-homomorphic encryption [5], and 3) federated learning (FL) [6], and so on [7–9]. In these cases, Methods 1) and 2) strictly preserve privacy by using data encryption and random numbers, and Method 3) partitions all data into subsets and distributes them to each server to distribute the data. This method executes learning by distributed processing without sending the data from each server. Each approach has advantages and disadvantages. Methods 1) and 2) are extremely confidential in terms of security. However, their utilization of machine learning is limited. Method 3) is highly utilizable for many machine learning problems owing to the simplicity of the procedure. However, the security level is low compared to Methods 1) and 2). Previous studies have shown that both security and utility are essential for machine learning using confidential data. Therefore, it is desirable to develop a method that strikes a balance between confidentiality and utility. In the previous paper, we proposed the method of secure distributed processing of BP with divided data [10,11]. The method randomly divides each of the learning data and parameters into multiple pieces and uses these divided data and parameters as learning data and parameters. However, this method fixes the division of data during learning. From the standpoint of data confidentiality, it is desirable that the division of data is updated during learning. This chapter proposes a secure distributed processing method with updatable decomposition data, "divided data" updated during learning. Specifically, a secure distributed processing of BP with updatable decomposition data is proposed. The accuracy of the proposed method is compared with the conventional methods in numerical simulations. The remainder of this chapter is organized as follows.

In Sect. 2, we define the divided data in the additive and product forms as secure divided data for the BP methods. In Sect. 3, we propose learning methods based on distributed processing using updatable decomposition data independently on each server. In Sect. 4, we compare the accuracy of the conventional BP methods with those of the proposed BP methods by numerical experiments. Finally, in Sect. 5, we summarize the contributions of this study and discuss future prospects.

2 Preliminaries

2.1 Secure Computation and a Configuration Used for the Proposed Method

We explain the system used in the proposed method. In this subsection, we show the system with $Q+1$ servers in Fig. 1. We denote x and $f(x)$ by a scalar data and target function (value). We divide each data x into multiple pieces and each of them is sent to a server.

First, we divide any data x into Q pieces randomly as $x = \sum_{q=1}^{Q} x^{(q)}$. Each piece $x^{(q)}$ is sent to Server q. We calculate each function $f_q(x^{(q)})$ in Server q and it is sent to Server 0, where $f_q(\cdot)$ means a function in Server q. In Server 0, we aggregate these results and calculate $h(x) = \odot_{q=1}^{Q} f_q(x^{(q)})$, where $\odot$ means an

integrated function. If $h(x)$ is sufficiently near to $f(x)$, then we terminate the process else the same process is repeated with updated $f_q(\cdot)$.

The problem is how to determine the calculation process $f_q(\cdot)$.

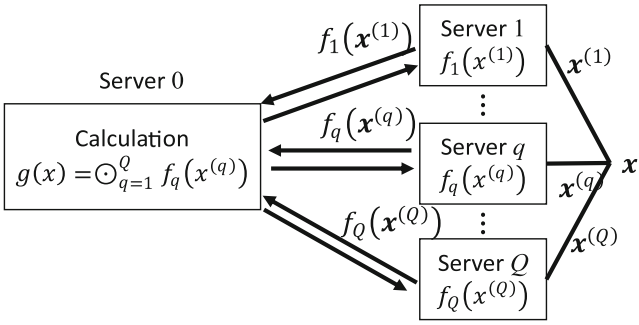


Fig. 1. An example of secure distributed processing method: the data x is divided into Q pieces. Each piece is sent to a server. Each server calculates $f_q(x^{(q)})$ and sends it to Server 0. Server 0 obtains the result $g(x)$ by integrating the partial calculations $f_q(x^{(q)})$. If $g(x)$ is sufficiently near to $f(x)$, the process terminates else the same process is repeated with updated $f_q(\cdot)$.

2.2 Decomposition Data for the Proposed Method

In this subsection, we explain how to use the divided data for the proposed method [8]. Let a and b be two real numbers. First, we divide two integers a and b into Q real numbers. We denote $a = \sum_{q=1}^{Q} a^{(q)}$ and $b = \sum_{q=1}^{Q} b^{(q)}$ as additive form and $a = \Pi_{q=1}^{Q} A^{(q)}$ and $b = \Pi_{q=1}^{Q} B^{(q)}$ as the product form. Then we have the following results.

1) $a + b = \sum_{q=1}^{Q}(a^{(q)} + b^{(q)})$, 2) $a - b = \sum_{q=1}^{Q}(a^{(q)} - b^{(q)})$

3) $a \times b = \Pi_{q=1}^{Q}(A^{(q)}B^{(q)})$, 4) $a/b = \Pi_{q=1}^{Q}(A^{(q)}/B^{(q)})$

It means that four basic arithmetic operations hold for integrating results computed independently on each server. In this case, every server cannot know the original data a and b. We explain how to divide and compute the data using an example.

Example 1. Let $a = 5$, $b = 6$ and $Q = 3$. We divide a and b randomly as $5 = 4 + (-1) + 2 = (-2) \times (-1) \times 2.5$ and $6 = 3 + (-2) + 5 = (-3) \times 2 \times (-1)$. In this case, we have $a^{(1)} = 4$, $a^{(2)} = -1$, $a^{(3)} = 2$, $A^{(1)} = -2$, $A^{(2)} = -1$, $A^{(3)} = 2.5$, $b^{(1)} = 3$, $b^{(2)} = -2$, $b^{(3)} = 5$, $B^{(1)} = -3$, $B^{(2)} = 2$ and $B^{(3)} = -1$, respectively. We can calculate $a + b$ as follows (See Table 1(a)).

$$a + b = (a^{(1)} + b^{(1)}) + (a^{(2)} + b^{(2)}) + (a^{(3)} + b^{(3)}) = 11$$

Likewise, we can calculate $a \times b$ as follows (See Table 1(b)).

$$a \times b = (A^{(1)} \times B^{(1)})(A^{(2)} \times B^{(2)})(A^{(3)} \times B^{(3)}) = 30 \qquad \square$$

Table 1. An example of decomposition data for the proposed method

(a) An example of decomposition data in additive form

	Data a	Data b	Addition
Server 1	$a^{(1)} = 4$	$b^{(1)} = 3$	7
Server 2	$a^{(2)} = -1$	$b^{(2)} = -2$	-3
Server 3	$a^{(3)} = 2$	$b^{(3)} = 5$	7
Addition	5	6	11

(b) An example of decomposition data in product form

	Data a	Data b	Multiplication
Server 1	$A^{(1)} = -2$	$B^{(1)} = -3$	6
Server 2	$A^{(2)} = -1$	$B^{(2)} = 2$	-2
Server 3	$A^{(3)} = 2.5$	$B^{(3)} = -1$	-2.5
Multiplication	5	6	30

2.3 Neural Network and BP Method

In this subsection, we explain three layered Neural Network (NN) and BP method without loss of generality [12]. Figure 2 shows an example of three layered NN. For any positive integer i, let $Z_i = \{1, 2, \cdots, i\}$ and $Z_i^* = \{0, 1, \cdots, i\}$. We define an output function $\boldsymbol{h} : J_{in}^n \to J_{out}^R$ for each $\boldsymbol{x} \in J_{in}^n$ as follows. $\boldsymbol{h}(\boldsymbol{x}) = (h_1(\boldsymbol{x}), \cdots, h_s(\boldsymbol{x}), \cdots, h_R(\boldsymbol{x}))$, where $J_{in} = [0, 1]$ or $[-1, 1]$, and $J_{out} = [0, 1]$, $[-1, 1]$ or $\{0, 1\}$. In this case, we determine weights of NN by using the set of learning data $X = \{(\boldsymbol{x}^l, \boldsymbol{d}(\boldsymbol{x}^l)) | \boldsymbol{x}^l \in J_{in}^n, \boldsymbol{d}(\boldsymbol{x}^l) \in J_{out}^R, l \in Z_L\}$ as follows, where $\boldsymbol{d}(\boldsymbol{x}^l) = \big(d_1(\boldsymbol{x}^l), \ldots, d_s(\boldsymbol{x}^l), \cdots, d_R(\boldsymbol{x}^l)\big)$ denotes the desired output for the input data $\boldsymbol{x}^l$.

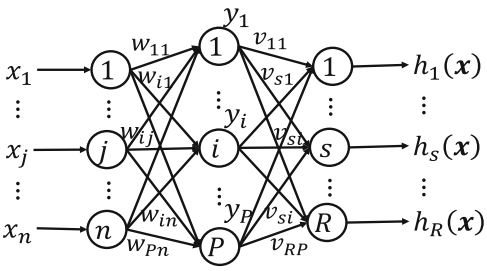


Fig. 2. An example of three layered Neural Network

We denote two sets $\{w_{ij} | i \in Z_P, j \in Z_n^*\}$ and $\{v_{si} | s \in Z_R, i \in Z_P^*\}$ by W and V, respectively. In this case, we calculate an output of NN as follows.

$$y_i(\boldsymbol{x}) = \frac{1}{1 + \exp\left(-\left(\sum_{j=0}^{n} w_{ij} x_j\right)\right))} \tag{1}$$

$$h_s(\boldsymbol{x}) = \frac{1}{1 + \exp\left(-\left(\sum_{i=0}^{P} v_{si} y_i(\boldsymbol{x})\right)\right)} \tag{2}$$

where $\boldsymbol{x} = (x_1, \cdots, x_j, \cdots, x_n)$ $(i \in Z_P,\ s \in Z_R)$ and w_{i0} and v_{s0} are thresholds, $x_0 = 1$ and $y_0 = 1$.

Then, we define the evaluation function as follows.

$$E(X, W, V) = \frac{1}{2L} \sum_{l=1}^{L} \sum_{s=1}^{R} \left(d_s(\boldsymbol{x}^l) - h_s(\boldsymbol{x}^l) \right)^2 \tag{3}$$

We update each weight of W and V using the update amounts of Eqs. (4) and (5).

$$\triangle w_{ij} = \alpha \sum_{s=1}^{S} (d_s(\boldsymbol{x}^l) - h_s(\boldsymbol{x}^l))(1 - h_s(\boldsymbol{x}^l)) v_{si} y_i(\boldsymbol{x}^l)(1 - y_i(\boldsymbol{x}^l)) x_j^l \tag{4}$$

$$\triangle v_{si} = \alpha (d_s(\boldsymbol{x}^l) - h_s(\boldsymbol{x}^l)) h_s(\boldsymbol{x}^l)(1 - h_s(\boldsymbol{x}^l)) y_i(\boldsymbol{x}^l) \tag{5}$$

Then, we show the flowchart of BP algorithm as Fig. 3. In this case, we denote the set of learning data, the maximum number of learning time, threshold, and learning rate by X, t_{max}, θ, and α, respectively.

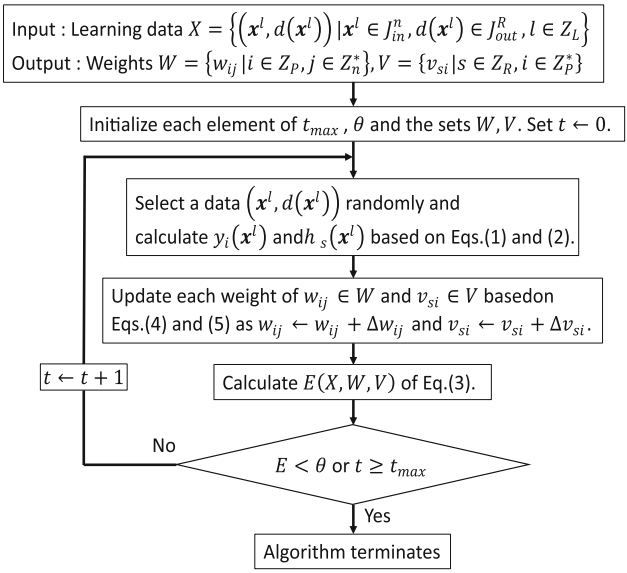


Fig. 3. The flowchart for BP method [12]

Likewise, batch and mini-batch types of BP are defined [12].

3 BP Method for Secure Distributed Processing with Updatable Decomposition Data

In this section, we propose a BP method for secure distributed processing with updatable decomposition data. The features of the proposed method are as follows.

(1) The conventional method fixes the decomposition data during learning, but we generalized it so that it can be updated at any step.
(2) Numerical simulations of function approximation and class classification show that the proposed method can be applied with almost the same accuracy as the conventional BP methods.

3.1 Data Structure for the Proposed Method

In this subsection, we divide any learning data $(\boldsymbol{x}^l, \boldsymbol{d}(\boldsymbol{x}^l)) \in X$ into Q pieces and each of them is stored in each server as follows [10].

$$x_j^l = \Pi_{q=1}^{Q} x_j^{l(q)} \tag{6}$$

$$d_s(\boldsymbol{x}^l) = \sum_{q=1}^{Q} d_s^{(q)}(\boldsymbol{x}^l) \tag{7}$$

where $\boldsymbol{x}^l = (x_1^l, \cdots, x_j^l, \cdots, x_n^l)$ and $\boldsymbol{d}(\boldsymbol{x}^l) = (d_1(\boldsymbol{x}^l), \cdots, d_s(\boldsymbol{x}^l), \cdots, d_R(\boldsymbol{x}^l))$.

In the same way, we divide any weight of W and V into Q pieces and each of them is stored in each server as follows.

$$w_{ij} = \Pi_{q=1}^{Q} w_{ij}^{(q)} \tag{8}$$

$$v_{si} = \Pi_{q=1}^{Q} v_{si}^{(q)} \tag{9}$$

In this case, we denote the sets $\{w_{ij}^{(q)} | i \in Z_P, j \in Z_n^*\}$ and $\{v_{si}^{(q)} | s \in Z_S, i \in Z_P^*\}$ by $W^{(q)}$ and $V^{(q)}$, respectively.

We have $\Pi_{q=1}^{Q} x_0^{(q)} = 1$. We calculate an output of NN for the divided data based on Eqs. (6), (7), (8), and (9) instead of (1) and (2) as follows.

$$y_i(\boldsymbol{x}) = \frac{1}{1 + \exp\left(-\left(\sum_{j=0}^{n} \Pi_{q=1}^{Q} w_{ij}^{(q)} x_j^{(q)}\right)\right)} \tag{10}$$

Further, we divide y_i as $\Pi_{q=1}^{Q} y_i^{(q)} (i \in Z_P^*)$ and $y_i^{(q)}$ is sent to Server q. We have $\Pi_{q=1}^{Q} y_0^{(q)} = 1$. We calculate an output $h_s(\boldsymbol{x}) (s \in Z_R)$ of NN in Server 0 as follows.

$$h_s(\boldsymbol{x}) = \frac{1}{1 + \exp\left(-\left(\sum_{i=0}^{P} \Pi_{q=1}^{Q} v_{si}^{(q)} y_i^{(q)}\right)\right)} \tag{11}$$

We divide the output $h_s(\boldsymbol{x})$ as $h_s(\boldsymbol{x}) = \sum_{q=1}^{Q} h_s^{(q)}(\boldsymbol{x})$ and each $h_s^{(q)}(\boldsymbol{x})$ is sent to Server q.

In this case, we calculate Mean Square Error (MSE) for the set of learning data X as follows instead of Eq. (3).

$$E(X) = \frac{1}{2L}\sum_{l=1}^{L}\sum_{s=1}^{R}\left(\sum_{q=1}^{Q}(d_s^{(q)}(\boldsymbol{x}^l) - h_s^{(q)}(\boldsymbol{x}^l))\right)^2 \tag{12}$$

Based on steepest descent method (SDM), we update each of W and V using Eqs. (13) and (14) instead of Eqs. (4) and (5) as follows [10].

$$\begin{aligned}\triangle w_{ij}^{(q)} = \alpha\sum_{s=1}^{R}\sum_{q=1}^{Q}(d_s^{(q)}(\boldsymbol{x}^l) - h_s^{(q)}(\boldsymbol{x}^l))(1-h_s(\boldsymbol{x}^l)) \\ \times\ \mathit{\Pi}_{q=1}^{Q} v_{si}^{(q)} h_s^{(q)}(\boldsymbol{x}^l)(1-h_s(\boldsymbol{x}^l))(\mathit{\Pi}_{q=1}^{Q} w_{ij}^{(q)} x_j^{l(q)})/w_{ij}^{(q)}\end{aligned} \tag{13}$$

$$\triangle v_{si}^{(q)} = \alpha\sum_{q=1}^{Q}(d_s^{(q)}(\boldsymbol{x}^l) - h_s^{(q)}(\boldsymbol{x}^l))h_s(\boldsymbol{x}^l)(1-h_s(\boldsymbol{x}^l))(\mathit{\Pi}_{q=1}^{Q} y_i^{(q)} v_{si}^{(q)})/v_{si}^{(q)} \tag{14}$$

In the previous paper, we proposed the secure distributed processing of BP with fixed decomposition data and parameters as shown in Table 2 [10].

3.2 Updatable Decomposition Data for the Proposed Method

In the previous paper, once we divided each of data x_j^l and $d_s(\boldsymbol{x}^l)$ into $x_j^{l(q)}$ and $d_s^{(q)}(\boldsymbol{x}^l)(q \in Z_Q)$, we need to use the same decomposition data during learning [10]. In this paper, we propose a secure distributed processing of BP with updatable decomposition data.

First, we show that any data $z(t)$ updated by adding or multiplying the random numbers to the divided data at the step t, becomes the same data at the next step $t+1$. That is $z(t+1) = z(t)$.

We consider $z_1(t)$ as a decomposition data in product form at the step t. That is, we define data $z_1(t)$ at the step t as follows.

$$z_1(t) = \mathit{\Pi}_{q=1}^{Q} z_1^{(q)}(t) \tag{15}$$

At the step t, we select $O_1^{(q)}(t)(q \in Z_Q)$ randomly satisfying the following equation.

$$\mathit{\Pi}_{q=1}^{Q} O_1^{(q)}(t) = 1 \tag{16}$$

Then, we hold the following relation.

$$\begin{aligned} z_1(t) &= z_1(t) \times 1 \\ &= \mathit{\Pi}_{q=1}^{Q} z_1^{(q)}(t) \times \mathit{\Pi}_{q=1}^{Q} O_1^{(q)}(t) \\ &= \mathit{\Pi}_{q=1}^{Q} O_1^{(q)}(t) z_1^{(q)}(t) \end{aligned} \tag{17}$$

Table 2. BP method. (online type) [10]

	Server 0	Server q ($q \in Z_Q$)
Initialize	Set the threshold θ and the maximum number of learning times t_{max}.	Store $\{x_j^{l(q)} \mid l \in Z_L, j \in Z_n\}$ and $\{d_s^{(q)}(\boldsymbol{x}^l) \mid l \in Z_L, s \in Z_R\}$. Initialize $W^{(q)} = \{w_{ij}^{(q)} \mid i \in Z_P, j \in Z_n^*\}$ and $V^{(q)} = \{v_{si}^{(q)} \mid s \in Z_R, i \in Z_P^*\}$.
Step 1	$t \leftarrow 0$	
Step 2	Select a number $l \in Z_L$ randomly and send it to each server.	
Step 3		Calculate $w_{ij}^{(q)} x_j^{l(q)} (i \in Z_P, j \in Z_n^*)$ and send them to Server 0.
Step 4	Calculate $y_i(\boldsymbol{x}^l)$ of Eq. (10). Divide $y_i = \Pi_{q=1}^{Q} y_i^{(q)}$ and send $y_i^{(q)} (q \in Z_Q)$ to Server q.	
Step 5		Calculate $v_{si}^{(q)} y_i^{(q)} (s \in Z_R, i \in Z_P^*)$ and send them to Server 0.
Step 6	Calculate $h_s(\boldsymbol{x}^l)(s \in Z_S)$ of Eq. (11), divide $h_s(\boldsymbol{x}^l) = \Pi_{q=1}^{Q} h_s^{(q)}(\boldsymbol{x}^l)$, and send $h_s^{(q)}(\boldsymbol{x}^l)(q \in Z_Q)$ to Server q.	
Step 7		Calculate $d_s^{(q)}(\boldsymbol{x}^l) - h_s^{(q)}(\boldsymbol{x}^l)(s \in Z_R)$ and send them to Server 0.
Step 8	Calculate the update amounts $p_{1(ij)} = \sum_{s=1}^{R} \sum_{q=1}^{Q} (d_s^{(q)}(\boldsymbol{x}^l) - h_s^{(q)}(\boldsymbol{x}^l)) \times (1 - h_s^{(q)}(\boldsymbol{x}^l)) v_{si} y_i(\boldsymbol{x}^l)(1 - y_i(\boldsymbol{x}^l)) \times (\Pi_{q=1}^{Q} w_{ij}^{(q)} x_j^{l(q)})$ and $p_{2(si)} = \sum_{q=1}^{Q} (d_s^{(q)}(\boldsymbol{x}^l) - h_s^{(q)}(\boldsymbol{x}^l)) h_s(\boldsymbol{x}^l) (1 - h_s(\boldsymbol{x}^l)) (\Pi_{q=1}^{Q} y_i^{(q)} v_{si}^{(q)})$ and send them to each server.	
Step 9		Update each of $W^{(q)}$ and $V^{(q)}$ as follows. $w_{ij}^{(q)} \leftarrow w_{ij}^{(q)} + \alpha p_{1(ij)} / w_{ij}^{(q)}$, $v_{si}^{(q)} \leftarrow v_{si}^{(q)} + \alpha p_{2(si)} / v_{si}^{(q)}$.
Step 10	Calculate $E(X, W, V)$ of Eq. (12).	
Step 11	If $E < \theta$ or $t = t_{max}$, the algorithm terminates. Otherwise go to Step 2 with $t \leftarrow t + 1$.	

That is, we update $z_1^{(q)}(t)$ by using Eq. (17) as follows.

$$z_1^{(q)}(t+1) = O_1^{(q)}(t) \times z_1^{(q)}(t) \tag{18}$$

In this case, by using Eq. (17), we obtain the following equation.

$$\Pi_{q=1}^{Q} z_1^{(q)}(t+1) = z_1(t) \tag{19}$$

Therefore, from this result, we can show that $z_1^{(q)}(t+1) = x_j^{l(q)}(t+1)$ is different from $z_1^{(q)}(t) = x_j^{l(q)}(t)$ and $\Pi_{q=1}^{Q} x_j^{l(q)}(t+1) = \Pi_{q=1}^{Q} x_j^{l(q)}(t) = x_j^l$.

The result shows that the divided data in product form is updated independently on each server.

Similarly, we consider $z_0^{(q)}(t)$ a decomposition data in addition form at the step t. That is, we divide data $z_0(t)$ into Q pieces at the step t as follows.

$$z_0(t) = \sum_{q=1}^{Q} z_0^{(q)}(t) \tag{20}$$

At the step t, we assume that $O_0^{(q)}(t)(q \in Z_Q)$ is selected randomly satisfying the following equation.

$$\sum_{q=1}^{Q} O_0^{(q)}(t) = 0 \tag{21}$$

In this case, we have the following relations.

$$\begin{aligned} z_0(t) &= z_0(t) + 0 \\ &= \sum_{q=1}^{Q} \left(z_0^{(q)}(t) + O_0^{(q)}(t) \right) \end{aligned} \tag{22}$$

By using the equations, we update $z_0^{(q)}(t)$ as follows.

$$z_0^{(q)}(t+1) = z_0^{(q)}(t) + O_0^{(q)}(t) \tag{23}$$

Then, from the Eq. (22), we have the following equation.

$$\sum_{q=1}^{Q} z_0^{(q)}(t+1) = z_0(t) \tag{24}$$

Therefore, from the results, we can show that $z_0^{(q)}(t+1) = d_s^{(q)}(\boldsymbol{x}^l)(t+1)$ is different from $z_0^{(q)}(t) = d_s^{(q)}(\boldsymbol{x}^l)(t)$ and $\sum_{q=1}^{Q} d_s^{(q)}(\boldsymbol{x}^l)(t+1) = \sum_{q=1}^{Q} d_s^{(q)}(\boldsymbol{x}^l)(t) = d_s(\boldsymbol{x}^l)(t)$.

The result shows that the decomposable data in additive form is updated independently on each server. To explain the above relation easily, we give an example of updatable decomposition data using data a in Example 1.

Table 3. Generalized BP algorithm for the proposed method. (online type)

	Server 0	Server q ($q \in Z_Q$)
Initialize	T_c denote the set of learning steps Set θ and t_{max} for decomposition data.	Store the sets $\{x_j^{l(q)} \mid l \in Z_L, j \in Z_n\}$ and $\{d_s^{(q)}(\boldsymbol{x}^l) \mid l \in Z_L, s \in Z_R\}$. Initialize the sets $W^{(q)} = \{w_{ij}^{(q)} \mid i \in Z_P, j \in Z_n^*\}$ and $V^{(q)} = \{v_{si}^{(q)} \mid s \in Z_R, i \in Z_P^*\}$.
Step 1	$t \leftarrow 0$	
Step 2	Select a number $l \in Z_L$ randomly and send it to each server.	
Step 3		Calculate $w_{ij}^{(q)} x_j^{l(q)} (i \in Z_P, j \in Z_n^*)$and send them to Server 0.
Step 4	Calculate $y_i(\boldsymbol{x}^l)$ of Eq. (10). Each $y_i(\boldsymbol{x}^l)$ is divided as $\Pi_{q=1}^{Q} y_i^{(q)}$. Each data $y_i^{(q)} (q \in Z_Q)$ is sent to Server q.	
Step 5		Calculate $v_{si}^{(q)} y_i^{(q)} (s \in Z_R, i \in Z_P^*)$ and send them to Server 0.
Step 6	Calculate $h_s(\boldsymbol{x}^l)(s \in Z_S)$ of Eq. (11). Each output $h_s(\boldsymbol{x}^l)$ is divided randomly as $\sum_{q=1}^{Q} h_s^{(q)}(\boldsymbol{x}^l)$. Each $h_s^{(q)}(\boldsymbol{x}^l)(q \in Z_Q)$ is sent to Server q.	
Step 7		Calculate $d_s^{(q)}(\boldsymbol{x}^l) - h_s^{(q)}(\boldsymbol{x}^l)(s \in Z_R)$ and send them to Server 0.
Step 8	Calculate the update amounts $p_{1(ij)} = \sum_{s=1}^{R} \sum_{q=1}^{Q} (d_s^{(q)}(\boldsymbol{x}^l) - h_s^{(q)}(\boldsymbol{x}^l))(1 - h_s^{(q)}(\boldsymbol{x}^l)) v_{si} y_i(\boldsymbol{x}^l)(1 - y_i(\boldsymbol{x}^l)) \times (\Pi_{q=1}^{Q} w_{ij}^{(q)} x_j^{l(q)})$ and $p_{2(si)} = \sum_{q=1}^{Q} (d_s^{(q)}(\boldsymbol{x}^l) - h_s^{(q)}(\boldsymbol{x}^l)) h_s(\boldsymbol{x}^l) \times (1 - h_s(\boldsymbol{x}^l))(\Pi_{q=1}^{Q} y_i^{(q)} v_{si}^{(q)})$ and send them to each server.	
Step 9		Each weight of $W^{(q)}$ and $V^{(q)}$ is updated as follows. $w_{ij}^{(q)} \leftarrow w_{ij}^{(q)} + \alpha p_{1(ij)} / w_{ij}^{(q)}$, $v_{si}^{(q)} \leftarrow v_{si}^{(q)} + \alpha p_{2(si)} / v_{si}^{(q)}$.
Step 10	If $t \in T_c$, $O_1^{(q)}$ and $O_0^{(q)}$ $(q \in Z_Q)$ are selected randomly, and sent to each server, where $\Pi_{q=1}^{Q} O_1^{(q)} = 1$ and $\sum_{q=1}^{Q} O_0^{(q)} = 0$.	
Step 11		If $t \in T_c$, each data of $\{x_j^{l(q)} \mid j \in Z_n, l \in Z_L\}$ and $\{d^{(q)}(\boldsymbol{x}^l) \mid l \in Z_L\}$ is updated as follows. $x_j^{l(q)} \leftarrow O_1^{(q)} \times x_j^{l(q)}$ and, $d^{(q)}(\boldsymbol{x}^l) \leftarrow O_0^{(q)} + d^{(q)}(\boldsymbol{x}^l)$.
Step 12	Calculate $E(X, W, V)$ of Eq. (12).	
Step 13	If $E < \theta$ or $t = t_{max}$, algorithm terminates. Otherwise go to Step 2 with $t \leftarrow t + 1$.	

Example 2. We divide data a as $a = 4 + (-1) + 2$, where $a^{(1)} = 4$, $a^{(2)} = -1$ and $a^{(3)} = 2$. Based on the above processing, we select $O_0^{(1)} = -5$, $O_0^{(2)} = 3$ and $O_0^{(3)} = 2$ randomly, where $\sum_{q=1}^{3} a^{(q)} = a$ and $\sum_{q=1}^{3} O_0^{(q)} = 0$. By Eq. (23), we update $a^{(1)}(t)$, $a^{(2)}(t)$ and $a^{(3)}(t)$ as follows.

$$a^{(1)}(t+1) = 4 + (-5) = -1,\ a^{(2)}(t+1) = -1 + 3 = 2,\ a^{(3)}(t+1) = 2 + 2 = 4.$$

In this case, we have $\sum_{q=1}^{3} a^{(q)}(t+1) = -1 + 2 + 4 = 5 = a$.

Further, we divide data a as $a = (-2) \times (-1) \times 2.5$, where $A^{(1)} = -2$, $A^{(2)} = -1$ and $A^{(3)} = 2.5$. Based on the proposed processing, we select $O_1^{(1)} = 0.5$, $O_1^{(2)} = -1$ and $O_1^{(3)} = -2$ randomly, where $\Pi_{q=1}^{3} A^{(q)} = a$ and $\Pi_{q=1}^{3} O_1^{(q)} = 1$. By Eq.(18), we update $A^{(1)}(t)$, $A^{(2)}(t)$ and $A^{(3)}(t)$ as follows.

$$A^{(1)} = -2 \times 0.5 = -1,\ A^{(2)} = -1 \times -1 = 1,\ A^{(3)} = 2.5 \times -2 = -5$$

In this case, we have $\Pi_{q=1}^{3} A^{(q)}(u+1) = -1 \times 1 \times (-5) = 5 = a$. □

3.3 Secure Distributed Processing of BP with Updatable Decomposition Data

We propose a secure distributed processing of BP with updatable decomposition data using the result of Subsect. 3.2.

We denote T_c by the set of learning steps for updatable decomposition data, where $T_c \subseteq Z^*_{t_{max}}$. We explain the outline of algorithm in Table 3. We divide the data and the weights and each of the divided ones is stored in a corresponding server. We perform calculation by using stored data on each server and calculate an output of NN on Server 0. We divide the resulting output into Q pieces and each piece is sent to a corresponding server. We calculates the error between the received data and the desired data on each server and sends the calculated error to Server 0. We integrates the errors on Server 0, calculate the update amounts of weights, and send them to each server. We update the weights using the amounts on each server. We update the divided data if the learning step is a member of the set T_c. Finally, we examine the termination condition.

We explain the algorithm of Table 3 in detail. As the initial condition, we divide each of data $\boldsymbol{x}^l$ and $\boldsymbol{d}(\boldsymbol{x}^l)$ and each of the sets W and V into Q pieces and each of them is stored on each server. At Step 3, we calculate each of products of weight and data on each server and we send the results to Server 0. At Step 4, we calculate each output y_i of the second layer in Server 0, we divide the data y_i into Q pieces and each of them is sent to each server. At Step 5 on each server, we calculate the product of $v_{si}^{(q)}$ and $y_i^{(q)}$ and sends the result to Server 0. At Step 6, we calculate each output $h_s(\boldsymbol{x}^l)(s \in Z_R)$ and divide it randomly into $\sum_{q=1}^{Q} h_s^{(q)}(\boldsymbol{x}^l)$ on Server 0 and each of pieces is sent to each server. At Step 7, we calculate the difference $d_s^{(q)}(\boldsymbol{x}^l) - h_s^{(q)}(\boldsymbol{x}^l)$ on each server and they are sent to Server 0. At Step 8, we calculate update amounts $p_{1(ij)}$ and $p_{2(si)}$ using the result of Step 7 on Server 0 and we send them to each server. At Step 9, we

update each weight of W and V on each server. At Step 10, if $t \in T_c$, we select the numbers $O_1^{(q)}$ and $O_0^{(q)}$ randomly and send them to each server. At Step 11, if $t \in T_c$, we update each of $x_j^{l(q)}$ and $d_s^{(q)}(\boldsymbol{x}^l)$ on each server and send them to Server 0. At Step 12 on Server 0, we calculate the MSE. At Step 13, if the final condition of learning is satisfied, the algorithm terminates otherwise go to Step 2 with $t \leftarrow t+1$.

In the method in Ref. [10], we fixed the divided data during learning. That is, it means that Steps 10 and 11 are removed in Table 3. In the proposed method, by inserting Steps 10 and 11, the algorithm is generalized so that decomposition data can be updated for $t \in T_c$.

4 Numerical Simulations

In this section, we compare the proposed BP method with the conventional methods in numerical simulations of function approximation and pattern classification.

4.1 Function Approximation

In the simulation, we use four systems specified by the following functions with $[0,1]^4$ for Eqs. (25) and (26) and $[-1,1]^4$ for Eqs. (27) and (28).

In this section, we have $P = 10$ and $S = 1$ in Eqs. (1) and (2). As simulation conditions, we have $t_{max} = 50000$, and constants K_w and K_v are 0.01. If the MSE for learning data is smaller than the threshold θ or the learning time is over the maximum number t_{max} of learning times, the algorithm terminates. Threshold θ is 0.0. As the initial condition, we divide data and weights into 3 pieces randomly, that is $Q = 3$. We select the numbers of learning and test data randomly as 1000 and 1000, respectively.

$$y = \frac{(2x_1 + 4x_2^2 + 0.1)^2}{37.21} \times \frac{(4\sin(\pi x_3) + 2\cos(\pi x_4) + 6}{12} \tag{25}$$

$$y = \frac{\sin(2\pi x_1) \times \cos(x_2) \times \sin(\pi x_3) \times x_4 + 1.0}{2.0} \tag{26}$$

$$y = \frac{(2x_1 + 4x_2^2 + 0.1)^2}{74.42} + \frac{4\sin(\pi x_3) + 2\cos(\pi x_4) + 6}{446.52} \tag{27}$$

$$y = \frac{(2x_1 + 4x_2^2 + 0.1)^2}{74.42} + \frac{(3e^{3x_3} + 2e^{-4x_4})^{-0.5} - 0.077}{4.68} \tag{28}$$

After learning, we calculate and compare MSEs of learning and test data for each method. We show the results of accuracy for each method in Table 4. In Table 4, we denote the conventional BP, the BP presented in Ref. [10] and

the proposed methods with $T_c = \{t|t \text{ is randomly selected}\}$ and $|T_c| = 5$ and $T_c = \{t|t \bmod 100 = 0\}$ by Methods A, B, C and D, respectively. In each box of Table 4, we denote MSEs ($\times 10^{-4}$) for learning and test by Learning and Test, respectively. The result of the simulation is the average value from twenty trials. As shown in Table 4, we show that the proposed methods are almost the same accuracy compared to the conventional methods.

Table 4. Simulation results for function approximation

		Eq. (25)	Eq. (26)	Eq. (27)	Eq. (28)
A	Learning	0.45	9.76	0.57	0.83
	Test	0.61	11.79	0.66	1.06
B	Learning	0.57	15.72	0.74	1.41
	Test	0.77	18.35	0.82	1.66
C	Learning	0.60	17.38	0.90	1.81
	Test	0.71	19.34	1.01	2.13
D	Learning	0.60	13.61	0.87	1.54
	Test	0.71	16.36	0.97	1.80

4.2 Pattern Classification

In this subsection, we perform pattern classification for Iris, Wine, Sonar, BCW, and Spam data as shown in Table 5 [13] using conventional and proposed methods. In Table 5, we denote the numbers of data, input #data:L, #input : n, and #output : R, and output are L, n, and R, respectively. In this section, we have $P = 10$ and $S = 3$ (for Iris and Wine) or $S = 2$(for Sonar, BCW, and Spam) in Eqs. (1) and (2). In the simulations, we use the 5-fold cross-validation as the evaluation method. As the simulation condition, we have $t_{max} = 50000$, and constants K_w and K_v are 0.01. If the MSE for learning data is smaller than the threshold θ or the learning time is over the maximum number t_{max} of learning times, then the algorithm terminates. Threshold θ is 3.0×10^{-2} for Iris and Wine, 4.0×10^{-2} for Sonar and BCW, and 8.0×10^{-2} for Spam, respectively. As the initial condition, we divide data and weights into 5 pieces randomly, that is $Q = 5$.

After learning, we compare the rates of miss-classified data (RM) for learning and test data for each method. In Table 6, we show the results of simulation. In Table 6, we denote the conventional BP, the BP in Ref. [10] and the proposed methods with $T_c = \{t|t \text{ is randomly selected}\}$ and $|T_c| = 5$ and $T_c = \{t|t \bmod 100 = 0\}$ by Methods A, B, C and D, respectively. RMs for Learning and Test denote the rates of miss-classified data for learning and test, respectively. We denote each value in Table 6 the RM(%) and it is the average from twenty trials.

Table 5. Data used in simulations

	Iris	Wine	Sonar	BCW	Spam
#data: L	150	178	208	683	4601
#input: n	4	13	60	9	57
#output: R	3	3	2	2	2

As shown in Table 6, we show that the proposed methods are almost the same accuracy compared to the conventional methods.

Table 6. Simulation results for pattern classification

		Iris	Wine	Sonar	BCW	Spam
A	Learning	2.08	0.52	0.91	2.36	4.60
	Test	2.93	1.78	17.83	3.01	6.43
B	Learning	1.90	1.06	2.66	2.26	4.66
	Test	2.12	2.51	17.80	2.48	4.81
C	Learning	1.90	1.10	2.90	2.23	4.69
	Test	4.50	4.22	18.57	2.96	5.86
D	Learning	1.80	1.18	2.81	2.20	4.72
	Test	4.43	3.81	19.26	3.14	5.86

5 Conclusion

This section proposed a secure distributed processing of BP with updatable decomposition data as a generalized learning method of our previously proposed method. In FL, which is the conventional method of distributed processing for machine learning, the data was partitioned into subsets, and the learning data itself were used in each server. However, unlike the conventional method, the proposed method is that learning is performed using the divided data and parameters without using the learning data itself. In our previously proposed method, once the learning data was divided into pieces, the division data were fixed during learning. On the other hand, our new proposed method updates the data division during learning, which improves the confidentiality of the data. Our numerical simulations demonstrate that the accuracy of the proposed method is almost the same as for conventional methods. In future works, we will develop methods with updatable decomposition for other machine learning models.

Acknowledgment. This work was partially supported by a research grant from Telecommunications Advancement Foundation (TAF) and Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number 20K11994.

References

1. Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., Ayyash, M.: Internet of things: a survey on enabling technologies, protocols, and applications. IEEE Commun. Surv. Tutor. **17**(4), 2347–2376 (2015)
2. Chen, J., Ran, X.: Deep learning with edge computing: a review. Proc. IEEE **107**(8), 1655–1674 (2019)
3. Sarwar Murshed, M.G., Murphy, C., Hou, D., Khan, N., Ananthanarayana, G., Hussain, F.: Machine learning at the network edge: a survey. arXiv:1908.00080 [cs.LG] (2019)
4. Canetti, R., Feige, U., Goldreich, O., Naor, M.: Adaptively secure multi-party computation. In: STOC 1996, pp. 639–648 (1996)
5. HElib: An Implementation of homomorphic encryption. https://github.com/shaih/HElib
6. Yang, Q., Li, Y., Chen, T., Tong, Y.: Federated machine learaning: concept and applications. ACM Trans. Intell. Syst. Technol. **10**(2), 1–19 (2019). Article 12
7. Aggarwal, C.C., Yu, P.S.: Privacy Preserving Data Mining: Models and Algorithms. Springer, Heidelberg (2009). https://doi.org/10.1007/978-0-387-70992-5. ISBN 978-0-387-70991-8
8. Miyanishi, Y., et al.: New methods to ensure security to increase user's sense of safety in cloud services. In: Proceedings of The 14th IEEE International Conference on Scalable Computing and Communications (ScalCom-2014), pp. 859–865 (2014)
9. Evans, D., Kolesnikov, V., Rosulek, M.: A pragmatic introduction to secure multi-party computation. Found. Trends Priva. Secur. **2**(2–3), 70–246 (2018)
10. Miyajima, H., Shigei, N., Miyajima, H., Miyanishi, Y., Kitagami, S., Shiratori, N.: New privacy preserving back propagation learning for secure multiparty computation. IAENG Int. J. Comput. Sci. **43**(3), 270–276 (2016)
11. Miyajima, H., Shigei, N., Miyajima, H., Shiratori, N.: Federated learning with divided data for BP. In: Lecture Notes in Engineering and Computer Science: Proceedings of the International MultiConference of Engineers and Computer Scientists, 20–22 October 2021, Hong Kong, pp. 94–99 (2021)
12. Gupta, M.M., Jin, L., Honma, N.: Static and Dynamic Neural Networks. IEEE Press, Wiley-Interscience (2003)
13. UCI Repository of Machine Learning Databases and Domain Theories. https://archive.ics.uci.edu/ml/datasets.php

Applications of Engineering 4.0 to Improve the Safety of Metalworking Operators: The Ansaldo Energia Case

Roberto Mosca[1], Marco Mosca[1], Saverio Pagano[1], Roberto Revetria[1], and Gabriele Galli[2(✉)]

[1] University of Genoa, 16126 Genoa, GE, Italy
{marco.tullio.mosca,roberto.revetria}@unige.it,
saverio.pagano@smarttrack.it
[2] University of Michigan, Dearborn, MI 48128, USA
ggalli@umich.edu

Abstract. The paper describes how, on behalf of Ansaldo Energia Spa, a multi-disciplinary team developed a methodology based on Industry 4.0 technologies, an approach that allows rescue teams to quickly intervene in the event of a man-down in isolated areas of the plant, where the unfortunate person would risk being found with significant delay and consequent problems for his physical well-being. Under the supervision of the team, a highly specialized supplier created a suitable hardware and software device to achieve this outcome. Such a device can immediately warn rescue crews in real time as soon as an incident occurs, as well as geo-locate the man on the ground with exceptional precision.

Once developed, the approach was standardized in a set of sequential and generic procedures in order to make it adaptable to any sort of firm, construction site, or workshop where a man-down event may happen. The methodology is set up as a real toolkit to protect operators from severe damage that can result from long waits for rescue teams, whenever operators experience negative events for their safety being them exogenous (fainting illnesses, heart attacks, epileptic attacks, strokes, etc.) or endogenous (accidents in the workplace).

Keywords: Artificial intelligence · Blockchain · Case-study · Connectivity · Digital Twin · Industry 4.0 · Man-down · Safety · Sensors

1 Introduction

An important objective of the project (indicated as OR6) was dedicated by the company to the study and subsequent implementation of tools-and-methodologies suitable for increasing the safety of its operators in the context of the Ansaldo Energia LHP Project (co-financed by the Ministry of Economic Development) for the adoption by the company of Industry 4.0 standards [1, 2].

It must be taken in consideration that a worker dies every 15 s on the planet, and 153 employees are injured in the workplace every day [3]. This is why Ansaldo has chosen

S. Ao et al. (Eds.): IMECS 2021, *Transactions on Engineering Technologies*, pp. 16–30, 2023.
https://doi.org/10.1007/978-981-19-7138-9_2

to improve the safety of its operators by using Industry 4.0 technology and applications as part of the "smart safety" initiative.

Protecting the man-down condition, which occurs when an operator is lying on the ground and is unable to move in isolated areas of the plant, outside the field of vision of any other operators or safety officers, is an aspect of safety that the project team, the company, and the workers' organizations consider to be of particular importance.

He may stay in this area permanently, with serious implications for his safety (e.g., in the case of a stroke the person injured must receive medical specialized intervention within 45 min to prevent permanent damages).

In addition to the Ansaldo Energia application in Genoa factory, the outcome is a generic technique that may be repeated in many different operational circumstances, as discussed in detail below.

The technique is built on cutting-edge hardware and software solutions that were custom-designed by the team and developed by the provider, and were chosen for their proven dependability and experience.

The suggested approach is able to offer consistent results after a cautious period of testing, followed by a number of early implementations in the organization, thus it can be regarded a legitimate instrument for ensuring the safety of the unlucky operators who encounter this situation.

2 Literature Review

The following three steps were used to conduct the literature review: (i) identification of keywords and their combinations; (ii) database selection; (iii) analysis of findings. Due to the limited amount of information accessible in the literature, "man-down" was the sole keyword employed, and its search was applied to the Article Title, Abstract, and Keywords. Two separate abstract and citation databases of peer-reviewed literature have been chosen for the second search stage (ii): Scopus and Google Scholar. Twenty-six articles were collected at the end of this phase.

Only six publications published between 2010 and 2021 have been included in the search. A fresh systematic study of key specifics was undertaken during the third stage (iii). This literature study demonstrates that researchers have spent only a few energy in this area, which is critical for worker safety.

Guilbeault-Sauvé et al. understood the need of having a dependable, durable, and simple-to-use system with a low false alarm rate, quick reaction time, and acceptable ergonomics. The suggested approach is based on three observable states: worker falls, worker immobility, and worker on the ground. These are based on the characterization of bodily movement and orientational data from accelerometer and gyroscope sensors. The technique was tested on a public database and found to be 99% accurate [4].

Tayeh et al. developed a system based on Long-Range (LoRa) technology to detect man-down situations in persons conducting tasks in areas where there is no network connectivity.

This is accomplished by sending an alarm message via LoRa that includes the individual's status as determined by a system that includes a GPS-enabled IoT device, a wristwatch, and a smartphone [5].

Cerruti et al. have proposed a system called Smart Safety. The authors developed a system for real-time position estimate that consists of a network, a number of IoT sensors, and a central monitoring system. It is particularly beneficial in a variety of circumstances, including man down. The technology sends different notifications to the worker based on their position, which is presented in real-time in a three-dimensional plant map. The efficiency of this system has been demonstrated in a variety of industries, including power plants, petrochemical plants, refineries and oil processing facilities [6].

Schlesser et al. developed a technique to improve worker safety, particularly in the blast furnace area, which is exposed to a variety of hazards. Because GPS cannot be considered as a good choice for inside position location, the authors created Smart Safety Guard, a tracking system that uses specific features like geofencing to enable live visualization of every worker's position in the form of a visual depiction [7].

For safety surveillance, Ugolini implemented a geo-localization and communication system. It can locate and identify all operators inside a production area that is covered by Wi-Fi, GSM-GPRS-UMTS, Bluetooth, and ZigBee networks and is monitored by a CCTV system. Each operator is outfitted with a unique helmet that includes a headset and microphone, as well as a Wi-Fi/GSM-COM device and a ZigBee belt tag for location tracking [8].

The oldest of the systems examined is Johnson et al.'s, which integrates Wi-Fi and location-based technology with gas detectors to allow companies to remotely monitor situations in areas where wireless networks were previously unavailable. The Accenture Life Safety Solution is intended for use in hazardous work areas such as refineries, chemical plants, and other sites where dense steel infrastructure makes wireless safety solutions difficult to implement [9].

3 Material and Methods

3.1 Detection and Surveillance of High-Risk Areas

Installation of man-down devices and accompanying acoustic/visual detection systems for the geo-location of the operator in the case of an emergency is part of this endeavor (e.g., illness). The installation of these devices will affect operators who work in remote locations where proper surveillance results objectively impossible.

The team's solution will enable the creation of a speedier and more effective warning system through the use of an intelligent tracking system based on logic and rules for the analysis of growing data correlations (Bayesian approach). The potential of a system certification carried out in complete conformity with the law's principles and directives in terms of privacy will be assessed. It should be noted that the technology does not monitor the operators' movements and is used exclusively to geo-locate man down events for this reason.

Furthermore, several prediction models based on machine learning techniques will be constructed and evaluated using the acquired data to offer evolutionary signals on the risk situations under consideration.

During the course of an operator's daily activities, one of the many accidental causes (such as a fortuitous fall, a bump on the head, an illness resulting in loss of consciousness,

the sudden onset of a serious pathology such as heart attack, stroke, epileptic attack, etc.) may result in the operator collapsing to the ground and being unable to move.

When such an event occurs in areas of the plant where other operators are present, any worker can issue an alert to summon the rescue team; however, if the same event occurs in isolated or confined areas (where continuous surveillance is not possible), the accidents may lead to tragic consequences for the unfortunate individual (as it is precisely the timeliness of the rescue that, in many cases, can make the difference between survival and death, or the possibility of avoiding permanent damage).

As a result, the team deemed a collection of devices capable of reporting to the operation center when any operator monitored is laying on the ground following an accident or sickness as a critical component for ensuring the operators' safety.

Therefore, an outstanding supplier (Smart Track) has been identified as capable of producing a tool based on the team's specifications that, as a consequence of the man-down event, allows to geo-locate the exact position where the event occurred, in real-time, using appropriate acoustic-visual systems, ensuring the rescue team's prompt intervention. To get to the point where the device could be manufactured, the team took many stages under the supervision of the adviser.

First and foremost, once the supplier was chosen, the team worked with Ansaldo and the supplier to define the technical specifications for a system that included a wearable device, a reference beacon, and management software that could fully meet the team's technological expectations, which were deemed essential by the team and whose unalienable goal was to provide the fastest possible assistance to injured operators.

3.2 Device

Different types of sensors have been selected and chosen for the detecting side, keeping in mind the functions intended by the device:

- 1 accelerometer to detect a sudden fall by an operator and/or a protracted lack of movement by the operator and the angle of inclination of the operator's body in relation to the expected vertical axis
- 1 temperature/barometric pressure sensor for the operator's safety by sensing the operator's distance from the ground and/or the unwanted event of any fires
- 1 GPS sensor for the wounded operator's geo-location

Figure 1 depicts the device.

Fig. 1. Device used

On the communication front, it should be taken in consideration that the device doesn't communicate with the anchors (beacons), but rather listens to them for geo-localization (a principle that's almost identical to that of a lighthouse in navigation). By "sniffing" the Bluetooth signal emitted by the anchors, the device is able to determine its precise location and communicate it in the event of an alarm. To activate the alarms, the device also talks with the server.

The communication device is equipped with the following features:

- New integrated BT (Bluetooth) 5.1 control (communication with anchors)
- 2.4 GH Wi-Fi (communication with the LAN)
- Machine-to-Machine roaming with GSM/GPRS Sim cards (communication with the central server)
- NB (narrowband) - IoT 4G (geolocation with ±5 m precision)
- Extendable to UWB (ultra-wideband) with 50cm precision (additional applications include COVID19 detection for gatherings, forklift man collision, hands-free entry management in some business areas by detecting inclination angle in the pocket without swiping the badge…)

The device is powered by a LiPo (lithium polymer) battery that can be recharged in 60 min through USB. Depending on how it's used, it can last from 2 days to 12 months. It's worth noting that the device's primary function is to send the information gathered by the sensors to the nearest security officer, concierge (guard), and infirmary via redundant real-time notifications (call, messaging, email, SW Smart Studio interface).

When the instrument identifies any risky incident, the operator wearing the device is notified by a buzzer and an appropriate vibration before the alerts are transmitted to the security personnel, the infirmary, and the reception.

In addition to the above-mentioned functions, the device must also supply the rescue squad with the required elements for the geo-location of the man-down. The device works in conjunction with other equipment (beacons) known as anchors to accomplish this task.

The current number of operators covered each shift under the LHP OR6.2 project is 29.

Anchors

Each machine or inside the isolated or limited regions to be monitored has a variable number of reference anchors (beacons) installed, so that each device worn by the operators present can continually refer to one of the anchors installed.

Figure 2 depicts the anchor.

Fig. 2. Anchor

The anchors send out a Bluetooth signal that the device can pick up. As a result, the anchors are unable to interact with the system. The anchors' technical equipment includes:

- BT 5.0, a beacon for connecting to devices (maximum communication distance: 80 m)
- 1 Amp lithium button cell batteries
- Ambient temperature sensor (5 years)

There are already 86 anchors put as part of the LHP OR6.2 project.

The functional system is depicted in Fig. 3.

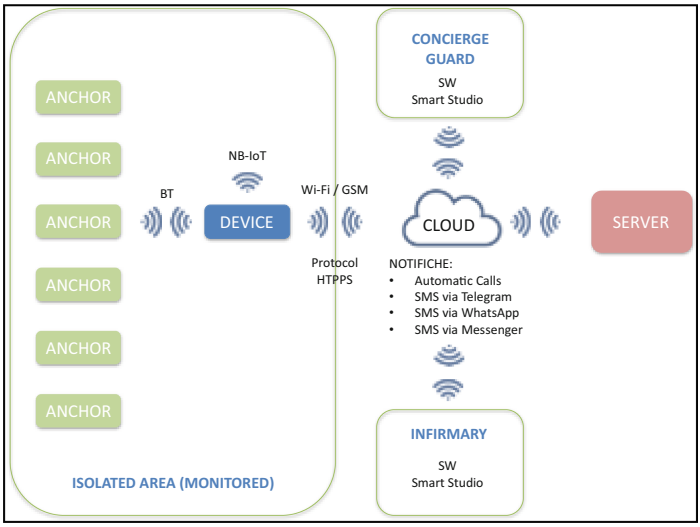


Fig. 3. Functional scheme

3.3 Smart Studio

It is a graphical interface software installed at rescue team desks, such as the heads' cabin (supervisors), the concierge (guard), and the infirmary, that allows for precise geo-location of the man-down by indicating the site, the area, the department, the machine, the position of the reference anchor on it, and the device ID worn by the operator.

The program also offers specific information on the impacted operator, including his or her name, last name, and IP address, as well as the accident's geo-localized position, address and device type (release), as well as the Bluetooth signal threshold, status, and battery level. The program may also provide a history of all man-down occurrences that have happened across the production.

3.4 Respect for Privacy

As previously stated, the device has been designed to fully respect the operators' right to privacy, so it does not monitor or record any operators' path or any interruptions to regular activities (such as breaks, toilet visits, coffee shops…), and instead is limited to geo-locating and disclosing detected man-down events.

3.5 Other Functions

The device has also two essential functions:

- the operator's voluntary and preventive alert submission through an appropriate button installed on the device, which may be used to alert the rescue team even though the scenario is unrelated to the man-down (e.g., injuries of various kinds, malaise, request for intervention for third parties)

- the processing of false alerts, with the option of canceling them

This is when the operator makes deliberate moves that the device may misinterpret, such as activities that he must undertake while lying down on the ground, sudden leaps to achieve a certain position, periods of stillness (e.g., the need to work on the computer or moments of rest), and so on.

3.6 Organization

On an organizational level, the team has assigned one device (complete with battery charger) per operator every shift to each department and has created a set of essential documentation for the proper administration of safety tasks, including:

- A list of training participants, which is required to determine whose operators are authorized to operate the device.
- Instruction manual, to ensure compliance with the correct use and maintenance of the device supplied
- Battery management manual, to improve the efficiency of the device and extend the life cycle of the power supply components
- Informed consent to participants, to ensure that operators and their safety representatives understand the importance of the system and approve its use in full respect of privacy
- A delivery and return register, in which the assignment of devices to operators is recorded.

4 Further System Developments

After the first cycle of implementations in Ansaldo, the authors carried on the phases of research and development necessary for the activation of some new functions previously identified, which consist of the following:

1. Extension and accuracy of connectivity
2. Integration and enhancement of the Software
3. Digital Twin
4. Artificial intelligence
5. Service 4.0
6. Blockchain

The interactions of all those technologies are shown in Fig. 4.

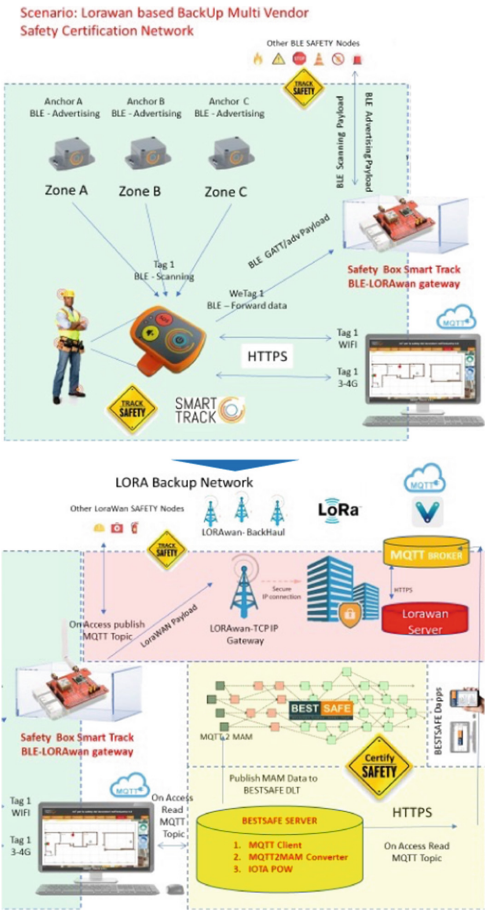


Fig. 4. Functional scheme of the safety system

4.1 Extension and Accuracy of Connectivity

The scope of this job was to extend the applicability of the service to companies that work on multiple sites, dislocated throughout the territory and not belonging to the same LAN (Local Area Network). The extension is applied to all multinational companies, consortia, business associations that want to guarantee a safety standard. The software is designed to offer an overview, with the possibility of drilldown to reach the target level of detail, with a geo-localization accuracy of 50 cm, through the extension from the previous NB-IoT 4G (Narrow Band Internet of Things), whose accuracy reaches ± 5 m, at the UWB (Ultra-Wideband) standard. In addition to the obvious benefits generated in the field of safety, the positive effects are not negligible in terms, for example, of scoring in the event of participation in tenders, an area in which site safety is of primary importance.

4.2 Integration and Enhancement of the Software

In accordance with the 4.0 standards, which provide for the maximum integration of company systems (Departmental, WMS, ERP, …), the authors decided to deeply investigate an already performing software, with the specific aim of maximizing the benefits produced by the system, by sharing the data collected and processed with other company platforms. A point of particular interest was identified in the transformation of system analytics, initially conceived exclusively as a data logger, now becoming progressively more suitable for the generation of usable information from the data source. The new functions include, for example, the creation of statistics on the frequency and type of injury in order to allow to trace the root causes of accidents, for improving prevention, on the one hand, and the timeliness of rescue operations on the other. Again, the system integration produces second-order positive effects, although no less important. In fact, by integrating the system with the internal quality function, the company benefits from an effective tool in the event of an AUDIT by customers, or certifying bodies, or in the event of participation in tenders.

4.3 Digital Twin

The replication of the system in cyberspace opens the doors to completely new horizons, in fact this represents the transition from the basic characteristics of monitoring and control to a system connected online and real time, able to "live the life" of the physical system in parallel (through the connection that allows continuous data collection) and an autonomous "cyber-life", being adaptive according to the collected data.

This means:

- both being able to take advantage of the real database, to project an accelerated replication of events into the future, thus being able to learn the critical issues and predict accidents
- being able to generate system perturbations, through the setting of new scenarios that can impact on the case history of accidents (such as the increase in work in confined areas, the extension of work to different shifts, the introduction of new machines, layout,…)

By associating basic Machine Learning algorithms, the system will thus be able to learn simultaneously from the experience and from the simulated, in view of a continuous improvement of the overall safety performance (e.g., trends, seasonality, cyclicality that can lead to risk phases, which can be mitigated through the introduction of ad hoc procedures).

4.4 Artificial Intelligence

Cognitive algorithms applied to the Digital Twin allow it to improve performance in order to ensure Continuous Improvement. In this view, the role of these algorithms is to raise the System to an Expert System capable of using the large amount of data, partly collected, and partly simulated, to learn from experience with logic comparable to how

the Human brain would do, with the difference of the processing capacity offered by the computers available today, on a scale that is not comparable to the performance achievable by man. The authors therefore consider the application of A.I. as the normal evolution to complete the Digital Twin. The algorithms used will range from neural networks to decision trees, to Bayesian classifiers, to bagging boosting and stacking techniques, to deep learning.

4.5 Service 4.0

In consideration of the trend of companies focus on core business, aimed at obtaining better business performance and, at the same time, on non-core services, the idea was to make the monitoring and emergency intervention relating to the Man Down event. The concept is based on the centralization of the continuous monitoring service (performed in outsourcing) in a service center shared by multiple companies, thus requiring a smaller number of pure local rescue operators (located as insourcing). In this way, in accordance with the BCG 11/2018 definition of Service 4.0, the service offered could adapt to the growing expectations of a constantly evolving clientele, transforming:

- the service offering from reactive to proactive, from standardized to custom (focused on the safety of the individual operator), from experience-based to data-driven.
- the provision of the service from remote to present (multi-channel), from predefined paths to dynamic paths (formed in real time) and from separate to shared systems.

Once again, thanks to the level of innovation brought about by the adoption of 4.0 standards, the benefits produced in terms of Operator safety are widely evident.

4.6 Blockchain

Smart Track (innovative Startup and spin-off of the University of Genoa that develops IoT systems for the safety and security of workers in indoor and complex environments in which the GPS does not work) is configured as a project partner, who developed the system, covered with a specific patent, called BEST SAFE (nr. 102019000021537 released on date 08/03/2022), whose purpose is to certify the safety of workers on blockchains. A further feature is therefore defined, namely the possibility of directing the tracking of data from sensors, equipment and IIoT applications in industrial and professional scenarios of various types (for example, and not exclusively Industrial Plants, Construction Sites, Logistics Systems…) in which operate multiple Suppliers, using heterogeneous tools, networks, and protocols, to make the reconstruction of events objectively reliable for:

- Insurance (e.g., objective reconstruction of the exact sequence and concatenation of events)
- Legal (e.g., assessment of any behavioral responsibilities)
- Maintenance (e.g., compensation for any coverage of the equipment and systems involved)

- Healthcare (e.g., reconstruction of social interactions between workers in case of need for "contact tracing" during pandemic crises)

In particular, the architecture at the base of BestSafe method provides for the presence of:

- A semantic protocol, called BestSafe 3S (Safety & Security Statements), for the preparation of worker Safety tracking messages (shareable, interoperable, and interpretable both by Human Operators and in M2M exchanges between third-Party applications and machinery).
- An external register called BestSafe Registry containing the different vocabularies and profiles, to be used for different sets of declarations that can be expressed according to the BestSafe 3S Language for different operational contexts (eg. Industrial Plants, Construction Sites, Construction Sites Shipping…)
- A DLT (Distributed Ledger Technologies) type distributed archive, called "BestSafe Ledger", on which various Providers can publish messages expressed in BestSafe 3S semantics, in order to guarantee both the necessary levels of interoperability and immutability and marking at the same time independent temporal of the same.

It is estimated by the Authors that BestSafe will facilitate the penetration and use of the blockchain to certify the use of innovative solutions for safety in the workplace with legal and insurance value, in particular with a distributed new generation blockchain infrastructure (IOTA TANGLE) designed directly to certify data transactions.

In particular, BestSafe 3S is a semantic representation of interoperability that allows Companies to describe data, metadata and observations coming from peripheral IIoT sensors for Personnel Safety and Plant Safety, making them uniformly interpretable and processable by heterogeneous Actors, both in Human communication scenarios and M2M.

5 Implementation Process and Process Phases

Given the necessity of monitoring man-down incidents through a system that can prompt rescue teams' action, thereby ensuring the unfortunates' survival, the test phase was given special attention.

A test step like this was required to ensure that the system met the expectations of the team who implemented it.

Both Ansaldo, the adviser, and the supplier who constructed the system demanded that the team examine the system's resilience in a very thorough and thorough manner.

This is because, once implemented, the system must be able to ensure its operation and efficacy in a wide range of conditions, depending on the type of location where the incident happened (distance from the receiving stations, size of the area, type and arrangement of the machinery…).

After conducting an in-depth analysis of the Genoa plant's production departments, Ansaldo and its team have chosen the following departments to conduct the pilot project: the Large Mechanical Department (MECG) in AREA CAMPI 1, the Medium Mechanical

Department (MEME), and the Rotors Department (ROTO), all of which are located in the FEGINO AREA.

Once the test phases were completed, the same Departments (MECG, MEME, and ROTO) were chosen as areas and departments to expand the project, gradually increasing the number of machines under observation, then the Ultra Speed Test Cell for Rotors in AREA FEGINO, the Diagnostic Center in AREA CAMPI 2, and the Mechanical Tests Workshop in AREA BOSCHETTO.

Following the identification of the departments, it was necessary to proceed with the list of machines to cover and, as a result, to determine the number of beacons (in relation to the machine's dimensions) and their relative positioning on the field in order to ensure adequate coverage of the Department's isolated areas.

Before beginning with the installation of the new system, it was agreed, in complete agreement with the company, to have a series of preparatory informational sessions with the area managers, the RLS (workers' safety managers), the security officers, the infirmary, and the concierge (guard).

The objective of these sessions is to deliver timely information on the new system to all operators of the Functions who will be actively involved in its implementation, as well as to raise awareness of the need of a tool capable of quickly activating assistance to accident operators.

Ansaldo and the team have defined the areas and departments to which the project will be extended in the future based on the results of the first phase of the test.

The next step was to install the beacons on the machines in the previously identified positions, which was not an easy task because, on the one hand, it had to avoid the risk of damaging complex and delicate machines, and on the other hand, it had to ensure stable beacon anchoring in any operating conditions of the machines, taking into account vibrations, large movements, and various configurations.

As a result, a double-sided sticky tape, adequately approved for such purpose, was chosen as a remedy, to be used after meticulous cleaning/degreasing of the afflicted surface.

To prevent the anchors from being accidentally removed, an informative sign was erected next to each one (reporting that they are part of the LHP Ansaldo project financed by MISE in the safety field).

The registration of the anchors to achieve geo-localization and placement is the final step in this phase.

The security personnel' mobile phones were then updated with notifications, and the smart studio software was installed on the computer in the chiefs' cabin.

However, in order to proceed with the pilot project, adequate training for the department operators was required, as well as having them sign the document on informed consent to the project and the participation register, after which the device, along with its battery charger, was delivered and notified to the infirmary and concierge (guard).

The favorable conclusion of the pilot project allowed us to gain team agreement for the system's adoption, the testing departments, and the project's expansion to the previously indicated regions.

6 Conclusions

Given the critical necessity for timely man-down rescue, any organization wishing to mitigate the effects of bad occurrences should use the recommended methodology.

The traditional management of Man Down events, with all of the well-known limits and associated safety risks, can be transformed into the best breed of an Engineering 4.0 application, as demonstrated by the methodology devised by the team, implemented by Ansaldo, and required by the new company, thanks to this system.

The system has been subjected to a cycle of continuous improvement, as per one of the fundamental principles of Engineering 4.0, such that the first model has been improved in progress by generating second-generation devices and anchors (or Bluetooth 5.0 Communication, vibration enhanced, impact-resistant polycarbonate injected shell, previously prototyped in resin with 3D printing).

Because the suggested system was broad in nature, it was also standardized at this time. By following the set of consecutive stages outlined below, it was feasible to modify it for any sort of Production and Processing Department:

1. The organization appoints a project leader who is an expert on worker safety problems (granting him the powers and budget appropriate to the project ojectives)
2. The corporation appoints a technical and scientific advisor to oversee the technologies selected and their implementation.
3. The company's selection of an acceptable technology supplier
4. Identifying the regions and departments to which the project will be applied
5. Identifying the machines included therein and their dimensions
6. Determination of the number and location of beacons required to guarantee constrained area coverage
7. Meetings with area leaders, RLS, security officers, infirmary, and concierge to provide information
8. Beacons are installed in the departments as part of the trial project.
9. Installation of security personnel' mobile phones with alerts
10. Installing the smart studio software in any rescue team's workplace is number ten.
11. Departmental operator training
12. System functioning test
13. Commissioning of the system when all of the preceding tests have been passed.
14. Extending the tested projects to other areas of the company

Lastly, the authors want to emphasize the importance of all those measures that can reduce the occurrence of man-down events, including, far from being secondary, the protective clothing and PPE (Personal Protective Equipment) that operators must wear, which vary depending on the type of company, job, task, and specific activities.

References

1. Mosca, R., Mosca, M., Pagano, S., Revetria, R., Galli, G.: Engineering 4.0 to improve the safety of plant operators in a metalworking company of international importance: the Ansaldo

Energia Case. In: Lecture Notes in Engineering and Computer Science: Proceedings of the International MultiConference of Engineers and Computer Scientists 2021, 20–22 October, Hong Kong, pp. 196–201 (2021)
2. Project Presentation. https://www.mecspe.com/it/appuntamenti/presentazione_ansaldo/. Accessed 08 Apr 2022
3. International Labor Organization (ILO), World Statistics. https://www.ilo.org/moscow/areas-of-work/occupational-safety-and-health/WCMS_249278/lang--en/index.htm. Accessed 08 Apr 2022
4. Guilbeault-Sauvé, A., De Kelper, B., Voix, J.: Man down situation detection using an in-ear inertial platform. Sensors **21**(5), 1–15 (2021). https://doi.org/10.3390/s21051730
5. Tayeh, G.B., Azar, J., Makhoul, A., Guyeux, C., Demerjian, J.: A wearable LoRa-based emergency system for remote safety monitoring. In: International Wireless Communications and Mobile Computing Conference, Limassol, Cyprus (June 2020)
6. Cerruti, C., Ferrari, M., Ferrari, A., Magurno, G., Taurino, R.: Application of digital technologies and IoT systems to enhance safety - smart safety project. In: Society of Petroleum Engineers - SPE International Conference and Exhibition on Health, Safety, Environment, and Sustainability 2020, HSE and Sustainability (2020)
7. Schlesser, N., Hansen, F., Reuter, Y.: Smart safety guard - improving safety through real-time people localization. In: AISTech - Iron and Steel Technology Conference Proceedings, vol. 1, pp. 10–18 (2020)
8. Ugolini, C.: SAFESTAR - safety improvements in any industrial environment. In: AISTech - Iron and Steel Technology Conference Proceedings, vol. 1, pp. 55–58 (2014)
9. Johnson, K., Bogard, K.: Unique safety approach to tracking personnel in the plant along with hazardous gas information. In: NPRA Annual Meeting Technical Papers, vol. 1, pp. 243–246 (2014)

Effect on E-Service Quality and Perspective of Customers on Customers' Loyalty in the Banking Industry in Hong Kong

Hon Keung Yau(✉) and Sin Yee Tang

Department of Advanced Design and Systems Engineering, City University of Hong Kong, Kowloon Tong, Kowloon, Hong Kong
honkyau@cityu.edu.hk, sintang4-c@my.cityu.edu.hk

Abstract. As the International Financial Centre, Hong Kong Bank sector is an important industry. With the rapid adoption of e-banking services in the bank industry, e-banking services should be addressed. Good e-banking service quality helps banks to gain a competitive advantage. There are different services provided by the bank sector online and offline. It is a trend that people are able to manage their own accounts using e-banking services. This new technology gives new challenges to the bank sector. Whenever new technologies and services is presented, it needs to be adopted by their users. The acceptable use of e-banking services can be influenced by different factors. This research aims to explore the variables that have influenced the quality of e-banking service and the perspective of customers about customers loyalty. A survey will be conducted, and the samples will be collected with varying personal characteristics such as income, gender, educational level, age, and employment situation in using e-banking services. Different attributes from scales have been selected to test the relationship among factors. This study would analyze the factors affecting the e-banking service's quality and the perspective of customers to know their attitudes of customers.

Keywords: Banking industry · Customers loyalty · Customer satisfaction · E-service quality · Perspective of customers · Responsiveness

1 Introduction

Internet technology has become a part of daily life in a few decades. The Internet enhances the quality of people. Various activities such as shopping, ordering food, and money transactions can be done through websites. Electronic banking is one type of internet banking. Consumers can have financial transactions online using an electronic payment system [1, 8]. E-banking service has become a major business for banks. In the situation of e-banking services in Hong Kong, the investigation of the interaction between customers and financial institutions (e.g. banks, etc.) is an important element. The quality of the e-service dimension model has been developed to determine the user's perception of the e-banking system [4]. A questionnaire survey will be conducted in order to help bank managers to know more about e-banking users' attitudes toward online e-banking

S. Ao et al. (Eds.): IMECS 2021, *Transactions on Engineering Technologies*, pp. 31–39, 2023.
https://doi.org/10.1007/978-981-19-7138-9_3

websites and identify the factors that have a significant impact. In this research, we focus on the following questions:

1. What dimensions of e-banking services and perspectives of customers are important to customer loyalty?
2. What kind of influence do the dimensions have on the degree of customer satisfaction and loyalty?

2 Study

2.1 Electronic Banking

Electronic Banking is all forms of banking services and products using electronic delivery channels. The common delivery platforms are Internet Banking, mobile phone app, automatic teller machines (ATMs), and telephone banking. The Hong Kong Monetary Authority defines electronic banking as a financial service provided to personal or business customers through an electronic device [3]. E-banking service includes 4 major categories such as Internet banking, financial services with self-service terminals, contactless mobile payments, and phone banking.

2.2 E-Service Quality Dimension

Initially, seven dimensions of e-SERVQUAL are the guideline for assessing the quality of e-service. The different unique e-service processes had been fitted, Parasuraman, Zeithaml & Malhotra split seven dimensions into two constructs which are E-S-QUAL and RecS-QUAL respectively. E-Service Quality (E-S-QUAL) contains 4 core dimensions while E-Recovery Service Quality (E-RecS-QUAL) includes 3 core dimensions [4]. Both scales provide properties for a variety of reliability and validity assessments based on the e-service platform. Dimensions of quality of e-service are chosen and shown below.

2.3 E-Service Quality (E-S-QUAL)

Efficiency: The speed and ease of accessing the e-service site is to find the information of their designated products or services.

2.4 E-Recovery Service Quality (E-RecS-QUAL)

Responsiveness: Handling problems and returns through the website should be effective.
Contact: Availability of assistance with instant communication tools such as telephone or online customer services.

2.5 Customer Perspective

Customers' Perspective measures the performance of the organization from an external customers' point of view. It is a crucial factor in creating financial success for banks. In the fierce competition among banks, customer retention is a significant strategy in the banking environment. The financial institution needs to obtain customers' requirements information for further marketing planning in order to limit customer defection. Lack of knowledge, poor service utilized, and poor criteria of the bank may lead to customer defection [5]. The WebQual is the scale for measuring customers toward the criteria of e-service quality [7]. WebQual scale consists of 22-item with 3 core dimensions: usability, quality of information, and quality of service interaction [7]. The WebQual scale aims at assessing e-commerce website quality from the perspective of customers.

2.6 Usability: The Association Quality with Site Design

Information Quality: The accuracy and suitability of the information in the content of the website. Service Interaction Quality: The interaction experience by customers with trust and empathy.

2.7 Customer Satisfaction

Customer satisfaction is a measurement of attitude towards the specific service by customers' perceptive. It is important that customer expectations should be well understand customer expectations on the experience of e-service. According to previous study of online transaction, satisfaction is positive related to loyalty. Satisfaction has been determined to influence loyalty of customers [6].

2.8 Customer Loyalty

Customer loyalty is positive related to customer satisfaction, customer experiences, and overall value of the service that meet their needs. Loyalty is one of main elements to e-banking recently. Loyalty has been defined as an attitude of customers' favorable and commitment towards the electronic banking service [6]. Owing to the quick accessible, district broadness, and real time online service, customers compare the service quality easily. Bank needs to perform differentiate from competitive bankers on influencing the satisfaction and loyalty of customers.

3 Development of Hypotheses

3.1 Three Service Quality Dimensions

H1: Efficiency is positively related to the quality of e-banking service.
H2: Responsiveness is positively related to the quality of overall e-banking service.
H3: Contact is positively related to the quality of overall e-banking service.

3.2 Three Perspectives of Customers

H4: Usability is positively related to the perspective of customers.
H5: The quality of Information is positively related to the perspective of customers.
H6: The quality of Service Interaction has a positive relationship with the perspective of customers.

3.3 Customer Satisfaction

H7: The quality of overall e-banking service is positively related to the satisfaction of customers.
H8: The perspective of customers is positively related to the satisfaction of customers.

3.4 Customer Loyalty

H9: The satisfaction of customers is positively related to the loyalty of customers.

4 Findings

In this study, we conducted a survey to collect data and information from respondents. The targets are the user who has experience in using the e-banking service.

4.1 Population and Sample

In this study, we focused on e-banking service users. They have experienced the products or services before the survey. They tried other's products and services and compare the differences among companies. 180 questionnaires were distributed to target respondents and 165 copies were returned. Table 1 shows the demographic characteristics of respondents.

4.2 Data Collection Method and Analysis

The questionnaire was distributed among universities with different backgrounds, including Bachelor's students and Postgraduate students through google Forms and paper. Most university students have experience in using e-banking services for paying the tuition fee. The questionnaire was also distributed to relatives and working partners. The research may find out some interesting patterns in using online services if the respondents are of different ages and employment statuses.

This questionnaire consists of three sections. The first section contains the usage data of e-banking services including length, frequency, and useful way for respondents. The second part is the issue of customers' perception of e-banking services. There are total 31 closed-ended questions. The general personal information of the respondents, including their gender, age, educational certificate, current employment, and average monthly income is required in the last part.

Table 1. Demographic information

Variables	Statistics
Gender	
Males	49.4%
Females	50.6%
Age	
18 to 30 years old	82.7%
30 to 50 years old	17.3%
Education	
Secondary school and below	0.6%
Undergraduate and postsecondary	82.1%
Master or above	17.3%
Employment	
Respondents' employment (full-time)	46.3%
Respondents' employment (part-time)	23.5%
Student	30.2%
Average monthly income	
<$1000	9.3%
$1000–$5000	25.3%
$5001–$10000	39.5%
$10001–$20000	16.7%
$20001–$30000	9.3%

4.3 Data Analysis and Findings

Moreover, we used Cronbach's alpha (Table 2) to evaluate the reliability of the measurement scale. It is recommended that the cutoff criteria, Since Cronbach's alpha is larger than 0.7 [2], means that the scales are reliable.

Table 3 shows the factor loading table. The values were close to 0.3 indicating an acceptable value for the measurement scales [2].

Table 2. Reliability

Scale	Cronbach's alpha
Efficiency	0.731
Responsiveness	0.716
Contact	0.752
Usability	0.735
Information quality	0.73
Quality of service interaction	0.734
Quality of e-banking service	0.752
Perspective of customers	0.747
Satisfaction of customer	0.727
Loyalty of customer	0.748

5 Hypothesis Testing

The Pearson correlation coefficient measures the strength of the relationship and the direction of the relationship with positive and negative between two variables. The calculation of the Pearson correlation coefficient is generated by SPSS. If the significance value (2-Tailed) is less than or equal to 0.05, there is a significant correlation between the two variables (H1–H9).

The Pearson correlation coefficient between efficiency and overall e-banking service quality was 0.236. The p-value is $0.003 < 0.05$, so efficiency is positively related to the quality of the overall e-banking service. Since the Pearson correlation coefficient is + 0.236, the relationship between efficiency and quality of overall e-banking service was positively significant.

H1: Efficiency has a positive association with the quality of overall e-banking service which was supported.

The Pearson correlation coefficient between responsiveness and the quality of overall e-banking service was 0.273. The p-value is $0.000 < 0.05$ so responsiveness is positively correlated to the quality of overall e-banking service. Since the Pearson correlation coefficient is +0.273, the relationship between responsiveness and overall e-banking service quality was positively significant.

H2: Responsiveness has a positive association with the quality of overall e-banking service, which was supported.

The Pearson correlation coefficient between contact and the quality of overall e-banking service was 0.156. The p-value is $0.048 < 0.05$ so contact is positively related to the quality of overall e-banking service. Since the Pearson correlation coefficient is +0.156, the relationship between contact and the quality of overall e-banking service was positively significant.

Table 3. Factor loading

Items	Factor loading
Efficiency 1	0.411
Efficiency 2	0.418
Efficiency 3	0.480
Fulfillment 1	0.778
Fulfillment 2	0.492
Availability of system 1	0.613
Availability of system 2	0.405
Availability of system 3	0.695
Privacy 1	0.411
Privacy 2	0.613
Responsiveness 1	0.496
Responsiveness 2	0.390
Contact 1	0.771
Contact 2	0.394
Usability 1	0.780
Usability 2	0.545
Quality of information 1	0.462
Quality of information 2	0.735
Quality of Information 3	0.566
Quality of service interaction 1	0.819
Quality of service interaction 2	0.591
Quality of service interaction 3	0.717
Quality of overall e-banking service	0.689
Perspective of customers	0.743
Satisfaction of customer 1	0.575
Satisfaction of customer 2	0.596
Satisfaction of customer 3	0.772
Satisfaction of customer 4	0.612
The loyalty of customer 1	0.612
The loyalty of customer 2	0.414
The loyalty of customer 3	0.837

H3: Contact has a positive association with the quality of overall e-banking service quality, which was supported.

The Pearson correlation coefficient between usability and the perspective of customers was 0.199. The p-value is 0.011 < 0.05 so usability is positively related to the perspective of customers. Since the Pearson correlation coefficient is +0.199, the relationship between usability and the perspective of customers was positive and significant.

H4: Usability is positively related to the perspective of customers, was supported.

The Pearson correlation coefficient between information quality and the perspective of customers was 0.224. The p-value is 0.004 < 0.05 the quality of Information is positively related to the perspective of customers. Since the Pearson correlation coefficient is +0.224, the relationship between information quality and the perspective of customers was positive and significant.

H5: Information Quality is positively related to the perspective of customers, was supported.

The Pearson correlation coefficient between service interaction quality and the perspective of customers was 0.177. The p-value is 0.025 < 0.05 so there is a significant correlation between the two variables. Since the Pearson correlation coefficient is + 0.177, the relationship between the quality of service interaction and the perspective of customers was positively significant.

H6: Service Interaction Quality is positively related to the perspective of customers, was supported.

The Pearson correlation coefficient between the quality of overall E-banking service and the satisfaction of customers was 0.289. The p-value is 0.000 < 0.05 so the quality of overall E-banking service is positively related to the satisfaction of customers. Since the Pearson correlation coefficient is +0.289, the relationship between the quality of service interaction and the satisfaction of customers was positively significant.

H7: The quality of overall E-banking service has a positive correlation with the satisfaction of customers, which was supported.

The Pearson correlation coefficient between the perspective of customers and the satisfaction of customers was 0.267. The p-value is 0.001 < 0.05 so the perspective of customers is positively related to the satisfaction of customers. Since the Pearson correlation coefficient is +0.267, the relationship between the perspective of customers and customer satisfaction was positively significant.

H8: The perspective of customers is positively related to the satisfaction of customers faction, which was supported.

The Pearson correlation coefficient between the satisfaction of customers and the loyalty of customers was 0.292. The p-value is 0.000 < 0.05 so the satisfaction of customers is positively related to the loyalty of customers. Since the Pearson correlation coefficient is +0.292, the relationship between the perspective of customers and customer satisfaction was positively significant.

H9: the satisfaction of customers is positively related to the loyalty of customers, which was supported.

6 Discussion and Conclusion

We aim at investigating the effect of the quality of overall e-banking service and the perspective of customers on customer loyalty. It is not easy to balance the service of

traditional and online banking methods. The effectiveness of the online e-banking service is an important element to gain the satisfaction of customers.

The banking industry provides instant e-banking services to users online in many countries. Hong Kong banks have launched 'Payme', a mobile phone app to facilitate users for managing their own accounts. The dimensions of this project (H1–H9) can help bank managers better decisions. The correlation of responsiveness with the quality of overall e-banking service and the quality of information from the perspective of customers is relatively higher. Banks can strengthen those elements in the service. A bank manager can put more effort into the prevention of emergency accidents. Problem handling methods and compensation also need to consider for online services. The content of the website can change to a better design in order to attach the attention of users. More detailed and accurate information can be provided on an online e-banking website.

Furthermore, both the quality of overall e-banking service and the perspective of customers have a positive effect on the satisfaction of customer satisfaction. The customers' expectation of the Internet banking service is high. The online service's quality needs to match the perspective of customers for developing their trust of users. Loyalty is developed by repeated usage of the online service. Customers may introduce the service to friends or relatives. It is easy to attract existing and potential customers in order to gain a competitive advantage in the bank sector.

References

1. Frankenfield, J.: Online Banking. https://www.investopedia.com/terms/o/onlinebanking.asp. Accessed 23 Apr 2020
2. Gerber, N., Price, J.K.: Measures of Function and Health-Related Quality of Life. https://www.sciencedirect.com/topics/medicine-and-dentistry/factor-analysis. Accessed 23 Apr 2020
3. Hong Kong Monetary Authority. Risk Management of E-banking. https://www.hkma.gov.hk/media/eng/doc/key-functions/banking-stability/supervisory-policy-manual/TM-E-1.pdf. Accessed 23 Apr 2020
4. Parasuraman, A., Zeithaml, V.A., Malhotra, A.: E-S-QUAL: A Multiple-Item Scale for Assessing Electronic Service Quality. https://journals.sagepub.com/doi/10.1177/1094670504271156. Accessed 23 Apr 2020
5. Saleh, M.: Bank Selection Criteria in a Customers' Perspective. https://www.researchgate.net/publication/314834507_Bank_Selection_Criteria_in_a_Customers'_Perspective. Accessed 23 Apr 2020
6. Sathiyavany, N., Shivancy, S.: E-Banking Service Qualities, E-Customer Satisfaction, and e-Loyalty: A conceptual Model. https://www.researchgate.net/publication/326045919_E-Banking_Service_Qualities_E-Customer_Satisfaction_and_e-Loyalty_A_conceptual_Model. Accessed 23 Apr 2020
7. Tarigan, J.: User Satisfaction Using Webqual Instrument: A Research on Stock Exchange of Thailand (SET). https://media.neliti.com/media/publications/75767-EN-user-satisfaction-using-webqual-instrume.pdf. Accessed 23 Apr 2020
8. Yau, H.K., Siu, M.K., Tang, H.Y.: Effects of e-service quality dimensions on overall e-banking service quality, customer satisfaction and customer loyalty in Hong Kong banking sector. In: Lecture Notes in Engineering and Computer Science: Proceedings of the International MultiConference of Engineers and Computer Scientists 2021, 20–22 October, Hong Kong, pp. 218–221 (2021)

Application of Airships in the Surveillance Field

Emanuele Adorni[1](✉), Anastasiia Rozhok[1,2], and Roberto Revetria[1]

[1] Università degli Studi di Genova, Via Balbi 5, 16126 Genova, Italy
emanuele.adorni@gmail.com, roberto.revetria@unige.it
[2] Bauman Moscow State Technical University, 2-Ya Baumanskaya Street 5, 105005 Moscow, Russia

Abstract. From the available review on the employment of unmanned aerial vehicles (UAVs) in the surveillance field and especially in the application of these in indoor environments, we developed a comparison between heavier than air (HTA) and lighter than air (LTA) vehicles, proposing active solutions to the disadvantages and how they can be employed in the surveillance field. Both drones (HTA) and airships (LTA) have been found to be reliable means for surveillance even given the meagre literature that can be found on the internet. The definition of a model representing a real-life scenario is proposed and analysed in all the factors that have been used. This model is presented also the pattern that the vehicles should follow to minimise their environmental impact. This study is developed inside a broader analysis of how to provide the right level of security for an automated port that would rely on an autonomous security system.

Keywords: Airships · Autonomous security · Drones · Simulation · Surveillance · Unmanned aerial vehicle

1 Introduction

Considering the misleading outcome of human reaction to specific issues, nowadays, it is possible to overcome false alarms with newly developed tools. In this work, we want to review the research and applications of unmanned aerial vehicles (UAVs) for surveillance in different scenarios. During the research phase of our project, we came across the issue that the literature regarding the application of UAVs in the surveillance field is minimal. A topic on which researchers have not been focusing a lot is the application of UAVs for the security of infrastructures. UAVs have already been studied for their application for safety. These types of studies are especially regarding their employment in scenarios dealing with natural disasters, for example, reconnaissance operations after earthquakes or avalanches. According to their lifting power, UAVs can be divided into two categories: Heavier-Than-Air (HTA) and Lighter-Than-Air (LTA). Within the first category belong drones, aircraft and copters which we will describe in the first part of our article. In the second part of the paper, we will talk about airships, among which we can find blimps, which belong within the LTA category. As we mentioned, the literature about UAVs used in the security and safety field is very small, and the literature about these two subcategories is even meagre. By dividing into these two categories we could differentiate

S. Ao et al. (Eds.): IMECS 2021, *Transactions on Engineering Technologies*, pp. 40–52, 2023.
https://doi.org/10.1007/978-981-19-7138-9_4

the advantages and disadvantages of each, also considering the perspective that we have for future applications in future projects. The will of employing both drones and airships is widely spread. The main advantage that many studies highlight is the efficiency of such technology with all the implemented with sensors. The third part is dedicated to the comparison of the two surveillance methods to provide a given level of autonomous security for outdoor and indoor environments.

The steps followed to develop this research have been the analysis of the literature review of UAVs employed in the surveillance field [23] and then developing a first model simulating the possible sprint-and-drift surveillance approach. This approach is thought of as an active solution to maximise the eco-friendly surveillance potential of the airships, the means of surveillance that we understood to be most effective. When in the future it will be possible to simulate the behaviour of these vehicles also implementing the system with artificial technology (AI) and machine learning tools, it will be possible to think about the active application of this technology in frameworks such as the one that we can find in the harbour of Genoa: the surveillance of the new terminal San Giorgio (TSG).

2 Drones

Nowadays, it is possible to buy or build a drone almost effortlessly. Drones are now used as a hobby since technically they are not anything complicated. For example, using drones to take photos or videos from various angles inaccessible from the ground is becoming more and more popular every day. If we consider a drone as a way to observe a certain environment, it cannot be done without integration with technologies such as artificial intelligence and machine learning algorithms. In the literature, one can find studies on the use of UAVs for surveillance, protection and rescue, but scenarios in the area of patrolling a specific territory for security purposes are not common. The studies we studied allowed us to analyze the state of the art and identify the main problems for which we want to propose specific solutions.

Even though nowadays it is easy to buy a drone, it is necessary to take into account the legal restrictions on its use. It is necessary to understand for which purposes the use of drones is allowed and for which it is strictly forbidden by law. For example, operating drones for commercial purposes requires a special license, since urban areas are often no-fly zones and the use of drones in these areas may interfere with specific operations. The use of drones for civilian purposes (weddings, photoshoots, videotaping, etc.) may nevertheless be unsafe within the constraints of established laws. A step toward the safe use of such technology was a study of how to determine the types and intentions of such objects [2]. When tested in various scenarios with increasing recognition complexity, Bisio et al. obtained an improved software with an astounding 96% protection rate. These and other studies suggest that a detection and protection system against external agents will always be necessary for an infrastructure surveillance scenario.

The use of drones began with military infrastructure, and later drones became available to civilians. For surveillance purposes, improvements were required that would take into account the speed of response and autonomy of UAVs. For commercial purposes, drones began to be used due to their cost-effectiveness and efficiency [13]. For example,

for the transportation and delivery of small goods (small parcels, packages, light objects), UAVs that do not require too much computing power can be used, for this purpose set waypoints, which the drone follows the designated route [3]. To develop an effective surveillance system using drones it is necessary to implement it in a better cloud system based on the cognitive Internet of Things [4], in which all elements must communicate and exchange information with each other.

The delay between recognition and tracking phases is a studied problem of drones in object recognition and tracking. In an industrial environment, this concept is important for surveillance scenarios, such as identifying the location of an object in a small manufacturing environment. The drone sends a signal to return the object to its location, and then successfully determines whether the action taken was successful or not. When the drone recognizes the problem autonomously, that is, without the involvement of the control support centre, this action is the most successful. An experiment was conducted by Kim et al. [10] that yielded reliable results for their dynamic computation offloading scheme for unmanned mobile surveillance systems. In 2017, a new enhancement technology in the field of object recognition technology, YOLO9000, was proposed [14]. YOLO9000 was designed to optimize joint detection and classification and was an improved version of YOLOv2 and then improved in terms of speed to YOLOv3 [15]. It is worth noting, however, that the researchers did not consider the problem of drone power consumption. The exhaustive capabilities of the drone in the versions presented above, in our opinion, can lead to problems of drone power consumption and the need to increase battery charges. In our study [24] we have already set ourselves the problem of reducing the energy consumption of airships in the emergency lighting system, in the future our work will also be aimed at studying this problem.

In search and rescue (SAR) operations, the use of drones is an effective technology, but it requires autonomous surveillance for which very efficient algorithms using deep learning must be developed. In search and rescue operations, drones can be used to identify relief areas, scan them and provide further assistance or send a signal to locate the location. The main obstacle here is the large amount of data required for training, because drones need to recognize even small gestures, such as a wave of the hand asking for help, body movement under a structure, etc. Then, this information should be transmitted to the operator to help determine the number of victims and possibly assess their condition [11]. The problem is further amplified by the fact that the number of studies on the detection and recognition of actions on aerial photographs is currently limited, so it is necessary to develop an appropriate algorithm to scan a specific area and recognize certain actions and objects that were in an accident or disaster situation.

When considering the scenario of using drones in an industrial environment, we should consider the complexity of the industrial system to understand the resources that are needed to improve system efficiency by minimizing the human factor. For navigation inside an industrial facility, the most relevant task is to fully automate drones [21]. For indoor surveillance operations, it is necessary to use appropriate sensors such as LIDAR or LEDDAR, cameras, ultrasound and others. The main challenge of implementing UAV technology indoors is considered to be ensuring safe flight in a confined environment. When drones move in open space, the issues of their orientation and positioning are solved by the joint use of GPS and Glonass, as well as real-time kinematic systems (to

recognize obstacles and objects) [7], which allows the UAV to perform many tasks in open space. However, when operating drones indoors, it is necessary to strictly determine their position and reference points. The most convenient tool can be considered to implement it in the IoT system and equip drones with autonomy and buoyancy systems (sensors), gyroscopes (as a combination of microelectronics and mechanics), accelerometers, magnetometers, piezoelectric barometers, ultrasonic range finders, stereoscopic cameras, LEDDAR to determine the distance to the object. According to theory [8], indoor localization methods can be divided into two main categories:

- Wave characteristics and propagation through various media do not give the drone an accurate position (with errors of the order of 5–9 cm). Therefore, it is very important to avoid collision between it and any element of the covered object.
- Vision/image-based localization, which uses computer vision to display objects in a global coordinate system. However, this technology has not been well studied in a dynamic indoor environment.

In an inertial navigation system, localization is possible with an initial position and various motion sensors. However, the use of inertial motion sensors (IMU) is not advisable due to the possibility of dangerous situations such as loss of control and subsequent collision with objects.

Ibrahim and Moselhi [8] in their work proposed a solution belonging to the second category and aimed at improving the autonomy of drones. It consists of AprilTags and coordinates known in advance using 3D BIM [12]. The main advantages of the solution proposed by the authors are the cost, as well as the resources associated with the use of wireless ultra-wideband (UWB) networks, as well as vision-based positioning cameras. This solution is in many ways more efficient than existing ones, and the use of tags is very relevant, but there are disadvantages associated with the distance between the tag and the UAV for effective communication and localization during drone takeoff and landing operations.

At present, the potential of drones is limited by the capacity of the batteries used, so a system should be developed to increase their performance. As we mentioned earlier, drones are widely used in military infrastructure. Williams and Yakimenko [22] proposed a military scenario, which defines many ideas that can be implemented: hydrogen fuel cells, solar-powered drones, in-flight laser beam recharging, etc. The process of recharging drones should be automated also to save money and time, and then both battery charging and replacement can be considered. Williams and Yakimenko proposed using 3 (or 4, depending on the need) drones that would communicate and work by swapping with each other when the battery drops below a certain threshold. This is a new mission concept that would keep data loss to a minimum, and the mission would be completed with a minimal computational effort from each drone.

3 Airships

The employment of blimps can be seen as another very effective means of surveillance. When we are talking about LTA vehicles we can consider blimps, which have a semi-rigid structure, and zeppelins, which have a fully rigid structure. For this analysis, we

will talk about the more general category of vehicles: airships. Thanks to technology nowadays, airships can be equipped with the necessary technology to avoid the human presence on board. It would need only a trained pilot from a base control station. Airships are characterized by the low acoustic noise level, structure simplicity, low energy consumption, long durability in air, high payload-to-weight ratio and vertical take-off and landing ability. Airships, in general, are composed of an envelope, which will be filled with the buoyant gas, and a nacelle, where we can find the equipment needed for the operations.

Airships can be both built and purchased. A necessary step for the choice of the airship is the definition of the lifting force that you will need. The lifting force can be static, dynamic or powered. For our research, we decided that the choice of a static lifting force generated by a gas (which can be hydrogen or helium) [5] was the right approach. The choice of one of these two gases is not an easy operation. The first thought has to be the safety parameters of these. Hydrogen and me-thane have ignition properties, which can result in being an issue in certain scenarios. An example is the one we want to conduct our research on, the surveillance of a porta area. IN this type of situation, the choice of an inflammable gas can be taken only under certain conditions. Another choice for lifting gas can be hot air, which results in having a lifting force of around 3.14 N/m^3, while other gases can also exceed 10 N/m^3. Given the necessity of ensuring safety and not having any economic constraints at the moment, the hypothesis of choosing helium as lifting gas seems acceptable. Helium is commercially available but expensive with low-flammability properties. A safety measure to ensure the safety of the airship, in this case, would be first pumping hot air to check for eventual leaks, and then pump the helium into the envelope. In the case of different missions, such as transport, the choice has to be taken considering the industrial application of other materials, such as liquefied petroleum gas or other already flammable materials.

Ganesh et al. [5] offered an effective methodology for the calculation of the static lift for airships, surface area and projection area, motion, and some structural parameters. This methodology was verified with designing and simulation processes.

Kadir et al. [9] approached the motion issue differently. By considering an airship equipped with two hulls, to increase the carrying capacity of the airship, they analyzed the temperature and pressure of the environment to calculate the lifting power of the airship. This process is usually necessary for every type of airship, but especially when talking about multi-hulled vehicles. This is because the compressibility of helium, the most commonly used gas, depends on these parameters of the environment. Another problem related to the control of airships is the deviation from the course. This happens due to the influence of wind speed. Through their research, the scientists showed that the airship deviated from the course after several meters for indoor and outdoor operations, but the deviation values decreased when using closed-loop control. This brought to the idea that, given the port area scenario that we consider, we need to be extremely careful on how to retrieve real-time data of the environment and transmit them to the airship system.

An example of surveillance and safety applications of airships is given by Gorkin III et al. [6]. The researchers, to keep a healthy marine ecosystem in Australia, studied how to develop recognition software to determine the threat level (given by sharks, stingrays,

etc.) for beach goers. Through eventual smart devices, patrolling and surveillance agencies may be able to share data through smart-wearable devices and phones connected to the system. In any case, the disadvantages related to battery-powered drones were quite important (limited battery life, the need for pilot training, equipment expertise, limited flight areas per air safety regulations, etc.) resulting in being crucial for the perfect functioning of the system.

Adam et al. [1] conducted experiments based on continuous wildlife monitoring. From both surveillance but also a touristic point of view, this feature would provide authorities and visitors with information about the state of fauna and simplify the authorities' work to ensure safety on beaches or other natural environments. The system employed was represented by an airship as the aerial platform. The authors focused on the advantages of such technology over other types of aerial survey devices as means of observation of ecosystems. The advantages pointed out were: silence and remoteness, at least 8 h of observation with no need for a recharge, zero licensing, and minimal training (to deploy the vehicle without reference to the aviation authority and without needing a drone pilot). The disadvantage understood by the researchers was that, given the technology employed, it was impossible to distinguish animal species from a height of 70 m, indicating the need to improve camera parameters. In the study, the experimental part was conducted only for 70% of the established time given that, for the 30% of the prearranged time, environmental conditions (rain and wind conditions) didn't allow the deployment of the airship.

When large-scale disasters, for example, earthquakes resulting in multiple casualties, occur around the world, the detection of people within 72 h from the accident is extremely important to have high chances of saving their life. Within the framework of surveillance, airships acquire a very important and essential role given their lower-sky availability, allowing high-quality 3D images also of areas that may be difficult to be reached fast by human operators. Saiki H. [16] proposed an autonomous airship system with a robust flight control system in longitudinal motion using H-infinity control, which considers wind deflection. The experimental procedure gave a result the application of control methods and dynamics of the airship's longitudinal motion considering the trim, with a reduction of the steady-state deviation about the altitude of the airship.

The technology for detection and location of objects can be already employed on airships, allowing to use them effectively for identification and surveillance (through the employment of thermal images, plate detectors, etc.). Through much researche, it is possible to identify the reasons for which airships are better than rotor drones, offering active solutions for disadvantages such as fling issues and costs. Airships can be employed as means of surveillance which offer advantages also from a structural point of view. It must be noted that, in this aspect, drones represent a better choice when needed to perform operations in small environments [18]. For other aspects, we have already listed the advantages of airships: overcoming the battery-related problems previously highlighted for drones, for example; having the possibility of longer periods of autonomy and durability in flight; presenting systems to ensure energy efficiency; the low environmental impact when employing renewable energy systems; and the lack of large spaces for take-off and landing operations.

A necessary step for the technology is the integration of airships with AI and machine learning algorithms. The first attempt at autonomous communication between airships and computers was conducted in 2018 by Shah et al. [17]. The researchers considered detection and tracking algorithms, which could be suitable for the mission, and the image processing concept. The scientists tested three algorithms ("edge detector," "canny operator," and "thresholding") thought to be the most suitable for the mission and found out that the thresholding algorithm was giving the simplest method of image segmentation, creating a binary image of black and white pixels, differentiating itself by the canny operator and the edge detector which use probability to find the error rate, improving the signal-to-noise ratio and allowing finding objects even when in noise condition. The authors developed then an autonomous UAV to be employed in indoor environments for surveillance and monitoring, where the system had the mission of recognizing objects and displaying the results of the acquisition on a graphical user interface.

To perform indoor environments previous analysis, have to be done. For example, through fluid flow simulation programs, it is possible to determine the drag coefficient of the airship. This will be then used to determine the optimal streamlined profile of the airship, which changes given the system conditions. Van Asares et al. [19] performed a study on different airship profiles to find the most suitable shape for indoor operations. Once determined the optimal shape, the airship was equipped with automatic manoeuvring devices. Through CFD, it was possible to determine the blimp design profile, according to drag coefficient and drag forces experienced by the streamlined shape at different speeds and for three different length-to-diameter ratios. This process brought the authors to the conclusion that the throttle level at which a balance of power and weight can be achieved. The next step of the research has been the testing under different environmental conditions, for example, changing the head-winds. The results obtained were considered satisfying for the conditions of an airship hanging in the air.

Wang et al. [20] performed a study on the indoor application of airships, investigating the modelling and control of the path movements. In this case, the testing scenario was the surveillance of industrial storage management. The factors considered by all the studies related to the indoor application were the keeping the stationarity and trajectory tracking of the airships, with particular attention to the development of a way to keep a stable trajectory in the airspace to follow a given path. The process adopted to reach this goal was to apply robust control, in other words, to suppress unknown uncertainty, and to estimate the unknown uncertainty with a dynamic system. All this would be then compensated by a closed-loop controller for estimation-based control. An interesting approach has been the decision in the two directions of flight: vertical and horizontal motion. These two have to be controlled by a regulator which will have output feedback and a perturbation compensation. The idea behind this is the simplification of a complex model by dividing it into two: a kinematic model, considering North-East-Down directions, and a dynamic model, derived from the Newton-Euler equations.

The control of airships can be done by OptiTrack system. This can capture and track the airship in indoor environments, solving problems related to the position and orientation of the vehicle, which data are then transmitted to the main cluster which collects all the data. An important concept repeated more than once is the closed-loop motion of the airship. To obtain reliable results, it is possible to simulate the motion of

the airships in different software, such as Simulink or Powersim, as used by us for our simulation. In any case, the model representing the "real system" can be replaced by two interfaces: one receiving the data of the position of the airship, from the OptiTrack system, for example, which will then be decoded and from which the data of the position will be extracted; the second one collecting the data on the motor commands and sending them to the control system of the airship. These data can be collected by an XBee module, for example, in which, after receiving the packet from the host computer, the microcontroller analyses the packet and controls the motors using PWM waves with a motor driver board.

Among all the research analyzed, the work of Wang et al. resulted to be complete and valuable. Their experiments show the effectiveness of the developed controller for stabilizing a point and following a trajectory in the presence of disturbances, and the results of practical tests coincide with the results of modelling. However, it is necessary to slightly expand the research to achieve the results we need, for example, to consider movements without assumptions to reduce the pitch angle when moving, develop general motion controllers without dividing it into vertical and horizontal, and add the use of a camera to achieve autonomy.

In further studies, we will also address the issues related to the legal problems and issues of the application of airships in private spaces. We will need to consider also this aspect given the nature of our project.

4 Simulation

Given our previous study [23] we suggested the high efficiency of the combination of blimps with drones. With further research [24] we understood how, in theory, was very efficient, for real applications just the employment of airships would be good. In particular, we considered the positive characteristics of the LTA vehicles:

- Long duration of the flight, with the aim of renewable energies and power-saving patrolling approaches
- The high reachable altitudes, given the proper permissions obtainable by the authorities
- The reasonable costs for good products, both in the case of purchasing from industry or building it on your own
- The possibility to obtain licenses for flying above certain altitudes
- The possibility to have a pilot not on board, considering the airships as UAVs
- The possibility of reaching high speeds, given a proper environmental friendly propulsion system
- The presence of low risks for the surrounding environments, because of the possibility of choosing different characteristics
- The possibility of employing small storage places, in proportion with the airship's dimensions.

Given our studies, we decided to develop a simulation where, given certain constraints and initial hypotheses, we realized the most efficient path for surveilling a certain area. Some of the assumptions have been: considering one airship as a unit of a

fleet, for which a wider area can be covered; the airship with a power generating system that allows extremely long periods of surveillance; the employment of the power saving patrolling approach of sprint-and-drift where the propulsion system is activated only during the sprint phase, and during the drift phase only the wind makes up the necessary speed. In this section, we will explain the development of the model used to study the efficiency of airships in the surveillance field. Figure 1 shows how we developed the sprint-and-drift approach.

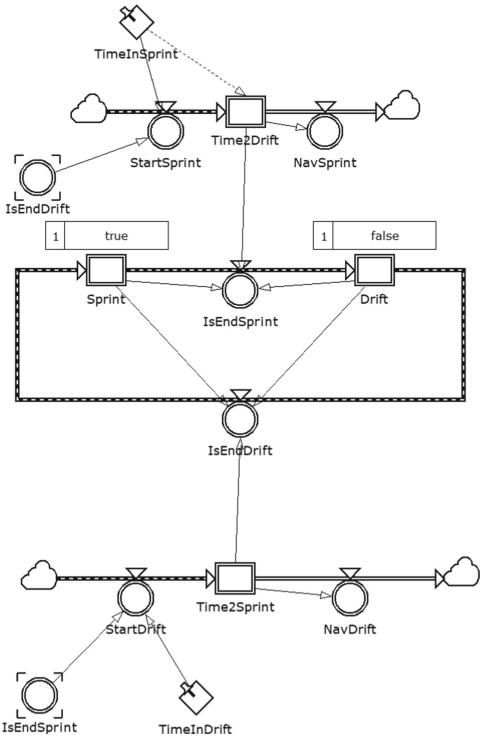


Fig. 1. Sprint-and-drift approach model for surveillance.

In particular, we started by defining two logical functions, *Sprint* and *Drift*. These two will acquire value *true* or *false* if the airship will have to sprint in a direction or drift in the wind direction. The entities *IsEndDrift* and *IsEndSprint* are auxiliary functions that have the role of switching between the two phases of the approach. *IsEndDrift* depends on *Time2Sprint* (level) and *IsEndSprint* depends on *Time2Drift* (level). These two levels are the results of the operations *StartDrift* and *StartSprint*. With these two operations, we defined a relationship between the auxiliaries *IsEndSprint* and *IsEndDrift* and the constant values *TimeInDrift* and *TimeInSprint*. While the auxiliaries have already been introduced, of great importance for the simulation is given to the two constants. These two identify how much time the airship will spend in the two phases because, when dealing with nautical orientation, the element used to calculate all the movements is

time. During the sprint phase, the airship will be propelled for a certain amount of time (*TimeInSprint*) and then during the drift phase, the airship will move along the wind direction for a fixed amount of time (*TimeInDrift*). The last two entities that we need to introduce are *NavDrift* and *NavSprint*. These two have the function of scanning the time necessary for concluding the sprint and drift phases.

In Fig. 2 it is possible to see the second part of the model. Here, the logical functions *Sprint* and *Drift* implement the actual functions that govern the surveillance path, *CCourseX* (for the movement in the West direction) and *CCourseY* (for the movement in the North direction). These two depend on the wind vector components (*CourseDriftX* and *CourseDriftY*) which are also the components of the drift direction. As we can see from the scheme, the components of the vector of the sprint are influenced by the wind. The vector of the sprint phase is influenced by the wind. The velocity vectors *CCourseX* and *CCourseY* are then integrated through the flows *UpdatingPosX* and *UpdatingPosY*. Th two flows then integrate the functions and keep as an integration constant the two elements *West* (for the vector components on the *X* axis) and *North* (for the vector component on the *Y* axis). These two integration constants are generated by the two variables *InitialX* and *InitialY*. These two represent the initial point of our surveillance operation. We can suppose that the airship will be deployed from the hangar and will sprint until the initial surveillance location, which will be a random point within a certain range (both in the West and North directions).

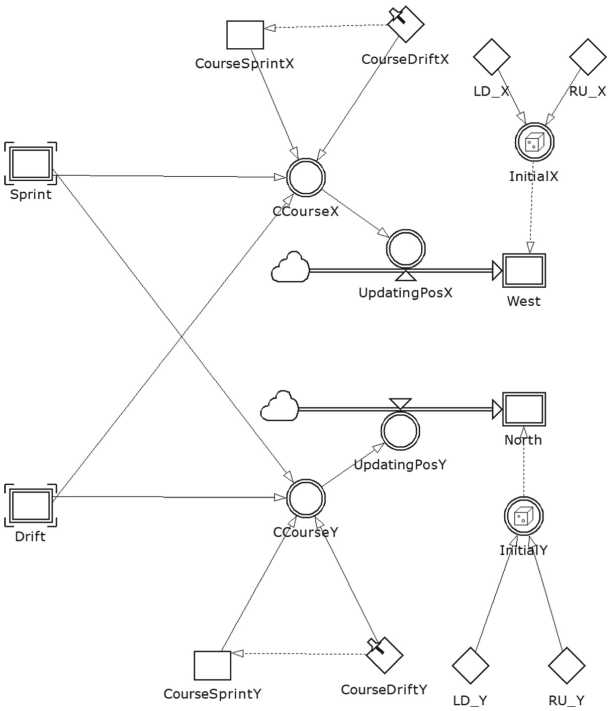


Fig. 2. Implementation of the functions describing the sprint and the drift phases.

Built then the necessary relationships, it is possible to develop a graph of a hypothetical path of surveillance (Fig. 3). In this case, we can see how, given a hypothetical wind with an inclination of 30° North, a vast area can be covered.

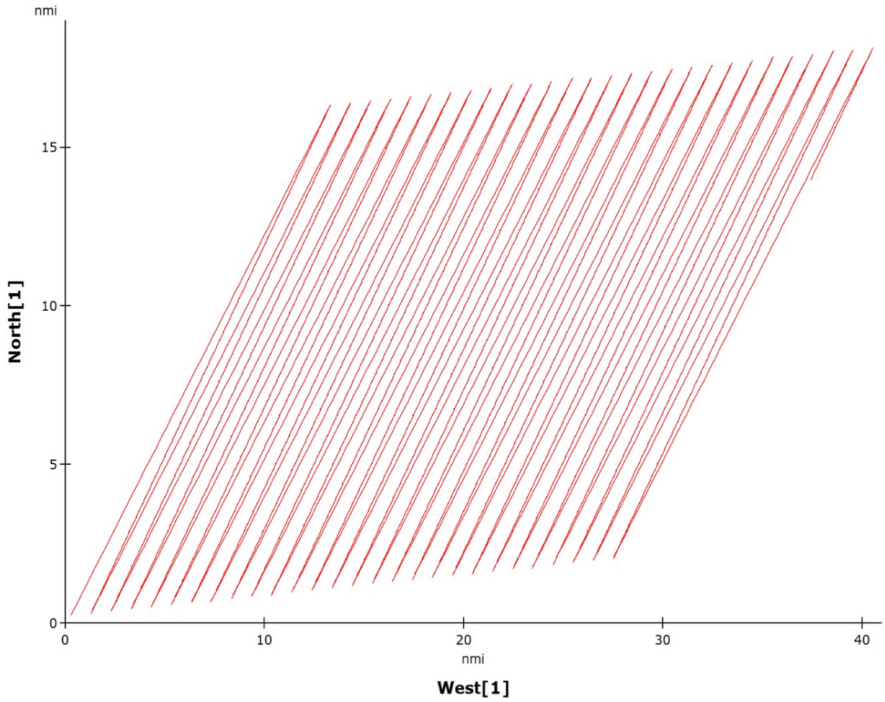


Fig. 3. Patrolling pattern starting from a base in sprint phase.

As the figure shows, the area patrolled is not filled. This inconvenience can be initially solved by employing more airships which, starting from different points, will cover a wider area.

5 Conclusion

The amount of research done on the application of HTA and LTA vehicles in the surveillance field resulted to be meagre. The first idea has been to employ a combination of drones-airship systems for the patrolling operation; after deeper research, we understood that, in reality, airships alone result be the best technology since they give the possibility of longer duration of the flight, higher reachable altitudes, lower costs both for buying and building, the low environmental impact and many other significant advantages. Starting from our previous literature review on the application of UAVs used in the surveillance field, we extended our study with the simulation model that represented the supposed ideal path described, the sprint-and-drift approach. This approach is thought to take advantage of the eco-friendly nature of airships: during the drift phase, the airship

will move in the direction of the wind using energy only for the surveillance equipment and not for the propulsion system.

We proposed a model The initial model proposed in this paper seems to resemble the first relevant step towards building a more complex model considering all the factors that would influence the surveillance operations.

6 Future Works

This article is an extended version of our previous research on the "Literature Review on Drones Used in the Surveillance Field". Given the relevance of our previous research and the knowledge acquired with this article, we would like to introduce here the topic of our future research: the possibility of employing airships not only as means of surveillance but also as technology for saving people lost at sea. In our future studies, we plan to implement the presented model with more controllable and probabilistic inputs to assess the efficiency of rescuing people lost at sea during coast patrolling operations.

References

1. Adams, K., Broad, A., Ruiz-García, D., Davis, A.R.: Continuous wildlife monitoring using blimps as an aerial platform: a case study observing marine Megafauna. Aust. Zool. **40**(3), 407–415 (2020)
2. Bisio, I., Garibotto, C., Lavagetto, F., Sciarrone, A., Zappatore, S.: Blind detection: advanced techniques for WiFi-based drone surveillance. IEEE Trans. Veh. Technol. **68**(1), 938–946 (2019)
3. Dinesh, M., Santhosh, K., Sanath, J., Akarsh, K., Manoj Gowda, K.: Development of an autonomous drone for surveillance application. Int. Res. J. Eng. Technol. (IRJET) **05**(08), 331–333 (2018)
4. Ding, G., Wu, Q., Zhang, L., Lin, Y., Tsiftsis, T., Yao, Y.-D.: An amateur drone surveillance system based on the cognitive Internet of Things. IEEE Commun. Mag. **56**(1), 29–35 (2018)
5. Ganesh, M., Kumara, S.: Design of airship for aerial surveillance and communication using knowledge based engineering. Int. J. Mech. Prod. Eng. Res. Dev. (IJMPERD) **8**(1), 17–26 (2018)
6. Gorkin, R., III., et al.: Sharkeye: real-time autonomous personal shark alerting via aerial surveillance. Drones **4**(2), 18 (2020)
7. Hell, P., Varga, P.: Drone systems for factory security and surveillance. Interdiscip. Descr. Complex Syst. – Sci. J. **17**(3A), 458–467 (2019)
8. Ibrahim, M., Moselhi, O.: Inertial measurement unit based indoor localisation for construction applications. Autom. Constr. **71**(1), 13–20 (2016)
9. Kadir, H.A., Arshad, M.R.: The development of unmanned small-size double hull blimp for low altitude surveillance system. In: 2016 IEEE International Conference on Underwater System Technology: Theory and Applications (USYS), pp. 127–132 (2016)
10. Kim, B., Min, H., Heo, J., Jung, J.: Dynamic computation offloading scheme for drone-based surveillance systems. Sensors **18**(9), 2982 (2018)
11. Mishra, B., Garg, D., Narang, P., Mishra, V.: Drone-surveillance for search and rescue in natural disaster. Comput. Commun. **156**, 1–10 (2020)
12. Nahangi, M., Heins, A., Mccabe, B., Schoellig, A.: Automated localisation of UAVs in GPS-denied indoor construction environments using fiducial markers. In: 35th International Symposium on Automation and Robotics in Construction, pp. 1–8 (2018)

13. Pető, R.: Some safety and security issues of UAVS–I. Sci. Period. Mil. Natl. Secur. Serv. **2**, 93–108 (2016)
14. Redmon, J., Farhadi, A.: YOLO9000: better, faster, stronger. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 6517–6525 (2017)
15. Redmon, J., Farhadi, A.: YOLOv3: an incremental improvemen. ArXiv (2018)
16. Saiki, H.: Longitudinal control considering trim of outdoor blimp robots for disaster surveillance. In: AIAA SciTech (2016)
17. Shah, H., et al.: Design and develop an autonomous UAV airship for indoor surveillance and monitoring applications. JOIV: Int. J. Inform. Vis. **2**(1), 1–7 (2018)
18. Surmin, A., Rozhok, A., Damiani, L., Giribone, P., Revetria, R.: Investigation about use of drone in a patrol purpose and applicability of this surveillance particularly to existing legislation. In: Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science 2018, 23–25 October, 2018, San Francisco, USA, pp. 762–765 (2018)
19. Van Asares, A., Seon Ko, P., Minlay, J.S., Sarmiento, B.R., Chua, A.: Design of an unmanned aerial vehicle blimp for indoor applications. Int. J. Mech. Eng. Robot. Res. **8**, 157–161 (2019)
20. Wang, Y., Zheng, G., Efimov, D., Perruquetti, W.: Disturbance compensation based control for an indoor blimp robot. In: 2019 International Conference on Robotics and Automation (ICRA), pp. 2040–2046 (2019)
21. Wawrla, L., Maghazei, O., Netland, T.: Applications of drones in warehouse operations. Whitepaper. ETH Zurich, D-MTEC, Chair of Production and Operations Management (2019)
22. Williams, A., Yakimenko, O.: Persistent mobile aerial surveillance platform using intelligent battery health management and drone swapping. In: 4th International Conference on Control, Automation and Robotics (ICCAR), pp. 237–246 (2018)
23. Adorni, E., Rozhok, A., Revetria, R., Ivanov, M.: Literature review on drones used in the surveillance field. Lecture Notes in Engineering and Computer Science: Proceedings of the International MultiConference of Engineers and Computer Scientists 2021, 20–22 October, 2021, Hong Kong, pp. 178–183 (2021)
24. Adorni, E., Rozhok, A., Revetria, R., Suchev, S.: Conceptual design of the emergency energy supply system for a new generation of airships. In: 2022 4th International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE), pp. 1–6 (2022)

Strengthening Strategy on Halal Certification Body Through Halal Inspection Agency

Fitra Lestari[1(✉)], Mohammad Dzaky Adzkia[1], and Irdha Mirdhayati[2]

[1] Industrial Engineering Department, Sultan Syarif Kasim State Islamic University, Pekanbaru 28293, Indonesia
fitra.lestari@uin-suska.ac.id,
11850215219@students.uin-suska.ac.id

[2] Animal Science Department, Sultan Syarif Kasim State Islamic University, Pekanbaru 28293, Indonesia

Abstract. The Indonesian government is currently concentrating its efforts on ensuring a halal product. A case study conducted in an Indonesian province shows an imbalance in the number of halal-certified and the number of SMEs that provide products for the customer. An in-depth analysis of the variables must be conducted to strengthen the halal certification body. This research aims to conduct the institutional strengthening of the halal certification body in Indonesia. Three sectors have the authority to inspect business units in Indonesia to obtain halal certification involving government, private, and collaboration. The study collected data from 6 resource persons who are specialists in their professions, such as academics, halal auditors, government, and private entities, through semi-structured interviews and focus group discussions. The analytical hierarchy process (AHP) method was used to determine strategy in decision making by considering four main criteria determining factors for strengthening the halal certification body, which consisted of barriers, drivers, organization, and competitiveness. The results of this study indicated that there were 32 sub-criteria which are the factors for institutional strengthening of the halal certification body. The government sector was chosen as a top priority in the strategy for strengthening the halal certification body. Future study is suggested to improve the performance of the Halal certification body by exploring various stakeholders and entity in the system.

Keywords: Halal certification · Institutional strengthening strategy · AHP · Government · Private · Collaboration

1 Introduction

This case study was conducted in Riau Province, which has a population of 6,394,087 million people with a Muslim majority of 87% [1]. This increases the demand for halal items, particularly among Muslims. Small and Medium Enterprises (SMEs) are critical foundations of the national economy that may boost economic growth and absorb labor. SMEs are also the types of enterprises or business organizations that connect with people regularly at different levels. The government announced on their website in November

S. Ao et al. (Eds.): IMECS 2021, *Transactions on Engineering Technologies*, pp. 53–68, 2023.
https://doi.org/10.1007/978-981-19-7138-9_5

2020 that there are at least 15,126 SME actors. It is not proportional to the large number of items that have been certified, as indicated in Fig. 1. There are 385 business actors that have carried out halal certification in 2020. On the other hand, this results in an imbalance between the number of business actors and the number of products that have carried out halal certification in Riau by 2.55%.

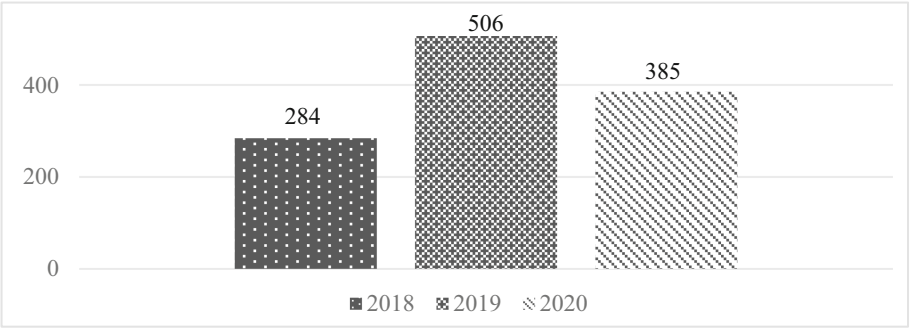


Fig. 1. Number of Halal products in Riau in 2018–2020

Halal assurance of a product is currently an issue that cannot be ignored. Given the large number of products circulating in the community, supervision and assistance related to cleanliness, halalness, and the feasibility of SME products in the food sector needs to be implemented. A study confirms that business actors must produce products that meet consumer needs and meet the standards set by the government [2]. The government has also made efforts to increase the number of certifications with various achievement parameters. The aim is to maintain the quality of imported products entering the Indonesian market and to keep local products competitive with imported products. In this case, halal certification can be done by government organizations, the private sector, or in partnership with universities. These three agencies have the power to control the current halal certification body in Indonesia. The halal certification body serves as a platform for encouraging community business units to achieve halal product certification [3]. The policy of halal certification is governed by Law No. 33 of 2014 on Halal Product Guarantee (UUJPH). A halal certificate is required for all products that enter, circulate or trade in Indonesia. According to UUJPH, the Halal Product Guarantee Agency (BPJPH) issues the formal halal certificate, which is required mandatory. The process of issuing halal certification in Indonesia is depicted in Fig. 2 [4]. There are several stakeholders involved in the process, including applicants or business actors, the government, halal supervisors or internal halal auditors, and Halal Assurance Institutions (LJH) in the form of BPJPH and MUI, Halal Inspection Agency (LPH) in the form of LPPOM and other related institutions.

The flow of the certification process above shows that a core problem can be drawn in this study. The halal certification body needs a strategy to increase the number of SMEs obtaining halal certification. A qualitative study on institutional strengthening elements for halal certification bodies is needed to address these issues. It aims to identify variables that affect SMEs undertaking halal certification and develop the optimal

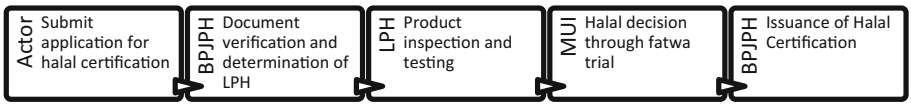


Fig. 2. Procedure for Halal certification in Indonesia

choice strategy for current issues. A study stated that the Analytical Hierarchy Process (AHP) methodology requires a decision-making method (Multiple Attribute Decision Making) to establish the strategy for a complicated problem. Multi-Objective Decision Making (MODM) and Multi-Attribute Decision Making (MADM) are the two types of MCDM. The Analytical Hierarchy Process (AHP) has been widely utilized to address problems in numerous industries such as education, industry, and engineering [5]. It is regarded as the most effective and widely used MCDM approach in various studies. AHP can be used in a variety of situations, including ranking order, prioritizing, resource allocation, and decision-making. AHP aids decision-makers by providing a numerical basis for quantifying the weight of the evaluated criteria. Each element's criterion weight is established by its relative value concerning other hierarchical elements. As a result, AHP can help decision-makers discover and prioritize important issues [6]. Decision-makers frequently employ AHP to examine the link between linguistic data [7]. Compiling a hierarchical structure, performing pairwise comparisons, computing partial weights, testing consistency, and doing synthesis are the primary processes in applying the AHP technique to solve problems [8]. As a result, this study is planned to provide policymakers with recommendations for developing methods to improve halal product certification in Indonesia.

2 Methodology

2.1 Selection of Indicators

The selection of criteria was carried out based on a review of literature studies that had been carried out by previous researchers, as stated by [9] stated in her research. She explained that there were at least 21 indicators or sub criteria with two variables or main criteria that influenced SMEs in carrying out halal certification. It is also equipped with additions from other literature studies to meet the barrier, driver, organizational, and competitiveness variables. A barrier variable is a sort of variable that causes or makes it difficult for business actors to do the halal certification. In this study, the drivers are in the form of drivers or elements that can help business actors succeed in halal certification. Organizational factors are those that are directly influenced by the company's internal management, whereas competitiveness is the polar opposite of organizational variables, arising as a result of conditions beyond the company's control. Thus, this case study found that four primary criteria and 32 sub-criteria were used to establish the optimal strategy for halal certification institutions to increase the number of halal certifications. Table 1 describes the dimensions of the criteria that are considered significant and are used as factors for strengthening the halal certification institution in increasing the number of halal certifications.

Table 1. Indicators in institutional strengthening in Halal certification body

Variable (MC)	Item (SC)	Code	Reference
Barriers (B)	Difficult licensing bureaucracy (MD distribution permit, IUMK, IUI, P-IRT)	B1	[9]
	Limited financial resources	B2	[9]
	Stages of the certification procedure that are unclear	B3	[9]
	Information is scarce regarding JPH	B4	[9]
	The raw materials for business players' products are assumed to be halal	B5	[9]
	Certification has a short validity term (2 years)	B6	[9]
	The time it takes for a halal certificate to be issued	B7	[9, 10]
	Institutional coaching and services are lacking	B8	[9, 10]
	Business actors have a low level of education	B9	[9]
	Government enforcement and affirmation are lacking	B10	[9, 10]
	Unavailability of raw materials according to SOP in certification	B11	[9]
Drivers (D)	Consumer awareness of the importance of safe and hygienic products	D1	[9]
	The increasing in the company's image and branding	D2	[9]
	Actors in the business world are aware of the regulations and Islamic law	D3	[9]
	The Halal logo adds value to the goods	D4	[9]
	Increased customer confidence and income	D5	[9]
	Halal certification is mandatory under UUJPH	D6	[9]
	Based on the UUJPH Islamic business model, halal certification is required (Halalan Thayyiban)	D7	[9]
	Halal certification as a commercial commodity	D8	[9]
	Program for halal certification	D9	[9]
	It is permissible to market things without fear of being prosecuted	D10	[9, 11]
Organizational (O)	Lack of knowledge regarding halal items among business actors	O1	[11]
	Top managerial commitment is lacking	O2	[11, 12]
	The unpreparedness of production facilities for a halal production system	O3	[12, 13]

(*continued*)

Table 1. (*continued*)

Variable (MC)	Item (SC)	Code	Reference
	Operations management capability in adopting halal services	O4	[13]
	Traceability documents are available and up to date	O5	[14]
Competitiveness (E)	Limited availability of halal-certified raw material suppliers	E1	[11]
	The shift in the lifestyle of the Indonesian people following the west	E2	[11]
	Government support to provide halal logistics services	E3	[12]
	Competitive pressure as a threat of losing the advantage	E4	[13]
	Market demand for halal products	E5	[13, 15]

2.2 Stages of Analysis

The method adopted in this study was the AHP method which is useful for decision-makers to solve complex problems by considering a supporting hierarchy model. In solving a problem using the AHP method, several steps starting from the preparation of a hierarchical structure to synthesizing the results [8].

Hierarchical Structure. The construction of the hierarchical structure is a crucial phase in the AHP method's processing; this hierarchy consists of numerous layers that typically start from a high level and work their way down, as shown in Fig. 3. In general, the AHP paradigm begins with level 1 as the primary goal, then moves on to level 2 as the primary criteria, level 3 as a sub criteria, and level 4 as alternate options.

Pairwise Comparison. Following the creation of a hierarchical structure, a pairwise comparison was performed between the indications discovered. At each matrix or level, comparisons were conducted, and nine scales were utilized in the process, as indicated in Table 2. This scale was used to quantify the relative relevance of indicators. It is crucial to consider which preference is more important than other items at the same hierarchical level or level when choosing between two indicators [8].

The calculation of the geometric mean was used in this case to determine the average group assessment of all respondents.

$$\begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \dots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{bmatrix} \tag{1}$$

$$GM = \sqrt[n]{X1 \times X2 \times \dots Xn} \tag{2}$$

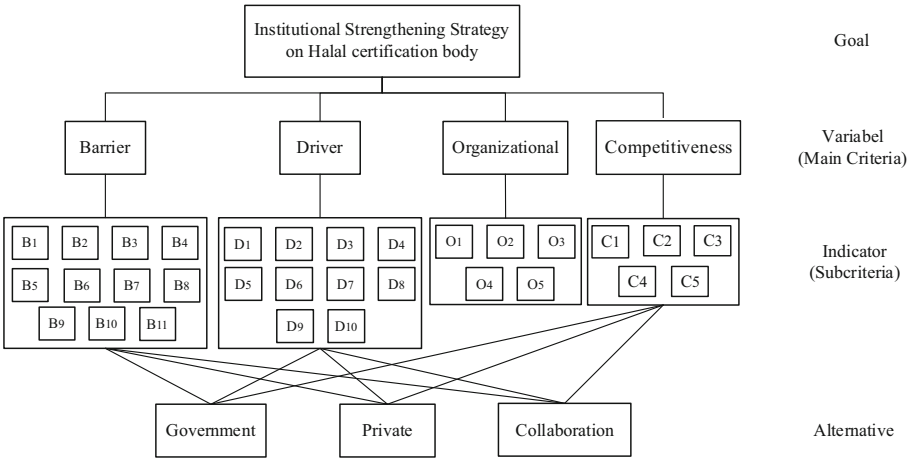


Fig. 3. The hierarchical structure of institutional strengthening at the Halal certification body

Table 2. Interest intensity scale

Importance intensity	Definition
1	Equally important
3	A little bit more important
5	More important
7	Highly important
9	Absolutely more important
2, 4, 6, 8	If doubt between two side by side values
Opposite	If element *i* has one digit above when compared to element *j,* then *j* has the opposite when compared to element *i*

The formula above is a geometric mean calculation, where "n" is the number of respondents while "X1" is the assessment of the 1st respondent in cell "*b11*" for each respondent's assessment, and "Xn" is the assessment of the nth respondent.

Partial Weight Calculation. Many methods for estimating the partial weights of comparison matrices have been devised by experts. The eigenvector and logarithm approaches, on the other hand, are appropriate methods for calculating partial weights. The founder of AHP developed the eigenvector approach which is derived from matrix theory in his research. The weight calculation is determined using this approach after normalizing the matrix by multiplying the values in each row by the number of columns in the matrix. The matrix "b" in Eq. (1) can be represented as, where "n" is the number

of rows.

$$\text{Partial weight} = \begin{bmatrix} (b_{11} + b_{12} + ... + b_{1n})/n \\ (b_{21} + b_{22} + ... + b_{2n})/n \\ (b_{n1} + b_{n2} + ... + b_{nn})/n \end{bmatrix} \tag{3}$$

Consistency Testing. Testing the consistency ratio on a matrix is very important to do. It is useful to find out whether the data being tested is more than 0.1 (CR $\leq$ 0.1). If it exceeds the limit, then the comparison between elements is inconsistent, and comparisons between elements must be re-done. The intensity of each level is calculated in the same way to calculate the priority and consistency of the ratio. In calculating the consistency ratio of a matrix, there are several elements to consider, namely consistency index (CI) and random index (RI). Equation (4)–(6) is a step in finding the consistency index value of a matrix.

$$Vector\ consistency = geometric\ mean \times partial\ weight \tag{4}$$

$$Eigen\ value\ (\lambda\ max) = \frac{\sum_{i=0}^{n} consistency\ vektor}{n} \tag{5}$$

$$CI = \frac{\lambda max - n}{n} \tag{6}$$

The CR was calculated by dividing the CI and RI values after obtaining the consistency index value of a matrix. The Random Index (RI) value utilized in this investigation is shown in Table 3.

Table 3. Random index [8]

n	1	2	3	4	5	6	7	8	9	10	11
RI	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49	1.51

Synthesis of Results. The partial weights of the main criteria matrix and the sub-criteria matrix are multiplied at this stage. The alternative priorities were synthesized using the value of the vector consistency and partial weights. The synthesis process begins at the bottom of a hierarchy and progresses to the top. Then, the total weight is added using the formula stated in Eq. (7).

$$Total\ overall\ weight = \sum weight\ of\ alternative\ criteria \tag{7}$$

It is crucial to note that the total sum of the possible weights must be 1.00. The weight value of an alternative determines its priority in developing the halal certification body. The bigger the weight value of an alternative, the more likely it is to become a priority in strengthening the halal certification body.

3 Results and Discussion

Focus group discussions and brainstorming with the LPH (Halal Inspection Agency) are the alternate technique chosen based on the findings of early observations. Halal Inspection Agency from The government, the private sector, and collaboration are three competing strategies that can be employed to solve existing problems. This alternative technique was chosen because all three options have the potential to become a halal certification body capable of regulating the Indonesian halal certification process. On the government side, the LPH is LPPOM. The private sector is defined as an entity or company that can audit SMEs and has been certified by the government as a halal guarantee institution in the second alternative plan. The last alternative is collaboration in the form of universities that can invite private parties or the government to conduct training for universities so that they can become LPHs capable of auditing SMEs. The three alternative options described previously have been summarized in Table 4 and will be evaluated for selection.

Table 4. Summary of alternative options.

Description	Government	Private	Collaboration
Experience	33 yr	2 yr	5 yr
Validity	4 yr	4 yr	4 yr
Model	A	B	C
Certified by	SNI ISO/IEC 17065:2012	ISO 17020	BNSP
Technology	Real time PCR, GC-FID, GC-MS	PCR-DNA, FTIR, UPLC	HPLC,ICP-MS
Publisher	BPJPH	BPJPH	BPJPH

Based on the AHP steps for the case studies discussed in the previous Sect. 2 and the basic model developed in this study, the results of the pairwise comparisons of the main criteria, subcriteria, and pairwise comparisons of the alternatives are presented and the results that have been synthesized are also presented.

3.1 Pairwise Comparison of Main Criteria (MC)

Barrier, driver, organizational, and competitiveness are all part of the MC matrix. The consistency ratio is verified when group assessments from respondents are included in the comparison matrix. If the consistency ratio (CR) is equal to or less than 0.10, the respondent's judgment can be accepted. The CR value for the six respondents' assessments was found to be less than 0.10, indicating that all of the respondents' assessments were consistent, and there was no need to perform pairwise comparisons again. Table 5 shows the pairwise comparison matrix where the geometric mean and matrix normalization calculations have been carried out, where the result of the CR calculation for this matrix is 0.012. The CR matrix is less than 0.10 or 10%, the matrix is considered consistent, and the calculation can be continued.

Table 5. Main criteria matrix normalization.

Main criteria	B	D	O	E	Partial weight	Consistency ratio
B	0.275	0.230	0.324	0.231	0.265	**0.012**
D	0.267	0.224	0.189	0.259	0.235	
E	0.320	0.446	0.377	0.395	0.385	
O	0.138	0.100	0.111	0.116	0.116	
Sum	1.000	1.000	1.000	1.000	1.000	

3.2 Pairwise Comparison of Subcriteria (SC)

After calculating the partial weights and checking the consistency of the ratios on the main criteria, the next step is to perform pairwise comparison calculations for the sub criteria of each indicator on the main criteria. Tables 6, 7, 8 and 9 shows the calculation of the sub criteria matrix that has been normalized and then weighted for all sub criteria.

Table 6. Normalization of the subcriteria barrier matrix

SC	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	Partial weight	CR
B1	0.06	0.03	0.09	0.09	0.05	0.23	0.04	0.05	0.05	0.06	0.07	**0.07**	**0.07**
B2	0.10	0.05	0.10	0.10	0.07	0.04	0.05	0.02	0.09	0.04	0.02	**0.06**	
B3	0.09	0.07	0.13	0.13	0.15	0.07	0.12	0.09	0.18	0.15	0.14	**0.12**	
B4	0.07	0.05	0.10	0.10	0.14	0.06	0.12	0.04	0.12	0.13	0.21	**0.10**	
B5	0.15	0.10	0.10	0.09	0.12	0.12	0.11	0.09	0.11	0.20	0.11	**0.12**	
B6	0.02	0.07	0.10	0.09	0.05	0.05	0.02	0.07	0.07	0.10	0.04	**0.06**	
B7	0.10	0.08	0.07	0.05	0.07	0.14	0.06	0.06	0.04	0.04	0.09	**0.07**	
B8	0.11	0.19	0.11	0.17	0.11	0.06	0.09	0.08	0.04	0.04	0.04	**0.09**	
B9	0.11	0.05	0.06	0.07	0.10	0.07	0.15	0.18	0.09	0.08	0.07	**0.09**	
B10	0.12	0.14	0.09	0.08	0.06	0.06	0.18	0.19	0.12	0.10	0.13	**0.12**	
B11	0.07	0.16	0.06	0.03	0.08	0.11	0.05	0.13	0.10	0.06	0.07	**0.08**	
Sum	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	

According to Tables 6, 7, 8 and 9, the partial weight and consistency ratio calculation results for the complete subcriteria matrix, in which the CR value produced does not exceed 10%, indicating that the pairwise comparison evaluation conducted by all respondents is consistent.

3.3 Pairwise Comparison of Alternatives and Synthesis of Results

Table 4 shows that three alternatives were chosen to enhance the halal certification body in this research case study. These three alternatives each feature laboratories with

Table 7. Normalization of the subcriteria matrix driver

SC	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Partial weight	CR
D1	0.14	0.11	0.14	0.09	0.10	0.15	0.12	0.17	0.19	0.21	**0.14**	**0.05**
D2	0.09	0.07	0.06	0.05	0.11	0.05	0.05	0.09	0.07	0.07	**0.07**	
D3	0.08	0.08	0.07	0.05	0.05	0.07	0.05	0.09	0.16	0.09	**0.08**	
D4	0.13	0.14	0.14	0.09	0.07	0.11	0.09	0.09	0.05	0.04	**0.09**	
D5	0.16	0.07	0.18	0.14	0.11	0.10	0.20	0.06	0.07	0.16	**0.12**	
D6	0.10	0.17	0.12	0.09	0.12	0.11	0.13	0.11	0.09	0.06	**0.11**	
D7	0.09	0.11	0.10	0.07	0.04	0.06	0.07	0.06	0.14	0.06	**0.08**	
D8	0.08	0.07	0.08	0.08	0.16	0.09	0.11	0.09	0.05	0.13	**0.09**	
D9	0.07	0.10	0.04	0.16	0.16	0.11	0.05	0.18	0.09	0.09	**0.11**	
D10	0.06	0.09	0.08	0.20	0.07	0.16	0.11	0.06	0.09	0.09	**0.10**	
Sum	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	

Table 8. Normalization of the organizational subcriteria matrix

SC	O1	O2	O3	O4	O5	Partial weight	Consistency ratio
O1	0.24	0.26	0.21	0.24	0.27	**0.24**	**0.03**
O2	0.12	0.13	0.22	0.12	0.10	**0.14**	
O3	0.21	0.11	0.18	0.18	0.26	**0.19**	
O4	0.30	0.32	0.30	0.29	0.24	**0.29**	
O5	0.13	0.19	0.09	0.17	0.14	**0.14**	
Sum	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	

Table 9. Normalization of the subcriteria competitiveness matrix

SC	E1	E2	E3	E4	E5	Partial weight	Consistency ratio
E1	0.18	0.16	0.39	0.19	0.09	**0.20**	**0.09**
E2	0.17	0.15	0.12	0.10	0.20	**0.15**	
E3	0.10	0.27	0.21	0.33	0.35	**0.25**	
E4	0.15	0.27	0.12	0.18	0.18	**0.18**	
E5	0.39	0.15	0.16	0.19	0.19	**0.22**	
Sum	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	

various testing methods. Respondents were explicitly educated about the information and distinctions between these three possibilities while filling out the questionnaire. The

three possibilities are represented by the symbols A, B, and C in this study. Table 10 shows the partial weight computation for each choice. Table 10 was used to do all pairwise comparisons, vector consistency calculations, and matrix normalization, after which the data were synthesized to provide the ranking results for each alternative using Eq. (7) from the preceding section. The total weight of each matrix is shown in Table 10. Based on the three options, agency "A" had the highest priority score of 0.423, indicating that government agencies are the top priority in the halal certification body's institutional strengthening plan, which can be seen in Fig. 4. Furthermore, in the case study in this research, agency "C," namely the university, has a priority score of 0.332, making it the second priority. Meanwhile, the agency "B," namely the private sector, receives a priority score of 0.246, making it the last alternative. The results were then subjected to sensitivity analysis to determine their stability and resilience.

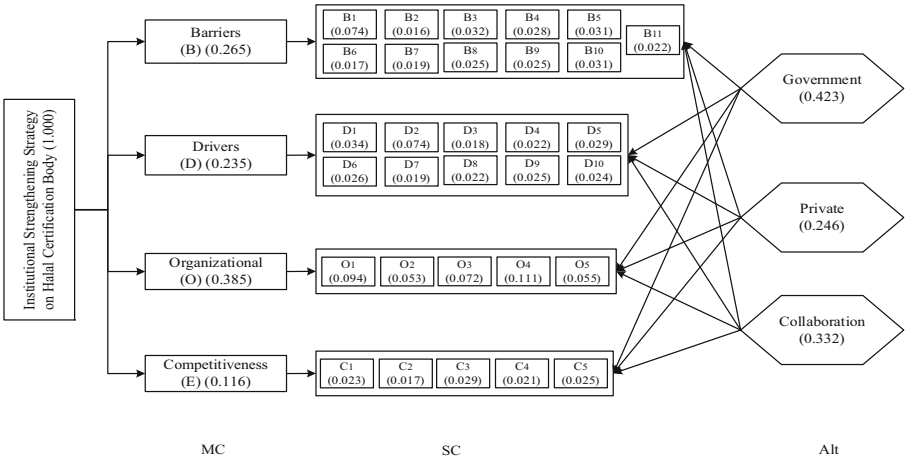


Fig. 4. Alternatives chosen

4 Stability Analysis

The AHP method was used to break down a complex and unstructured situation into multiple components in a hierarchical order, namely by providing a subjective assessment of each variable's relative importance and determining which variable has the highest priority to influence the situation's outcome. Changes in hierarchy and scoring can have a direct impact on the framework's outcomes. As a result, a stability analysis should be used to verify the robustness of an AHP architecture. The rating stability of various projection scenarios was tested. Dynamic sensitivity analysis was utilized to test rating stability by shifting the importance of the objectives. The weight of the criteria is increased or decreased in dynamic sensitivity analysis, resulting in changes in alternative priorities. Five scenarios are investigated and simulated in this research to see how these alterations affect the final ranking of options. This scenario is designed to see if adding weight to a variable has an impact on the case study's final alternative results.

Table 10. The weighting of institutional strengthening matrix for Halal certification body

MC	Partial weight (1)	SC	Partial weight (2)	The total weight (3)	Alternative partial weight			Total overall weight		
					A	B	C	A	B	C
Barrier	0.265	B1	0.074	0.020	0.56	0.20	0.24	0.011	0.004	0.005
		B2	0.061	0.016	0.50	0.22	0.28	0.008	0.004	0.005
		B3	0.120	0.032	0.35	0.36	0.30	0.011	0.011	0.009
		B4	0.104	0.028	0.44	0.24	0.31	0.012	0.007	0.009
		B5	0.118	0.031	0.55	0.21	0.24	0.017	0.007	0.007
		B6	0.063	0.017	0.42	0.37	0.22	0.007	0.006	0.004
		B7	0.073	0.019	0.40	0.20	0.40	0.008	0.004	0.008
		B8	0.093	0.025	0.27	0.30	0.44	0.007	0.007	0.011
		B9	0.094	0.025	0.40	0.33	0.26	0.010	0.008	0.007
		B10	0.115	0.031	0.22	0.25	0.53	0.007	0.008	0.016
		B11	0.084	0.022	0.47	0.20	0.33	0.011	0.004	0.007
Driver	0.235	D1	0.143	0.034	0.51	0.22	0.28	0.017	0.007	0.009
		D2	0.071	0.017	0.44	0.22	0.34	0.007	0.004	0.006
		D3	0.078	0.018	0.25	0.33	0.43	0.005	0.006	0.008
		D4	0.095	0.022	0.33	0.27	0.40	0.007	0.006	0.009
		D5	0.125	0.029	0.48	0.23	0.29	0.014	0.007	0.009
		D6	0.110	0.026	0.45	0.19	0.35	0.012	0.005	0.009
		D7	0.080	0.019	0.45	0.20	0.36	0.008	0.004	0.007
		D8	0.093	0.022	0.50	0.23	0.27	0.011	0.005	0.006
		D9	0.106	0.025	0.35	0.29	0.36	0.009	0.007	0.009
		D10	0.101	0.024	0.30	0.36	0.34	0.007	0.009	0.008
Organizational	0.385	O1	0.243	0.094	0.46	0.24	0.30	0.043	0.023	0.028
		O2	0.138	0.053	0.29	0.26	0.45	0.015	0.014	0.024
		O3	0.188	0.072	0.40	0.23	0.37	0.029	0.017	0.027
		O4	0.288	0.111	0.49	0.19	0.32	0.055	0.021	0.035
		O5	0.143	0.055	0.46	0.24	0.31	0.025	0.013	0.017
Competitiveness	0.116	E1	0.201	0.023	0.43	0.23	0.34	0.010	0.005	0.008

(*continued*)

Table 10. (*continued*)

MC	Partial weight (1)	SC	Partial weight (2)	The total weight (3)	Alternative partial weight			Total overall weight		
					A	B	C	A	B	C
		E2	0.150	0.017	0.40	0.29	0.31	0.007	0.005	0.005
		E3	0.252	0.029	0.39	0.27	0.34	0.011	0.008	0.010
		E4	0.180	0.021	0.50	0.25	0.25	0.010	0.005	0.005
		E5	0.216	0.025	0.46	0.24	0.30	0.012	0.006	0.007
Total	1.00									
Priority level								0.423	0.246	0.332
Alternative ranking								1	3	2

4.1 Scenario 1

In the scenario in the first simulation, the four main criteria are given the same weight with a weight of 25% each, as shown in Fig. 5. From the graph, it can be observed that the final ranking of the alternatives remains or does not change.

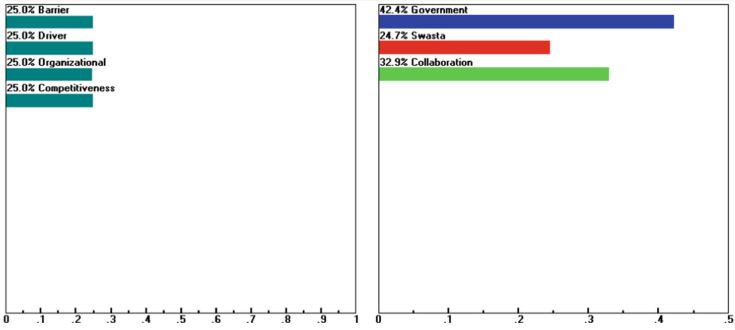


Fig. 5. Dynamic sensitivity graph for AHP framework (all criteria equal weight).

4.2 Scenario 2

In the second scenario, the main barrier criterion is given a weight of 50%, which changes the weight of the other criteria, as shown in Fig. 6. However, the ranking of the first alternative is still the same or consistent with the previous scenario.

4.3 Scenario 3

Even though the important weight for the driver criteria is set to 50% in this third case, as illustrated in Fig. 7, the ranking of alternatives remains stable.

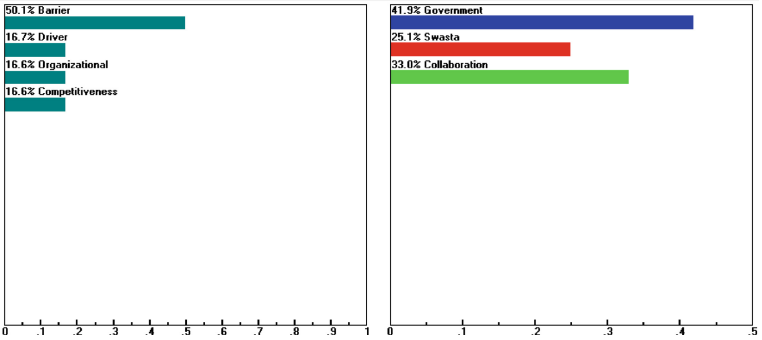


Fig. 6. Dynamic sensitivity graph for AHP framework (with 50% barrier weight).

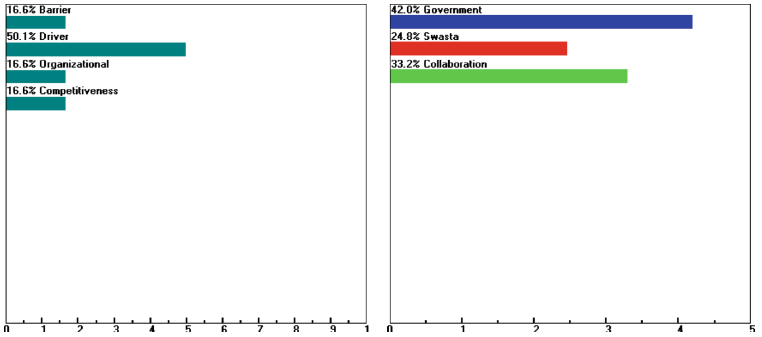


Fig. 7. Dynamic sensitivity graph for AHP framework (with 50% driver weight).

4.4 Scenario 4

In this scenario, 50% importance is given to the main organizational criteria (O). Figure 8 represents a dynamic sensitivity graph, where the graph still shows alternative "A" or government as the first alternative option.

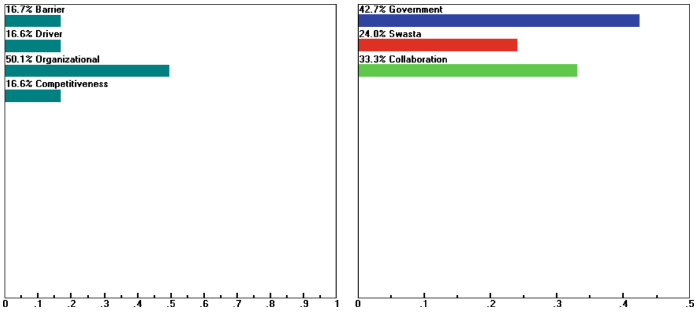


Fig. 8. Dynamic sensitivity graph for AHP framework (with 50% organizational weight).

4.5 Scenario 5

The major criteria competitiveness (E) has been given a weighting level of 50% in this scenario, and Fig. 9 depicts an alternate trend with alternative consistency "A" or government as the main ranking.

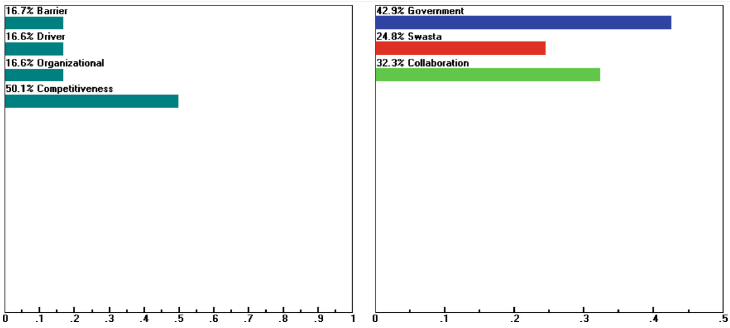


Fig. 9. Dynamic sensitivity graph for AHP framework (with 50% competitiveness weight).

Alternative "A" was chosen as the main priority in the strategy of strengthening the halal certification body with a weighting result of 0.423 or 42.3%. This is because when making pairwise comparisons with the other three variables, most of the respondents chose the interest intensity scale in the range 5–9 by considering the agency's 32-year experience in handling the halal certification process. The dynamic sensitivity graph's results are shown in Tables 5, 6, 7, 8 and 9. The proportion in each alternative can be changed by reducing or increasing the percentage of weighting on variables or major criteria, as shown in the graph of the test results. However, because it does not affect the final ranking outcomes, the AHP framework in this case study is stated to be stable, with a weighting percentage of 38.5% and the biggest weight from the cause indicator of SME for not doing halal certification, namely on the indicator of lack of willingness of the businessman on halal product and percentage weight 24.3% to overcome the problem on the indicator needs to conduct dissemination on the businessman or business unit about the importance of the halal product. Thus, the three sectors that were given the authority to carry out inspections of business units found that the government sector is believed to be able to play its role in increasing the number of halal product certifications in Indonesia.

5 Conclusion

This research aims to determine the strategy of institutional strengthening of halal certification bodies by considering different variables obtained from this study involving barriers, drivers, organization, and competitiveness. Three alternatives were selected in this case study research involving the government, private, and collaboration sectors. The government was selected as the main priority in the strategy of strengthening the halal certification body. Future study is suggested to make performance improvement

to the Halal certification body by exploring various stakeholders and entity within the system.

References

1. Central Bureau of Statistics: Population Census Riau Province in 2020. Off. News Stat. **19**(27), 1–8 (2020)
2. Fourie, L.: Customer satisfaction: a key to survival for SMEs? Probl. Perspect. Manag. **13**(3), 181–188 (2015)
3. Ab Talib, M.S., Siti, S.S., Abdul Hamid, A.B., Ai Chin, T.: Emerging halal food market: an institutional theory of halal certificate implementation. Manag. Res. Rev. **39**(9), 987–997 (2016). https://doi.org/10.1108/MRR-06-2015-0147
4. Government Regulation of the Republic of Indonesia: Government Regulation No. 31 of 2019 Concerning Halal Product Guarantee, pp. 1–60 (2019)
5. Waris, M., et al.: An application of analytic hierarchy process (AHP) for sustainable procurement of construction equipment: multicriteria-based decision framework for Malaysia. Math. Probl. Eng. **2019**, 72 (2019). https://doi.org/10.1155/2019/6391431
6. Akadiri, P.O., Olomolaiye, P.O.: Development of sustainable assessment criteria for building materials selection. Eng. Constr. Archit. Manag. **19**(6), 666–687 (2012). https://doi.org/10.1108/09699981211277568
7. Ishizaka, A., Labib, A.: Analytic hierarchy process and expert choice: benefits and limitations. OR Insight **22**(4), 201–220 (2009). https://doi.org/10.1057/ori.2009.10
8. Saaty, R.W.: The analytic hierarchy process-what it is and how it is used. Math. Model. **9**(3–5), 161–176 (1987). https://doi.org/10.1016/0270-0255(87)90473-8
9. Lestari, F., Nurainun, T., Kurniawati, Y., Adzkia, M.D.: Barriers and drivers for halal supply chain on small-medium enterprises in Indonesia. In: International MultiConference of Engineers and Computer Scientists, IMECS 2021, pp. 184–189 (2021)
10. Noordin, N., Md Noor, N.L., Hashim, M., Samicho, Z.: Value chain of halal certification system: a case of the Malaysia halal industry. In: Proceedings of European and Mediterranean Conference on Information Systems, EMCIS 2009, vol. 2009, no. 2008, pp. 1–14 (2009)
11. Prabowo, S., Rahman, A.A., Rahman, S.A., Samah, A.A.: Revealing factors hindering halal certification in East Kalimantan Indonesia. J. Islam. Mark. **6**(2), 268–291 (2015). https://doi.org/10.1108/JIMA-05-2014-0040
12. Din, R.C., Daud, S.: Critical success factors of MS1500:2009 implementation. Procedia - Soc. Behav. Sci. **121**, 96–103 (2014). https://doi.org/10.1016/j.sbspro.2014.01.1111
13. Azmi, F.R., Musa, H., Sihombing, H., Fen, F.S.: Adoption factors of halal standards: the Malaysian perspectives. In: Proceedings of the 3rd International Halal Conference (INHAC 2016) (September 2018)
14. Zulfakar, M.H., Anuar, M.M., Talib, M.S.A.: Conceptual framework on halal food supply chain integrity enhancement. Procedia - Soc. Behav. Sci. **121**, 58–67 (2014). https://doi.org/10.1016/j.sbspro.2014.01.1108
15. Shafie, S.: Halal certification: an international marketing issues and challenges. La Rev. Med. interne **23**(12), 957–959 (2002)

The Identification of Important Variables Which Affect the Satisfaction of Learner Towards Canvas

Hon Keung Yau(✉) and Kwing Nam Ng

City University of Hong Kong, Kowloon Tong, Kowloon, Hong Kong
honkyau@cityu.edu.hk, kwingnng2-c@my.cityu.edu.hk

Abstract. As technology has been blended into different aspects of our lives, the education aspect is obviously one of them. e-learning is no longer a rare technology in the learning environment. With the increasing popularity of e-learning, the way of education Hong Kong has also been influenced over the past years. Although the primary and secondary education may not seem to be influenced greatly by the technology advance yet, the tertiary education has already adopted e-learning. Therefore, this research concentrates on one university in Hong Kong. This research aims to identify the Canvas's important dimensions which affect the satisfaction of learners it. We conducted a survey to identify the thirteen variables influencing the student's perceived satisfaction. The results can show the factors that actually influence that satisfaction toward canvas system, therefore also show that how the institution could improve the canvas system.

Keywords: Canvas · E-learning · Learner satisfaction · Self-efficacy · Student anxiety · Student dimension

1 Introduction

In recent years, technologies have been evolving in an incredibly fast speed. With the fast-paced changing technologies, our lives are able to be more and more convenient. The application of technologies is becoming more and more common in different fields of our lives, therefore there is no exception regarding to the field of education. In our early stage of life, which is about twenty years, education simply takes up the whole part of it. As a result, the efficiency of learning is of paramount importance for students as it would make a huge difference in our early life. The application of technologies such as applying educational software is first applied in the tertiary education in Hong Kong. One study [1] has studied the canvas application in Hong Kong's university. As the intention of applying technologies in education is to increase the efficiency of student learning, the critical factors influencing learner satisfaction are crucial for successful e-learning. This project aims to examine the users' perception and experience towards e-learning of university students and analyzing the data collected, by dividing the investigation into different dimensions including learner, instructor, course, technology, design and environment dimension.

S. Ao et al. (Eds.): IMECS 2021, *Transactions on Engineering Technologies*, pp. 69–76, 2023.
https://doi.org/10.1007/978-981-19-7138-9_6

2 Hypotheses Development

Fifteen hypotheses were developed as shown as follows:

2.1 Gender Difference

Researches have shown that gender difference exists when it comes to learning with technology [2]. Studies also have shown that the males' 'computer self-efficacy', 'perceived usefulness', 'perceived ease of use', and 'behavioral intention to use' e-learning are better than females [3]. As a result, females tend to lose interests in using computers or technology as their 'self-efficacy' or 'perceived ease of use' to computers are lower, thus leading to less computer usage.

Hypothesis 1: Gender difference exists on computer usage habit

Hypothesis 2: Gender difference exists on student dimension

2.2 Student Dimension

For the student dimension, we first take the student attitude towards computers into account. Multiple researches showed that student attitude towards computers is important for the e-learning satisfaction [4–7]. Student attitude means that students' impression of learning through the usage of computer. The more aggressive of using IT is claimed that would lead to a better satisfaction and effectiveness towards e-learning [7].

Hypothesis 3: Student attitude towards computers will positively influence perceived canvas satisfaction

In addition to student attitude, student anxiety towards computer is also claimed to affect the e-learning satisfaction by students [7]. Computers anxiety means that the negative emotional feeling such failure of operating computers effectively [8]. Therefore, we suppose that the higher computer anxiety will lead to lower e-learning satisfaction.

Hypothesis 4: Student anxiety towards computers will negatively influence perceived canvas satisfaction

In addition, we suppose that the students' self-efficacy of computers is also an important factor affecting e-learning satisfaction. The self-efficacy of computers means that the evaluation on his own possibility success before performing a task using computer [9]. It is believed that student self-efficacy towards computers will positively influence perceived canvas satisfaction.

Hypothesis 5: Student self-efficacy towards computers will positively influence perceived canvas satisfaction

2.3 Instructor Dimension

The second dimension for my study is the instructor dimension. The first variable of the dimension is the instructor response timeliness. Instructor response timeliness means that if a student perceived instructor response to his enquiry quickly. A study shows that

Instructor's timely response influences student's e-learning satisfaction [5, 10], as the timely responses will result in encouragement to the student's learning.

Hypothesis 6: There is positive association between Instructor response and perceived canvas satisfaction

The second variable of the instructor dimension is the attitude of instructor towards e-learning. It is said that individual's perception is affected by other people such as other group members or supervisors attitude towards e-learning [11]. As one's behaviour is believed to be built on observing others' actions. Therefore, we suppose that the instructor attitude towards e-learning will affect the student's satisfaction.

Hypothesis 7: Instructor attitude towards e-learning will positively influence perceived canvas satisfaction

2.4 Course Dimension

The third dimension of my study is the course dimension. The first variable of the dimension is the course flexibility. The meaning of course flexibility is the elimination of physical barriers, hence allowing more dynamic learning. A research has shown that, by having a better course flexibility in time, location and methods, there are better participation and satisfaction of e-learning from students [5, 12]. Therefore, we suppose that the course flexibility will affect the student's satisfaction.

Hypothesis 8: There is positive association between course flexibility and perceived canvas satisfaction

The second variable of the dimension of course is the quality of course. With high quality courses in e-learning, the students may be able to construct high-level thinking models and develop conceptual theory with special learning model provided by IT, which is the interactive communications and media presentation [13]. When the students are able to do so, their effectiveness of learning is believed to increase, and so they are motivated to keep on using e-learning resulting in a higher satisfaction.

Hypothesis 9: There is positive association between course quality and perceived canvas satisfaction

2.5 Technology Dimension

The fourth dimension is the technology dimension. The first variable is the technology quality. Studies have shown that technology quality affect the e-learning satisfaction significantly [7, 14]. If the software used in e-learning process is user-friendly, the users are believed to be more willing to adopt such softwares and thus adopt e-learning. Therefore, the satisfaction of e-learning will be increased with better technology quality.

Hypothesis 10: There is positive association between technology quality and perceived canvas satisfaction

The second variable is the internet quality. Research also shows that as e-learning should allow more participation of discussion and learning by employing other computer technology such as video conferencing which may require good internet quality.

As lacking of internet quality may cause technical problems during those e-learning activities.

Hypothesis 11: There is positive association between Internet quality and perceived canvas satisfaction

2.6 Design Dimension

The fifth dimension is the design dimension. The first variable in this dimension is the perceived usefulness of the e-learning. The technology acceptance model (TAM) was developed in previous research [15], the perceived usefulness is defined as how much work can be improved after a system is adopted. If the perceived usefulness increases, the e-learning satisfaction would also increase as the students would be motivated to continue using e-learning when they see results.

Hypothesis 12: Students perceived usefulness will positively influence perceived canvas satisfaction

The second variable is the students' perceived ease of use. The technology acceptance model (TAM) was developed by previous research [15], the perceived ease of use is the ease of adopting new systems. When the ease of use increases, the perceived satisfaction is also believed to increase.

Hypothesis 13: Positive association between students perceived ease of use and perceived canvas satisfaction should be found

2.7 Environmental Dimension

The sixth dimension is the environmental dimension. The first variable of this dimension is the diversity in assessment. The diversity in assessment would affect the students' satisfaction as the diversity in assessment would give a feeling to the students that their efforts put are properly assessed [10].

Hypothesis 14: There is positive association between diversity in assessment and perceived canvas satisfaction

The second variable is the students' perceived interaction with others. More interaction between students could facilitate learning satisfaction [12]. As through the interaction, students could help to solve each other's problems.

Hypothesis 15: There is positive association between students' perceived interaction and perceived canvas satisfaction

3 Research Methodology

The survey was used in this study. 165 questionnaires in total were collected. For the gender of the 165 participants, 60.6% of them are males while 39.4% of them are females. For the Age range of the 165 participants, 93.9% of them aged 18–22, 3.0% of them aged 23–29, 3.0% of them aged >29. For the Grade of the 165 participants, 6.1% of them

were Year 2 students, 24.2% of them were Year 3 students, 66.7% of them were Year 4 students and 3% of them were Postgraduate students. For the College of the participants, 24.2% of the students majored in the College of Business, 21.2% of the students majored in the Liberal Arts and Social Sciences Department, 48.5% of the students majored in the College of Science and Engineering, 3.0% of the students majored in the School of Creative Media and 3.0% of students majored in the school of law.

The acceptable value of the factor loading is considered as 0.3 [16]. According to the Table 1, the factor loading for all 24 items from the survey are greater than 0.3, therefore, the scale is valid and all items has a large variance for the factor loading.

Table 2 shows the reliability analysis of six variables, The Cronbach's alphas of six dimensions are greater than 0.7 and they can considered as reliable [16].

Table 1. Factor loading

Item	Factor loading
SD1	0.851
SD2	0.682
SD3	0.750
SD4	0.582
SD5	0.828
SD6	0.553
ID1	0.699
ID2	0.762
ID3	0.796
CD1	0.630
CD2	0.712
CD3	0.709
TD1	0.636
TD2	0.744
TD3	0.791
TD4	0.661
DD1	0.783
DD2	0.707
DD3	0.632
DD4	0.695
ED1	0.882
ED2	0.720
SAT1	0.853

Table 2. Reliability analysis

Item	Cronbach's alpha
Student dimension	0.780
Instructor dimension	0.701
Course dimension	0.757
Technology dimension	0.763
Design dimension	0.748
Environmental dimension	0.803

4 Results and Discussion

The value of t-test for H1 and H2 is ($t = -1.580$, $p = 0.116$). Therefore, the result indicates that male in the student dimension is not significantly difference from that of female. Therefore, H1 and H2 are rejected.

The relationship between student attitude towards computer and canvas satisfaction is 0.612. Therefore, H3 is supported.

The relationship between student computer anxiety and canvas satisfaction is − 0.506. Therefore, H4 is supported.

The relationship of student internet self-efficacy and canvas satisfaction is 0.61. Therefore, H5 is accepted.

The relationship between instructor response timeliness and canvas satisfaction is 0.639. Therefore, H6 is accepted.

The relationship between instructor attitude toward canvas and canvas satisfaction is 0.655. Therefore, H7 is accepted.

The relationship between course flexibility toward canvas and canvas satisfaction is 0.701. Therefore, H8 is accepted.

The association between course quality toward canvas and canvas satisfaction is 0.559. Therefore, H9 is accepted.

The association between technology quality and canvas satisfaction is 0.566. Therefore, H10 is supported.

The association between internet quality and canvas satisfaction is 0.507. Therefore, H11 is supported.

The association between perceived usefulness and canvas satisfaction is 0.611. Therefore, H12 is supported.

The correlation between perceived ease of use and canvas satisfaction is 0.501. Therefore, H13 is supported.

The relationship between diversity in assessment and canvas satisfaction is 0.672. Therefore, H14 is accepted.

The association between student perceived interaction with others and canvas satisfaction is 0.503. Therefore, H15 is accepted.

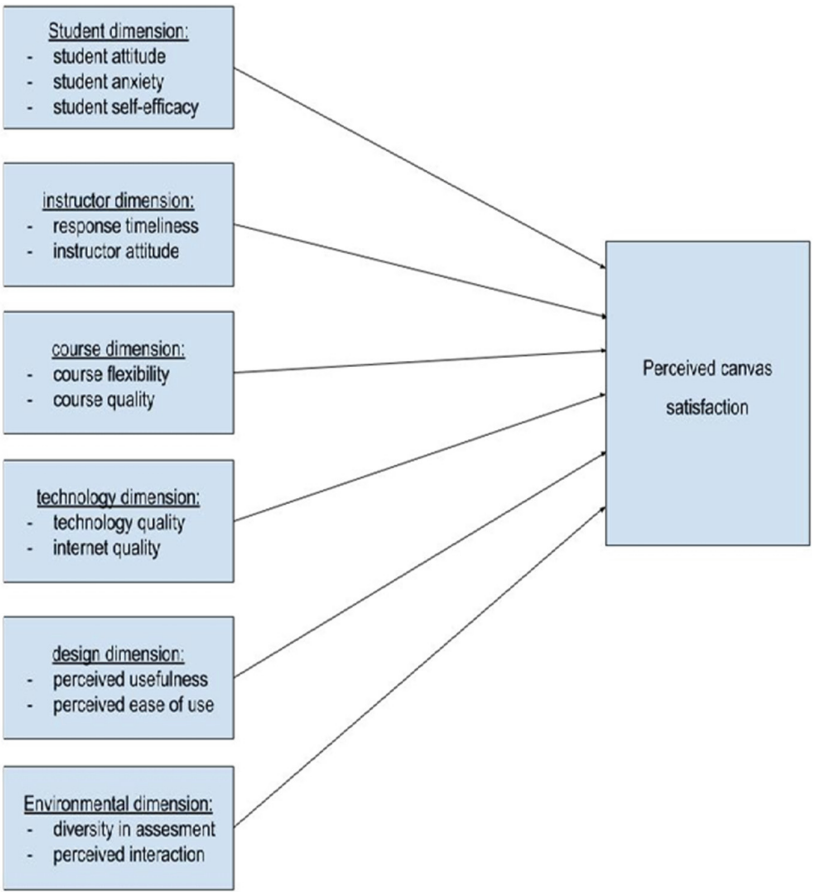


Fig. 1. The theory framework

The main achievement of this research is to finalize the theory model as shown in Fig. 1.

5 Conclusions

The difference between e-learning and face-to-face learning in classroom has been found as it involves internet, and making it a lot of complicated. Therefore, when we investigated the important variables of perceived satisfaction of the canvas system, an integrated model developed from previous study was used this previous model consists of 13 variables from six dimensions. The difference between male and female to the student dimension was examined, as people usually stereotype male as dominating sex with computer usage. Finally, the theory framework (Fig. 1) has been developed and this framework is useful for the university administration and professors to develop their courses via Canvas in near future.

References

1. Yau, H.K., Wu, L.C., Tang, H.Y.: Analysis of acceptance of canvas e-learning system. In: Lecturer Notes in Engineering and Computer Science: Proceedings of the International MultiConference of Engineers and Computer Scientists 2021, 20–22 October, Hong Kong, pp. 168–171 (2021)
2. Kahveci, M.: Students' perceptions to use technology for learning: Measurement integrity of the modified Fennema-Sherman attitudes scales. TOJET: Turkish Online J. Educ. Technol. **9**(1), 185–201 (2010)
3. Ong, C.S., Lai, J.Y.: Gender differences in perceptions and relationships among dominants of e-learning acceptance. Comput. Hum. Behav. **22**(5), 816–829 (2006)
4. Arbaugh, J.B., Duray, R.: Technological and structural characteristics, student learning and satisfaction with web-based courses – an exploratory study of two on-line MBA programs. Manag. Learn. **33**(3), 331–347 (2002)
5. Arbaugh, J.B.: Managing the on-line classroom: a study of technological and behavioral characteristics of web-based MBA courses. J. High Technol. Manag. Res. **13**, 203–223 (2002)
6. Hong, K.S.: Relationships between students' and instructional variables with satisfaction and learning from a web-based course. Internet High. Educ. **5**, 267–281 (2002)
7. Piccoli, G., Ahmad, R., Ives, B.: Web-based virtual learning environments: a research framework and a preliminary assessment of effectiveness in basic IT skill training. MIS Q. **25**(4), 401–426 (2001)
8. Barbeite, F.G., Weiss, E.M.: Computer self-efficacy and anxiety scales for an Internet sample: testing measurement equivalence of existing measures and development of new scales. Comput. Hum. Behav. **20**, 1–15 (2004)
9. Marakas, G.M., Yi, M.Y., Johnson, R.D.: The multilevel and multifaceted character of computer self-efficacy: toward clarification of the construct and an integrative framework for research. Inf. Syst. Res. **9**(2), 126–163 (1998)
10. Thurmond, V.A., Wambach, K., Connors, H.R.: Evaluation of student satisfaction: determining the impact of a web-based environment by controlling for student characteristics. Am. J. Dist. Educ. **16**(3), 169–189 (2002)
11. Fulk, J., Schmitz, J., Steinfield, C.W.: A social influence model of technology use. In: Fulk, J., Steinfield, C. (eds.) Organizations and Communication Technology, pp. 117–141. Sage, Newbury Park (1990)
12. Arbaugh, J.B.: Virtual classroom characteristics and student satisfaction with Internet-based MBA courses. J. Manag. Educ. **24**(1), 32–54 (2000)
13. Leidner, D.E., Jarvenpaa, S.L.: The use of information technology to enhance management school education: a theoretical view. MIS Q. **19**, 265–291 (1995)
14. Webster, J., Hackley, P.: Teaching effectiveness in technology-mediated distance learning. Acad. Manag. J. **40**(6), 1282–1309 (1997)
15. Davis, F.D.: Perceived usefulness, perceived ease of use and user acceptance of information technology. MIS Q. **13**(3), 319–340 (1989)
16. Gerber, N., Price, J.K.: Measures of Function and Health-Related Quality of Life. https://www.sciencedirect.com/topics/medicine-and-dentistry/factor-analysis. Accessed 12 May 2020

Survey Categorizing Paper on Education Question Answering Systems

Teotino G. Soares[1](✉), Azhari Azhari[2], and Nur Rokhman[2]

[1] Computer Science Department, Dili Institute of Technology, Dili, Timor-Leste
tyosoares@gmail.com
[2] Computer Science and Electronics Department, Universitas Gadjah Mada, Yogyakarta, Indonesia
{arisn,nurrokhman}@ugm.ac.id

Abstract. Question Answer (QA) is domain study that has related to other science, especially in the fields of learning. This is due to the increasing need for users to complete various tasks, the increasing number of teaching-learning activities that are currently, and increasing attention to efficient computing for creating the information it can be provided to real-time based users necessary. So as natural languages processing (NLP) is an alternative because have power to manipulate, replicate for complex requests and can combine with a variety of other methods, including sharing data automatically by helping to obtain language support for humans. This QA categorization survey provides universal reviews on the teaching-learning question answering systems that concentrate on four (4) fields, science of information & communication technology, science of social, science of linguistics, and science of biomedical, as well as reviewing research results, in accuracy, precision, recall and F1 score.

Keywords: Education · Question answering · Natural languages processing · Information retrieval

1 Introduction

The evolution of technology information is very influential on human interaction with computers, especially in the formal field very quickly, this occurs in the need for users to finish tasks of teaching-learning, if found in the digital world, thus affect efficient information to produce great information, so that it can help user needs. This case requires a large amount of heterogeneous data integration, to communicate with each other in order to obtain accurate information for users, this is natural languages processing (NLP) as an alternative because have expertise to process and presenting complex query, thus leading to the generation of question-answer (QA) from mapping various information based on identification and rejoinder [1], besides that QA is collaborations studies from various disciplines [6], but interrelated namely, information retrieval (IR), natural languages processing [NLP], information extraction (IE), and [2]. This is invite the researcher's attention, because the next stage is a process to obtain information or automated answers

S. Ao et al. (Eds.): IMECS 2021, *Transactions on Engineering Technologies*, pp. 77–91, 2023.
https://doi.org/10.1007/978-981-19-7138-9_7

to answer questions posed by humans in natural language and receive short answers that can improve the learning process in the digital world.

Besides that, survey study related with QAS have been performed by [2] using an metrics of evaluation to make different results from the earlier investigation with methods and procedures used in improving QAS, in order to help future researchers to select good models and procedures for improving the new system, and [3] make a survey as a development from [2], but [3] using evaluation metrics accompanied statistical analysis of system performance based on question types, methods, algorithms, and NLP procedures.

At the time [4] in his survey used metrics evaluation on the question answering systems in the biology science equating some system related to life such as biomedicine, but this survey was not devoted to answering questions in the biological field, so it became a new object to study for the next. In the future, [5] previous survey analysis in the QAS field also using metrics of evaluation and analyzing system implementation approach, algorithm used in the system, procedures, characteristics, also limitations of each research beforehand in the QAS field, and [6] Conducting surveys QAS which focuses on teaching and learning by presenting the methods, results, and limitations of each survey and research.

Four survey mentioned that above [2] to [5] using same model of evaluation survey and survey [4] equipped with analysis statistical and also survey [6] stated the methods, results, and limitations of each survey and research. However, all five surveys have the same goal to help share knowledge with future researchers to identify the procedures, characteristics, methods, and limitations of previous researchers, so it is necessary to meet the shortcomings of each survey, in order to obtain a better system performance in the future, although there are limitations because it has no combined all previous study, the scope is only the leading scientific works in the education QAS domain, as well as a wide area coverage so that there is not enough concentration to explain the categorizing of education QAS papers.

The aims of this survey categorization papers is to share knowledge the categorization of education QAS domain, namely technology information & communication science, science of social, linguistics science, and biomedical science with related previous study survey so that can be used as guidelines for further research in the future.

Complete the survey categorization of this article, is divide into several parts, namely Part 1 describe several problems related to the previous study as a reference in completing this study. Part 2 describes the settlement mechanism in this study. Part 3 categorizes education QAS domain. Part 4 describes the conclusions from the results of the review along with follow-up activity as a follow-up to the result of this study.

2 Research Methods

The process for completing this paper requires several procedure and criteria as guidelines to complete it, because education QAS is spread various journal and various disciplines, both formal and non-formal education. The survey research flowchart diagram can be seen in Fig. 1.

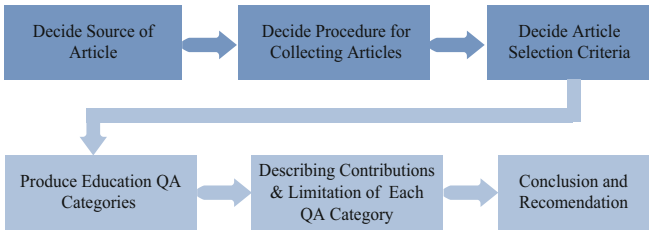


Fig. 1. Categorization study survey stage, is a process for the preparation of this paper, starting with the selection of article sources, selection of procedures and criteria for article discovery, categorizing, then analyzing contributions and limitations, until producing conclusions that become references for the next study, the details will be present in the next.

A. Articles Source

Sources of education QAS papers are obtained from proceedings, journals, and conference papers, lecturer note that can be accessed from leading sources, namely springer link online, IEEE Explorer, science direct (Elsevier), association for computing machines (ACM), Iaeng publication, Mendeley library and Association for computational languages (ACL).

B. Procedure for Articles Collection

Utilize the electronic catalogue mentioned above to sort articles in the last eleven (11) years, begin from 2011–2021 by input papers keywords in the link of electronic library text boxes searching like as question answering systems, education question answering systems. Result of the searching are displayed in the electronic library by showing the title of the paper, abstract, book chapter, conferences papers, and journal.

C. Article Selection

There are three (3) parameters used to determinate and accept the categorization of QA papers in the domain of education for further review. Papers will be eliminated if they don't meet the parameters to decide:

- The compulsory paper discussing research topic trends in the Education QAS domain. To meet these parameters, research must be uploading and reviewed between 2011 to 2021. The last eleven (11) year period may be considered for inclusion.
- Articles will be eliminated if obtained from unreliable sources, such as workshop description, theses, dissertation, articles conference, proceeding, and book chapters.
- Be sure to include only peer-reviewed journal articles, which have a high level of research quality, so that are often used by research expert to obtain good result.

3 Categorizing Education Question Answering System

The categorization of the education QAS paper in this survey is divided into four (4) sections, namely categories for teaching-learning domains in technology information &

communication science, social sciences, language sciences, and biomedical sciences. Each category of paper will be discussed in the following sub-headings.

A. Education Question Answering System in ICT Science

Research [7] use ontology knowledge to develop a smart systems in responding to recommendations to students to get answers easily, so that it can help increase student interest in online QA, the system can work well because the database that was built is able to provide answers automatically to questions asked without tutors having to be online to provide answers to the questions asked. However, the limitation of this study is the focus on non-formal education for Microsoft office application courses as well as other IF software modules. Furthermore, [8] used IR, information clustering, Relation Extraction (RE), and Named Entity Recognition (NER) to create an Ipedagogy system to answer questions on website-based data grouping, the results of the mean reciprocal rank test reached 0.73. The results obtained are different from other electronic learning applications but have limitations that are needed to improve the validity of data extraction and optimize data cluster analysis from website record search results using a better algorithm. Quite different from the research conducted by [7 In addition, [9] using vector space model (VSM), feature selection, Chinese segmentation, feature weighting to develop QAS on the e-learning support system, to share correct answers and return them to students for the questions are asked quickly so that they reach a precision of 82.34%, recall 8l.21%, and F1 75.73%. But needed to improve system performance with an appropriate algorithm or models to be applied to other fields. And [10] utilizes text mining and knowledge base to build smart QAS to respond to questions about the ICT domain, the result of the implementation shows that the system has have perform to find the best answers from all candidates, it is important for user's request. However these studies are only for junior high school students focused on the ICT domain, not only that the system needs to be improved in answer retrieval, speech recognition and knowledge base Next [11] utilize knowledge-based, Cosine Similarity, TF-IDF, Cuckoo Search Optimization, and explicit semantic analysis to build an intelligent QAS based mobile for education domain, can increase results accuracy of 79.2%, however, research has have problems in examining single word and multi words questions and concentrates on the technology information domain.

Researches by [12] create a system for QA in Portuguese to discussion software engineering materials, with improved (Seq2Seq) algorithms, and dynamic memory network (DNM+) and also using corpus of software engineering, with the assessment result 77% accuracy. But, needed to improve the system performance with the others algorithm and methods to response questions complex and also develop a corpus for others subjects, not only that[13] build a QAS for software engineering courses, using framework Context-Aware Question & Answer System (CAQAS), the experiment results show that can resolve for students, which could increase students' participation in multi cluster studies, however, has the limitation that doesn't have the means to track answers to the next question, and also concentrate on the subject of software engineering, and [14] develop a smart QAS network model, whose main module is designed based on the characteristics and necessity of the Massive Open Online Course (MOOC) to conduct QAS for computer network courses, however, this study centralize on computer networking and also need improve voice communication feature to achieve efficiency for accurate answers,

and the next [15] utilize the Bidirectional Gate Recurrent Unit (BiGRU)-Attention and Gated Recurrent Unit Convolutional Neural Network (GRU-CNN) to building biological QAS for junior high school, result of test return the model validity of datasets biological reached GRU-CNN 73% and BiGRU-Attention 74%, needed other increase performance of system, including study [16] combines Decision Tree (DT), Naive Bayes Multinomial (NBM), Logistic Regression (LG), Boosting Method, and Support Vector Machines (SVM) for automation of learning question quality in the field of computer science as well as to sharing facts about the facilitating of feasibility the large-scale online automated question and answer in Turkish junior high school, particularly on geography subject matter, but it is necessary to increase the accuracy of automatic classification with a more complex classification method.

All investigation mentioned above, known that separately study has had limitation, however all of them are committed in the same field, that is information and communication technology teaching-learning area. To comprehend better the comparison with the researches have been described above will be shown in Table 1.

Table 1. Comparing education QAS in ICT field

Ref.	Description	Methods/technics	Result and improvement	Limitation
[7]	Research articles on the education QAS, specifically in technology information and communication field with various techniques and methods, help students in process of teaching-learning with various subjects in science of technology information and communication	Information Retrieval (IR), Knowledge Ontology, Information Clustering	Can respond to problems automatically without the tutor has to be online then students can get answers to the questions asked	Learning software such as Microsoft office and other IT modules
[8]		Means Reciprocal Rank (MRR), Named Entity Recognition, Relation Extraction	MRR test results obtained 0.73	Combine more precise algorithms to improve validity of data extraction also optimize information cluster analysis
[9]		Feature Weighting, Feature Selection, Vector Space Model, Chinese Segmentation	Matrix evaluation result, precision 82.34%, recall 81.21, and F1 75.73%	To improve performance systems, needs others algorithm to applied in other domains
[10]		Knowledge Base Text Mining	Result of implementation shows that QAS can shows the best answer to the questioner so he systems performance is expected	Focus on ICT for junior high school students, in addition to improving system performance, adding voice recognition features to help people with disabilities in the teaching and learning process

(*continued*)

Table 1. (*continued*)

Ref.	Description	Methods/technics	Result and improvement	Limitation
[11]		TF-IDF, Cosine Similarity, Knowledge-based, Explicit Semantic analysis, also Cuckoo Search Optimization	The results of the assessment reached 79.2%	Limitations on the reading question deeply and long clause inquiries
[12]		sequence to sequence (Seq2Seq), Improved dynamic Memory Network (DNM+)	Satisfactory results reached 77%	Focus on software engineering subjects in Portuguese, and improved performance of algorithm respond to complex questions also add a corpus for other subjects
[13]		Context-aware Question & answer System (CAQAS) Framework	The implementation result shows that can minimize student's problems, so that increase their participation in of the teaching-learning process	Concentrate on the subjects if software engineering and answer follow-up questions unanswered because they don't have the means to track them
[14]		Massive Open Online Course (MOOC)	Built model of network QA, include design main module based on the characteristics of MOOC	Focus on computer network subjects, including features of voice communication to be improve accurate of the answer
[15]		Bidirectional Gate Recurrent Unit-Attention, Gated Recurrent Unit Convolutional Neural Network	Results of the experiment for a model of validity on biological datasets for junior high school students, reaching for GRU-CNN 73% and BiGRU-Attention 74%	Focus on Biology lessons for junior high school, besides that it improve performance of systems with other methods
[16]		SVM, Decision Tree (DT), Logistic Regression (LG), Boosting Method and Naive Bayes Multinomial	Research results can contribute to the clear automation of things that are on the computer and explained automatically on the online QA automation on large computers	Only applied in small large-scale computer, so it is necessary to develop other classes, by increasing the accuracy of automatic classification with ensemble approach, neural network feature of N-gram and stopwords

B. Education Question Answering System in Social Science

Teaching-learning in the social field [17] analyzes the classification of questions for the question answering systems in the learning realm utilized question processing, IR, and generate answering, also knowledge base techniques for the subjects of History and Geography, the results of the classification system using triggers and connecting keywords achieve 0.771 accuracy until 0.861. However, have problems, with the parsing connection dependence also resources of linguistic, so it cannot handle questions that are not following the correct grammatical structure, in addition to developing systems for grade 5 and 6 students with history and geography subjects. And also [18] building QAS for geography subject in the Turkish language, utilizing using the Rule-based technic, TF-IDF, Forward and Backward Models, and Sequence Classification Method (HMM-Glasses) for developing Hidden Markov Model (HMM) for high school students, the experimental results show that mode The formed l obtains decent accuracy, but requires manual collection of standardized information for further research, and can be developed on the subject.

The following is an explanation of the methods, results, and improvements, as well as the limitations of research in the domain of social teaching-learning described in Table 2.

Table 2. Comparing of education QAS in social field

Ref.	Description	Methods/technics	Result and improvement	Limitation
[17]	Research articles on the education QAS in the Social science with various procedures and methods can give solutions to student problems in the teaching-learning process in fields of social science	Knowledge-based, Multiple Choice Questions (MCQs)	The experimental results informs that the data training using MCQ corpus and student corpus could get an accuracy of 59%	The distribution of question types on the MCQ corpus and students' corpus affects performance so it needs to be explored to describe dependency parsing and linguistic features in order to improve system performance
[18]		HMM, Rule-based Method, Forward and Backward Models, TF-IDF	Results of the experimental informs that the former model has obtained a fairly accurate accuracy	For geographic domain of the Turkish language, and also requires manual collection of standardized information for research further

C. Education Question Answering System in Languages Science

Researchers [19] using a morphological approach, Norder, TNT tagger to build Malayan QAS can get 70%, but this research has limitations, namely focusing on closed domain QA with factoids in Malayalam, and [20] develop a QAS utilized machine learning, knowledge-base, Naive Bayes, and smith waterman model in Hindi, with experimental results up to 90% accuracy, limited implementation in Hindi, need to use classification methods and other collections of information. Furthermore, [21] utilized ontology and knowledge-based to build a QAS in Marathi, the results of the overall precision evaluation trial reached 93.95%, the recall was 94.55% and accuracy was 89.28%, the limitations of this research are using an ontology that is not automated in Marathi.

Furthermore, [22] developing QAS utilized SPARQ, NLP, ontology, resource description framework (RDF) and semantic web in Arabic, can produce 81% accuracy, but has a limitation that there is only one semantic rule. In the same year [23] carried out research on website-based QAS with a machine learning approach utilizing the support vector machine (SVM) classifier model and mean reciprocal rank (MRR), thus obtaining of 88.4% accuracy, greater than the studies [22]. But needed improve the accuracy of system with be developed types of question and feature training. In addition, [24] researched in the same field, namely building a system utilized Pattern Matching models and SVM in Arabic Islamic Hadith. So that it could significantly improve system performance 88.39% for precision, 87.66% for recall, and 87.93% for f-measure. But, this research concentrates on the questions classification such as *who*, *where* and *what*, and the future can be developed other questions classification also mores data training to increase performance systems. And researchers [25] researched QAS in Chinese in the medical domain, by utilizing the Word2Vec model, knowledge-base, word embedding, and Fast Text, can be facilitate users to ask questions and answers retrieval, but needed develop in others domain. Furthermore, [26] utilized rule-based extractor, word segmentation, logistic regression, stop words, part-of-speech tagging, to build an application questions answer in universal knowledge on Vietnam languages, with result of implementation 76.90% accuracy for training datasets, but the research focus is on open-ended questions, so it is necessary to try with other methods to get answers by reading to justifications the advantages of ontology questions, to improve performance system better, not only [27] utilizing query-based attention CNN (QACNN), end-to-end memory network (MemN2N), and multi-choice QA (MCQA), to create QAS with transfer learning for TOEFL learning and independent learning labeling. Results of the implementation display that transfers learning on various question types are effective in answering question correctly because it utilizes the concept of self-labeling transfer. But in the future, it is necessary to try to generalize the result of transfer learning to the dataset for other QA models because have problems with the explanation original information test result in TOEFL audio, it is a good step to try to initialize the QACNN embedding arrangement by acoustic words or embedding semantic directly learned audio based reading.

For all study that has been done, know that the studies has had no same, however all the studies move in the same domain, that is on language learning. Know clearly more about the comparing between the studies that have been described Table 3.

Table 3. Comparing of education QAS in languages field

Ref.	Description	Methods/technics	Result and improvement	Limitation
[19]	Research articles on the Education QAS and QA Classification in the field of linguistics by using various techniques and methods, can help to answer questions from users in the natural language quickly and accurately both offline and online	TnT tagger, Morphologies, Norder	Assessment results reach 70% accuracy	Focus on factoid domains in the Malayalam language
[20]		Naïve Bayes, knowledge-based, Smith-waterman	Assessment results reach 90% accuracy	Focus In the Hindu domain, various classification methods are needed
[21]		Knowledge-based, Ontology	Results of the testing get precision of 94%, recall 95%, and accuracy of 89%	Restrictions on using manual ontologies in the Marathi language domain
[22]		NLP, Semantic Web, Ontology, Resource Description Framework, SPARQ	Obtaining 81% accuracy results	Has a limited domain with few semantic rules
[23]		SVM, Mean Reciprocal Rank	Achieving 88.4% accuracy	Focus on factoid types of problems
[24]		SVM and Pattern Matching	Achieving Precision 88%, Recall 88% and F-measure 88%	Focus on the Question classification of *Who*, *Where* and *What* questions
[25]		Word Embedding, Word2Vec, Knowledge Base, Fast Text	Provide answers to users who ask questions in languages in the medical domain	Focusing on Chinese in the medical domain
[26]		Stop Words, Logistic Regression, Word Segmentation, Part-of-speech Tagging, Rule-based Extractor	The Systems can answer a variety of questions universally with an accuracy of 76. 90% on the training dataset	Limited to open-domain questions in Vietnamese, need another approach that can generate answers directly by asking oncologists, to improve system performance

(*continued*)

Table 3. (*continued*)

Ref.	Description	Methods/technics	Result and improvement	Limitation
[27]		End-to-end Memory Network, Query-Based Attention CNN, and Multi-Choice QA	Transfer learning is very influential on the questions classification, therefore self-labeling technique for transfer learning are effective	Development is needed to transfer audio from text, to help disabled students in TOEFL learning

D. Education Question Answering System in Biomedical Science

Research by [28] utilizing a semantic relations approach to create QAS in the medical domain, has had effective to response the questions immediately, with the results of evaluation for answers reaching 68%, needed having problems because only for biomedical realm use semantic relations. Studies [29] doing questions classification for the realm of medicine in QAS by utilizing a question classifier approach, Rule based and syntactic pattern, results of evaluation accuracy for each question classification, for How:44.54%, What: 5.33%, Why:66%, When:50%, Where:55%, and Why:66%, but there are problems in recognizing the types of patterns for good question classification, and [30] build a QAS for the biomedical domain utilizing Biomedical (BM25), Term Frequency Metric, Unified Medical Language System (UMLS) model, sentiwordnet Lexical, and metathesaurus. Results of implementation display that has effective to answer and extract tasks compared to systems participating, however needed assessment system in question classification, component of answer extraction, section search and document search, not only that [31] using Rapid Methods on epistemonikos Database and Rapid Methods was Based on PubMed to help QAS tools are based on patient recommendations, with the evaluation results for two rapid methods using the grading of recommendation assessment development and evaluation (GRADE) showing that both methods are effective in providing answers to 100 questions out of 200 questions, by obtaining 93.5% (95% CI 90% to 96.9%) classify answers as reasonable and 6.5% (95% CI 3% to 9.9%) classify answers that are potentially misleading but require other procedure in order to guarantee maximum recommendation result with satisfactory performance of system. Next [32] developing medical QAS utilizes support vector machine (SVM), interpolated smoothing, conditional random field (CRF), and obtained result as expected, namely (a) having good accuracy in identification answers and (b) increasing medical articles on PubMed by 40%, but it is necessary to improve articles from other sources of knowledge related to the medical world in order to improve the system performance.

Research conducted by [33] develop a QAS-based MEDLINE article for the biomedical domain utilizing SentiWordNet Lexical, Unified Medical Language System (UMLS), Mean Reciprocal Rank (MMR), Semantic Relations, Extract Medical Entities, Support Vector Machine (SVM), and Semantic Graph. The experimental results, the overall performance of the system for the evaluation of real questions and answers published from MEDLINE articles obtained significant results, but required a special process to answer questions containing when, why, as well as adding papers from other aspects, by combining various appropriate to improve overall system performance, in the same domain

study [34] utilized interpolated smoothing, reciprocal rating article scores (RRAS), clinical part and therapy graphs (CPTG), factored Markov network, as a follow-up to study conducted by [32] with utilizing another approach, which is able to increase biomedical articles by 49%, but it is necessary to automatically identify semantic attributes, so that it is easy to build a knowledge graph, and the knowledge base is identified from the totally of the articles taken in order to improve performance system, then researchers [35] use sparse auto-encoders (SAE), text convolution neural network (TEXTCNN), with significant impact assessment results with an accuracy of 0.915 and can avoid late installations, but need further research applying deeper semantic analysis for feature extraction, feature creation automation semantics, and researchers [36] combined Word2Vec, Expected Answer Type (EAT), Deep Neural Network (DNN), word embedding, Lexical Answer Type (LAT). The results of the analysis display that the wrong answers will be wastes in the calculation times and multiply each criterion. But need word embedding and neural network models to improve QAS performance in biomedical field.

For all the study that has been describes, it is understood that each study has differences, but has a mutually relevant domain, that is on biomedical science. The differences between the studies will be described in Table 4.

Table 4. Comparing of education QAS in biomedical field

Ref.	Description	Methods/technics	Result and improvement	Limitation
[28]	Research articles education QAS in the field of biomedical science using various techniques and methods, can help users in the process of obtaining targeted information through the system by asking questions	Semantic Relations	Claiming that the biomedical system built can provide quick answers to user questions	Only use semantic relationships in the biomedical domain
[29]		Rule-Based Approach, Question Processing Module, Syntactic Patterns, Question Classification	The proposed method for classifying questions can increase the accuracy for What:5.33%, Where:55%, When:50%, Why:66%, How:44.54%	Limitations in identifying question categorization and question classification patterns
[30]		Term Frequency Metric, Biomedical (BM25) Model, Lexical of SentiWordnet, UMLS)	The results of the implementation are very exciting and competitive in the task of extracting answers	It is necessary to evaluate the question classification, search of document, search section, and components of extraction answer

(*continued*)

Table 4. (*continued*)

Ref.	Description	Methods/technics	Result and improvement	Limitation
[31]		Rapid method based epistemonikos database, rapid method was based on PubMed	Results application from two methods able to respond to 100 inquiries from the 200 inquiries asked	Combination with other methods to improve system performance, to optimize the recommendations from the system
[32]		SVM, Interpolated Smoothing, Conditional Random Field (CRF)	The results of system testing show (a) able to recognize answers accurately and (b) increase medical literature on PubMed to 40%	Limitations of articles from other disciplines in order to improve system performance
[33]		Extract Medical Entities and Semantic Relations, Semantic Graph, SVM, UMLS, Lexical of SentiWordNet, MRR	The results of the system performance assessment are very promising so recommending using a demand relaxation strategy can improve overall system performance	However, it requires a more specialized process with complex questions (why, when), in addition to including other aspects to incorporate appropriate methods to improve system performance
[34]		Clinical picture and therapy graph (CPTG), factorized markov network, interpolated smoothing, and RRAS	Assessment results show 49% compared to more sophisticated results	However, it is necessary to automate the identification of semantic attributes so that it is easy to insert into a knowledge graph and think about paragraph parts rather than all articles in creating a knowledge base
[35]		Sparse Auto-encoders (SAE), Text Convolution Neural Network (Text CNN)	The assessment results reach 0.915% accuracy	However, requires deeper semantic analysis on feature extraction, as well as automating semantic feature generation
[36]		Word2vec, Word Embedding, Expected Answer Type (EAT) and Lexical Answer Type (LAT), Depp Neural Networking (DNN)	The results of the analysis conclude that the wrong answer will be taken when adding parameters and the calculation of word embedding in the system	However, the research is still in the form of analysis, so in the hereafter, it needs to be applied to a system

E. Discussion

In the period 2011–2021, education QA research has increased compared to before, due to it is considered a new domain of research with trends and challenges that are quite interesting to study. So it is mostly associated with architecture system, matrix evaluation, and QA components namely question classification, answer extraction, and document retrieval, involving NLP techniques, IR, and IE, further started with deep learning, reinforcement learning and machine learning in classification of the question and answer include answer validation, with various new method.

4 Conclusion and Recommendation

This Survey article category no significant difference with previous research surveys, namely providing knowledge to future researchers by categorizing, and summarizing the result of the latest studies, both methods and results of improvements and their limitations. The survey category papers on education QAS in this study found four (4) categories, such as technology information & communication sciences, sciences of social, linguistics sciences, and biomedical sciences, and also analyze the assessment results in accuracy, precision, recall, and F1 score. However, this category survey paper doesn't cover everything, but only top scientific papers for education QAS domain, according to the procedures and criteria determined in this paper category survey. This survey categorizing paper will serve as a guide for the next studies in education QA domain, concentrating on the classification of education QAS papers, involving various models, methods, and/or good procedures, so that satisfactory results can be obtained.

References

1. Karthika, K., Radhakrishnan, D.: Developing a natural language interface with semantic matching to complex data. J. Adv. Res. Dyn. Control Syst. **13**, 102–106 (2017)
2. Allam, A.M.N., Haggag, M.H.: The question answering systems: a survey. Int. J. Res. Rev. Inf. Sci. **2**, 1–12 (2012)
3. Bouziane, A., Bouchiha, D., Doumi, N., Malki, M.: Question answering systems: survey and trends. Procedia Comput. Sci. 366–375 (2015). https://doi.org/10.1016/j.procs.2015.12.005
4. Neves, M., Leser, U.: Question answering for Biology. Methods **74**, 36–46 (2015). https://doi.org/10.1016/j.ymeth.2014.10.023
5. Alkholy, E., Haggag, M.H., Aboutabl, A.: Question answering systems: analysis and survey. Int. J. Comput. Sci. Eng. Surv. **09**, 1–13 (2018)
6. Soares, T.G., Azhari, A., Rokhman, N., Wonarko, E.: Education question answering systems: a survey. In: Proceedings of The International MultiConference of Engineers and Computer Scientists 2021, Lecture Notes in Engineering and Computer Science, 20–22 October 2021, Hong Kong. IAENG Open Access Publication (2021)
7. Mei, L.H.: Intelligent question answering system of research based ontology on excellent courses: university computer basis course as an example. In: 2012 Fourth International Conference on Computational and Information Sciences, Chongqing, China, pp. 784–787 (2012). https://doi.org/10.1109/ICCIS.2012.177
8. Perera, R.: IPedagogy: question answering system based on web information clustering. In: 2012 IEEE Fourth International Conference on Technology for Education, Hyderabad, India, pp. 245–246 (2012). https://doi.org/10.1109/T4E.2012.48

9. Xia, T., Chai, Y., Lu, H., Wang, T.: E-learning support system aided by VSM based question answering system. In: 2013 8th International Conference on Computer Science and Education, Colombo, Sri Lanka, pp. 1281–1285 (2013). https://doi.org/10.1109/ICCSE.2013.6554118
10. Pudaruth, S., Boodhoo, K., Goolbudun, L.: An intelligent question answering system for ICT. In: 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), Chennai, India, pp. 2895–2899 (2016). https://doi.org/10.1109/ICEEOT.2016.7755228
11. Karpagam, K., Saradha, A.: A mobile-based intelligent question answering system for education domain. Int. J. Inf. Eng. Electron. Bus. (IJIEEB) **10**, 16–23 (2018). https://doi.org/10.5815/ijieeb.2018.01.03
12. Calijorne Soares, M.A., Cardoso Brandão, W., Silva Parreiras, F.: A neural question answering system for supporting software engineering students. In: 2018 XIII Latin American Conference on Learning Technologies (LACLO), São Paulo, Brazil, pp. 201–207 (2018). https://doi.org/10.1109/LACLO.2018.00047
13. Knobloch, J., Kaltenbach, J., Bruegge, B.: Increasing student engagement in higher education using a context-aware Q&A teaching framework. In: ACM Proceedings of the 40th International Conference on Software Engineering: Software Engineering Education and Training, pp. 136–145 (2018). https://doi.org/10.1145/3183377.3183389
14. Xue, H.: Research on intelligent question answering system of network course for mu. In: Proceedings of the 2017 4th International Conference on Education, Management and Computing Technology, vol. 101, pp. 872–875 (2017). https://doi.org/10.2991/icemct-17.2017.187
15. Wang, L., Wang, W.: Research and construction of junior high school subject Q&A system model based on deep learning. In: International Conference on Information Systems and Computer Aided Education (ICISCAE 2018), pp. 504–508 (2018). https://doi.org/10.1109/ICISCAE.2018.8666853
16. Hao, Q., Galyardt, A., Barnes, B., Branch, R.M., Wright, E.: Automatic identification of ineffective online student questions in computing education. In: Proceedings of Frontiers in Education Conference FIE, vol. 2018–October, pp. 1–5 (2019). https://doi.org/10.1109/FIE.2018.8658642
17. Lee, Y.-Y., et al.: Question type analysis for question-answering applications in education. In: Proceedings of 22nd International Conference on Computers in Education (ICCE 2014), pp. 132–134 (2014)
18. Derici, C., et al.: Question analysis for a closed domain question answering system. In: Gelbukh, A. (ed.) Computational Linguistics and Intelligent Text Processing, vol. 9042, pp. 468–469. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-18117-2
19. Seena, I.T., Sini, G.M., Binu, R.: Malayalam question answering system. Procedia Technol. **24**, 1388–1392 (2016). https://doi.org/10.1016/j.protcy.2016.05.155
20. Nanda, G., Dua, M., Singla, K.: A Hindi question answering system using machine learning approach. In: International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT), pp. 5–8 (2016)
21. Govilkar, S.S., Bakal, J.W.: Question answering system using ontology in Marathi language. Int. J. Artif. Intell. Appl. **8**, 53–64 (2017)
22. Albarghothi, A., Khater, F., Shaalan, K.: Arabic question answering using ontology. Procedia Comput. Sci. **117**, 183–191 (2017). https://doi.org/10.1016/j.procs.2017.10.108
23. Ahmed, W., Ajusha, P.V.: Web-based Arabic question answering system using machine learning approach. Int. J. Adv. Res. Comput. Sci. **8**, 40–45 (2017)
24. Hasan, A.M., Rassem, T.H.: Combined support vector machine and pattern matching for Arabic Islamic Hadith question classification system. In: Saeed, F., Gazem, N., Mohammed, F., Busalim, A. (eds.) Recent Trends in Data Science and Soft Computing, vol. 843, pp. 278–290. Springer, Cham (2019). https://doi.org/10.1007/978-3-319-99007-1_27

25. Feng, G., Du, Z., Wu, X.: A Chinese question answering system in medical domain. J. Shanghai Jiaotong Univ. (Sci.) **23**(5), 678–683 (2018). https://doi.org/10.1007/s12204-018-1982-1
26. Le-Hong, P., Bui, D.-T.: A factoid question answering system for Vietnamese. In: ACM First International Workshop on Hybrid Question Answering with Structured and Unstructured Knowledge, pp. 1049–1055 (2018). https://doi.org/10.1145/3184558.3191535
27. Chung, Y.-A., Lee, H.-Y., Glass, J.: Supervised and unsupervised transfer learning for question answering. In: Proceedings of NAACL-HLT, pp. 1585–1594 (2018)
28. Hristovski, D., Dinevski, D., Kastrin, A., Rindflesch, T.C.: Biomedical question answering using semantic relations. BMC Bioinform. **16** (2015). https://doi.org/10.1186/s12859-014-0365-3
29. Dodiya, T., Jain, S.: Question classification for medical domain question answering system. In: IEEE International WIE Conference on Electrical and Computer Engineering, no. December, pp. 204–207 (2016). https://doi.org/10.1109/WIECON-ECE.2016.8009118
30. Sarrouti, M., Ouatik, S., Alaoui, E.: A biomedical question answering system in BioASQ 2017. In: ACM Proceedings (2017)
31. Izcovich, A., et al.: Answering Medical Questions at the Point of Care : a Cross-SectionalStudy Comparing Rapid Decisions Based on PubMed and Epistemonikos Searches with Evidence-Based Recommendations Developed with the GRADE Approach, pp. 1–9. CrossMark (2017). https://doi.org/10.1136/bmjopen-2017-016113
32. Goodwin, T.R., Harabagiu, S.M.: Medical question answering for clinical decision support. In: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, pp. 297–306 (2016)
33. Ben, A., Zweigenbaum, P.: MEANS: a medical question-answering system combining NLP techniques and semantic web technologies. Int. J. Inf. Process. Manag. **51**, 570–594 (2015). https://doi.org/10.1016/j.ipm.2015.04.006
34. Goodwin, T.R., Harabagiu, S.M.: Knowledge representations and inference techniques for medical question answering. ACM Trans. Intell. Syst. Technol. **9** (2017). https://doi.org/10.1145/3106745
35. Yan, G., Li, J.: Mobile medical question and answer system with improved char-level based convolution neural network and sparse auto encoder. ACM Trans. Intell. Syst. Technol. 20–24 (2019)
36. Kamath, S., Grau, B., Ma, Y.: Verification of the expected answer type for biomedical question answering. In: ACM First International Workshop on Hybrid Question Answering with Structured and Unstructured Knowledge, vol. 2, pp. 1093–1097 (2018)

Technological and Organisational Factors Influencing Alignment of Information Technology and Business Objectives

Millicent M. Motheogane(✉) and Agnieta B. Pretorius

Tshwane University of Technology, Pretoria 0001, South Africa
millicent.motheogane@wits.ac.za, pretoriusAB1@tut.ac.za

Abstract. IT is one of the most important tools for increasing a company's competitiveness by increasing efficiency. Strong IT skills are required to support company strategy. This quantitative analysis focused on the technological and organizational variables that determine how information technology is aligned with strategic business objectives. This study employed a closed-ended questionnaire to collect data on the elements that contribute to the success of alignment of information technology and business objectives in the workplace. The data was analyzed using a factor analysis tool and the SPSS application. The perceived factors were ranked and weighed according to their importance. The study's conclusions will be used by organizations, particularly parastatals to make future technology investment decisions.

Keywords: Business strategy · Business strategic objectives · Government · IT strategy · Technological factors · Organisational factors

1 Introduction

The majority of organizations rely on information technology to carry out the procedures and activities necessary to service consumers and engage with partners. To accomplish this, information technology solutions must support the organization's strategy (Motheogane and Pretorius 2021). Business strategies are critical to any firm's success since they help the organization achieve a competitive edge (Porter and Canada 1985). IT systems is one of these competencies.

Businesses can benefit significantly from information technology. However, it is critical that organizations pay attention to the IT competencies in order to maximize their IT investments (Strnadl 2006). By using the appropriate IT solutions, a business can save money and time while enhancing productivity.

Businesses are shifting their sales channels to an omnichannel business model in today's highly competitive global environment. This is a business strategy in which online and offline operations are combined (Reinartz et al. 2019). Customers can benefit from both speedy internet transactions and traditional face-to-face service (Reinartz et al. 2019). These kind of business models require an advanced information technology

S. Ao et al. (Eds.): IMECS 2021, *Transactions on Engineering Technologies*, pp. 92–102, 2023.
https://doi.org/10.1007/978-981-19-7138-9_8

and internet-based platforms (Chen et al. 2021).Businesses, particularly e-Commerce enterprises that strive to connect and share information in order to serve customers and suppliers along the supply chain, must be able to incorporate modern IT resources into operational activities in this respect. Government entities should follow suit (Chen et al. 2021).

State-owned firms, sometimes referred to as 'parastatals,' are crucial to South Africa's strategic objectives. To attain economic development, the National Development Plan promotes parastatal policies to be linked with the national vision, they therefore, require IT capabilities to operate effectively. However, it is unclear how IT and business strategic objectives may be matched, successful alignment continues to elude practitioners (Tornatzky and Fleischer 1990; Luftman et al. 2002; Thackrah 2008; Mongale et al. 2021). The world has evolved, government parastatals struggles to align their business strategy with information technology, which results in unfinished projects and systems that do not support the business (Elmorshidy 2013).

This book chapter was created in response to the preceding discussion of IT and business alignment and technological progress. It is an enhanced version of a conference paper presented at the Engineers and Computer Scientists International Multi-Conference in 2021 (Motheogane and Pretorius 2021). The persistence of this problem shows how vital it is for organizations to understand the elements that influence successful alignment of IT and business. This type of situation necessitates examination and reporting, particularly in South African state-owned enterprises. This book chapter analyzes and interprets the results in greater detail. It is crucial to note that these findings were published in 2018, implying that much has changed as a result of technological advancements and the impact of the COVID 19 outbreak in South Africa. The objective of this paper is therefore to investigate the technological and organisational factors influencing the alignment of IT and business objectives.

2 Alignment of IT and Business

Aligning business and IT strategies is a collaborative activity, which means that the two cannot be accomplished independently (Mongale et al. 2021; Motheogane and Pretorius 2021). Alghazi et al. (2019) assert that an organization's efficiency can be increased by having well-aligned IT and business strategy.

Management is accountable for the connection of IT projects and business strategy (Schwalbe 2004; Alghazi et al. 2017; Motheogane and Pretorius 2021; Jambari and Hamid 2017). As a result, managers must educate themselves about technology capabilities and keep IT experts informed about changing business requirements (Schwalbe 2004). For instance, when an organization redesigns or renews itself, it must be prepared for chances for improvement, which implies that the organization's strategic strategy should drive the selection of IT initiatives. Management is responsible for designing strategies that enable information technology to support the business (Schwalbe 2004).

According to Serfontein and Hough (2011), the depth of any strategy is defined by the actions and steps taken by management to improve the organization's financial performance. The writers continue by defining strategy and stating that it is evident when management works to increase its competitive position in order to outperform

competitors (Serfontein and Hough 2011). In other terms, a strategy is a tool used by corporations to surpass their rivals. According to the authors, there are four effective strategies for differentiating a business from its competitors: aspiring to be the industry's lowest-cost provider in comparison to competitors, outperforming competitors on value proposition, gaining a competitive advantage through outperformance, and finally, developing expertise that is difficult to imitate. Additionally, these four ways enable organizations to develop strong customer loyalty, resulting in a long-term competitive advantage (Serfontein and Hough 2011).

The function of information technology in today's economy is constantly evolving and has shifted intensely from its early days as data processing (Roztocki et al. 2019). IT's job in the organization is to arm employees with the tools and resources necessary to deliver services successfully and efficiently. Organizations that employ information technology to gain a competitive edge sometimes differentiate themselves from their competition (Motheogane and Pretorius 2021). It is past time for businesses to acknowledge information technology as a significant tool and resource. The purpose of an organization's information technology strategy is to ensure that efficient and effective systems are in place.

Public organizations are supposed to conserve money while being efficient under the South African Public Finance Management Act (PFMA). Budgets are frequently slashed radically. As a result, managers must pool their resources and establish ways to ensure that they continue to provide superior service (Motheogane and Pretorius 2021). The PFMA's goal is to regulate financial administration by requiring the efficient and effective management of all revenue spending (assets and liabilities). Additionally, it acts as a guide for individuals charged with financial management duties (Treasury 2010). Therefore, organizations must enhance management procedures and harness existing investment instruments, with the primary criterion being that business units understand their business needs and connect IT with those goals. This is the point at which the vast majority of businesses fail. Businesses that outperform their competitors do more than just tighten their belts to keep expenses down. They are businesses that invest in their operations in order to develop new or enhanced products and services (Roztocki et al. 2019; Motheogane and Pretorius 2021).

As a result, organizations require technology to drive strategy and ensure that all established objectives are completed. They demand systems that are effective, dependable, and pertinent (Motheogane and Pretorius 2021). According to Yeh et al. (2014), organizations are obliged to link their internal value chain systems with external demand activities in order to respond swiftly to changing market conditions. Critical to the organization's success is the environment in which it functions. Organizations must evolve to keep pace with changing times and trends. For instance, if the environment no longer permits the use of manual methods, the organization must transition in order to remain relevant to its consumers and stakeholders.

The following can collectively be taken as benefits of alignment of IT with strategic business objectives (Yeh et al. 2014; Mithas and Rust 2016; Motheogane and Pretorius

– Cost savings - Organizations can save money on travel, time, and other expenses by utilizing solutions such as virtual systems, online systems, and electronic systems.

- Standardised processes - because they are clear and well documented, standardised processes make it relatively easy for organizations to establish activities and eliminate lack of understanding on projects.
- Increased productivity - better systems equate to increased productivity.
- Enhanced workflow and communications - improved systems enable improved workflow and communications.
- Sustaining repeatable service levels - this demonstrates dedication to the project.
- Enhancements to risk management mechanisms - enables organizations to mitigate potential risks and dangers.
- Developing a competitive advantage through the exploration and application of innovative technology.

Contributing Factors to the Successful Alignment

- IT as a business enabler - IT is considered as a business enabler rather than a cost (Tornatzky and Fleischer 1990).
- Communication - Ongoing contact between information technology and the company (Luftman et al. 2002; Thackrah 2008; Mogase and Kalema 2015; Roztocki et al. 2019).
- Organizational structure - The structures that facilitate the alignment of IT and business (Tornatzky and Fleischer 1990; Henderson and Venkatraman 1999; Thackrah 2008; Mongale et al. 2021; Motheogane and Pretorius 2021).
- Alignment facilitation process - The decision-making mechanism that facilitates alignment (Tornatzky and Fleischer 1990; Henderson and Venkatraman 1999; Luftman et al. 2002; Thackrah 2008; Motheogane and Pretorius 2021).
- Management capabilities - The organizations level of management capabilities, management capacity, and IT leaders (Schwalbe 2004; Thackrah 2008; Mogase and Kalema 2015).
- Individuals – Involve individuals in the alignment process (Carpenter et al. 2011).
- Competencies and skills - Capabilities of individuals to carry out critical tasks that support corporate strategy (Henderson and Venkatraman 1999; Luftman et al. 2002).

3 Related Work

Researchers from all over the world are interested in the synchronization of information technology with business strategies. The following are common obstacles and experiences, as well as proposed aspects that could aid in successfully aligning IT and business.

Shanks (2007) conducted a study to measure the level of alignment in Australian firms and to identify the factors that contributed to alignment. The author mentions that IT and business strategy alignment can be facilitated by a number of factors including IT as an organizational tool, communication, organizational culture and organizational structure (Shanks 2007).

Thackrah (2008) conducted a survey on "Alignment of business and information strategies in the financial services sector in South Africa". The study analyzed organisational, people, and processes to ascertain which of the three elements was critical in creating alignment between business and information technology initiative as well as determining the degree of success gained as a result of these variables. The study

revealed that the process factors were the most important and were successfully applied in South African financial institutions (Thackrah 2008).

Chen (2010) also conducted a study looking at the maturity of IT alignment within firms. According to the survey, senior executives were also uninformed of information technology capabilities, the technological competency of information systems departments, and the information systems department's role. The study's findings proposed the strategic alignment maturity assessment. This assessment provides a means of assessing an organization's current state and where it needs to go in order to achieve and maintain business-IT alignment. The author continues to mention that in order to guarantee that IT is being used to support or drive a company's business plan, it is critical to conduct a thorough assessment of a company's alignment maturity.

The internet's popularity and dependability have grown as more businesses operate digitally. Electronic collaboration is likewise advancing at a rapid pace. People rarely shop in physical stores, preferring to shop online. ICT has revolutionized the way we connect with one another, seek out necessary information, work, do business, interact with government authorities, and manage our personal life. Without a phone and a computer, one cannot function; this is how digital the world has become.

4 Conceptual Model

The conceptual framework unifies all components of the study, the purpose, associated literature, proposed methodology including the problem being addressed (Creswell and Clark 2007). Additionally, it assists in highlighting the links between the researched constructs. This model illustrates factors derived from the literature on the basis of Henderson & Venkatraman's definition of strategic alignment (1993). However, in light of SA's government parastatals such as the NRF and the fact that technology advances, the following technical aspects were jointly assessed to be affecting IT and business alignment. These factors are classified using categories derived from Tornatzky and Fleischer's (1990) TEO model and the literature.

This study concentrated exclusively on technological and organizational variables affecting South African parastatals. The world evolved and times have changed. The success of company plans and innovations is influenced by technological and organizational factors. These aspects are the drivers behind organization achieving their strategic objectives.

5 Methodology

The data for this study was gathered by closed-ended questionnaires distributed to the parastatal organization, the National Research Foundation. Participants were asked to rank the perceived factors in terms of their importance in attaining IT and business alignment. The SPSS software was used for the analysis, then perceived criteria were prioritized and weighted using excel spreadsheet to establish the critical factors.

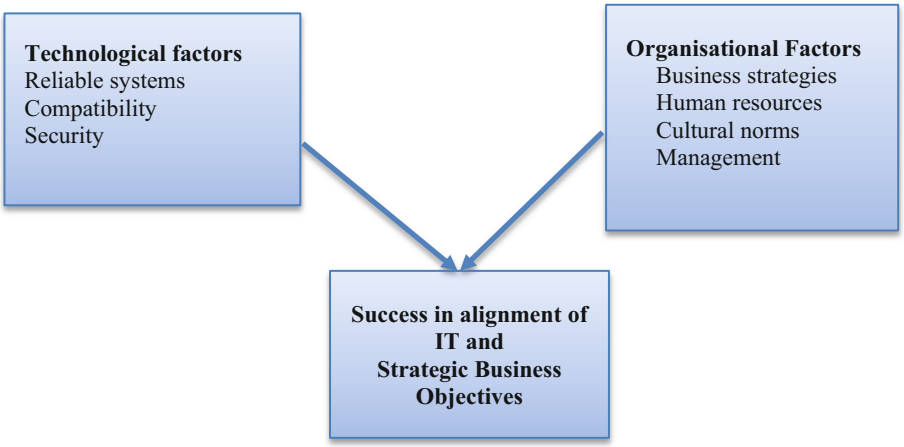


Fig. 1. Conceptual model

Results

In this investigation, non-probability random sampling was used. This sampling technique enables the researcher to verify that an adequate number of individuals is picked from each category so that the overall sample appropriately reflects the known proportions of each group (Goddard and Melville 2004). This option was selected because this study was intended to target a specific group of persons who possessed specific relevant knowledge or were expected to possess knowledge in the disciplines of information technology and business. A key objective of this research was to look at how organizational and technological factors interact to influence the alignment of information technology with long-term business goals. Participants were selected from five different departments: Reviews and Evaluations, Grants and Management Systems and Administration, Knowledge and Management Systems, International Relations, and Information Technology.

The data collection questions were derived from the conceptual framework given in Fig. 1, which was guided by the relevant literature. The measuring items were developed in accordance with the construct factors.

Construct validity was addressed by researching pertinent literature and utilizing variables previously utilized by researchers as the basis for questionnaire measurements. Additionally, construct validity was examined at a coarser level using correlation analysis to determine the interdependence of the constructs as indicated in the conceptual model for this study.

Green and Salkind (2011) define a measure as reliable if it consistently yields scores across administrations, implying that the results may be relied upon. The reliability analysis is done to verify the scales' correctness and their consistency over time. Cronbach's Alpha is the most often used reliability metric.

Reliability for each construct was assessed and produced a reliability of 0.967, which was much greater than the threshold of 0.7 and thus sufficient to trust the questionnaire answers. The dependability stats for the elements is shown below (Table 1).

Table 1. Constructs' stats reliability

Construct	Cronbach's Alpha	Standardised items – Cronbach's Alpha	# Items
Technological			
Reliability (TECREL)	.878	.880	4
Compatibility (TECCOMP)	.858	.864	4
Security (TECSEC)	.916	.920	4
Organisational			
Business strategy (ORGSTR)	.858	.864	3
Human Resources (ORGHRRSP)	.749	.754	3
Cultural norms (ORGCUL)	.883	.884	3
Management Responsibility (ORGTMR)	.916	.917	4

As shown in Table 2, all technological elements received a high ranking, with average perception ratings ranging between 1 and 2 on the Likert scale, indicating that the weighted score is between 3.17 and 3.52, corresponding to a weight of between 3 and 4. These findings reflect participants' assessment of the relative relevance of several factors in the alignment process.

Table 2. Technological elements based on rakings

Technological factors		
Very important	Very important + important	Score weightings
Technological reliability (65.50), (64%)	Technological reliability (91.50), (89%)	Technological reliability (3.52)
Technological compatibility (56), (54%)	Technological compatibility (84), (82%)	Technological compatibility (3.34)
Technological security (50), (49%)	Technological security (75.5), (73%)	Technological security 3.17

As shown in Table 3, all organizational aspects were highly valued, with average perception ratings ranging between 1 and 2, indicating a weighted score of between 3.23 and 3.29, which corresponds to a weight of between 3 and 4. These findings reflect participants' assessment of the relative relevance of several factors in aligning IT and strategic business objectives.

The perceived criteria were prioritized in terms of their importance. This aimed to provide a summary of the participants' responses, which reflected their assessments of the factors' relative importance.

Table 3. Ranking of organisational factors

Organisational factors		
Very important	Very important + important	Score weightings
Organisational strategy (60.33), (59%)	Organisational strategy (89.67), (87%)	Organisational strategy (3.29)
Organisational top management responsibility (64.00), (62%)	Organisational top management responsibility (91.00), 88%	Organisational top management responsibility (3.48)
Organisational human resource/employee (54.33), (53%)	Organisational human resource/employee (89.33), (87%)	Organisational human resource/employee (3.39)
Organisational culture (48.33), (47%)	Organisational culture (82.00), 80%	Organisational culture 3.23

The first column of Tables 2 and 3 ranks factors according to the percentage of respondents who indicated that a factor is very important to them while the second one ranks the factors based on very important or important. The last column ranks factors by averaging their replies base on the following descriptions; very important = 4, important = 3, moderately important = 2, slightly important = 1 and not important = 0.

Technological reliability was raked 64% in terms of very important and 89% in terms of very important or important, and the average weighted rating examined was 3.52.

Accountability for organizational top management raked 62% in terms of very important and 88% in terms of very important or important with a weighted average rating of 3.48. The technical security ranked the lowest with 49% in terms of very important and 73% in terms of very important or important with the average weighted rating of 3.17, otherwise all analyzed constructs were shown to be important in aligning IT and business objectives.

6 Interpretation of the Statistical Findings

The answer to the question "What technological factors influence the alignment of information technology with the NRF's strategic business objectives?" This research question was answered by conducting a study of the literature on technological factors affecting the alignment of information technology with strategic business objectives. Technology reliability, compatibility and security were analyzed, an average score of between 73 and 89% perceived these factors as very important or important.

The conclusions of this study corroborate those of Bouchbout and Alimazighi (2009), who assert that information technology systems facilitate an organization's business strategy objectives and that it is vital for organizations to have dependable information technology systems. This is also consistent with DeLone and McLean's (2003) results that, while system quality has an effect on system success, its effect on user satisfaction is relative and highly dependent on the flexibility, availability, dependability, response speed, and usability of the system.

Technology compatibility, as defined by Sithole (2014), is the process of ensuring that an organization's infrastructure is compatible with new technology. Sithole (2014) continues by stating that system innovation in organizations should be interoperable with other technologies and should improve how the organization runs, which is consistent with these findings.

The answer to the question "What organizational factors influence the alignment of information technology with the NRF's strategic business objectives?" This research question was answered by conducting a review of the literature on the organizational elements such as organizational strategy, human resources/employees, top management, and organizational culture.

Participants were asked about these factors, an average of between 80 and 87% of respondents responded that organizational elements are 'very important' or 'important'.

Organizational strategy, according to Yeh (2014), gives organizations a competitive edge. The author mentions that it is difficult for the organisation to exist without a well-defined business strategy, supportive human resources, and workers capable of assisting the organization in achieving its objectives. Regardless of the circumstances, employees must constantly support the organization's vision. This is also supported by Hough et al. (2011), who emphasise the importance of management comprehending their business and remaining competitive. Top management accountability and human resource management are similar in that both promote business and IT alignment. According to the DPSA's (2012) Governance of ICT policy framework, ICT decisions should be made by the state's senior political and management leadership, rather than delegated to technology professionals. This demonstrates the critical nature of managing information technology on a same footing with other resources.

Additionally, the current study corroborates Hoffman et al. (2014), who assert that organizational culture and environmental factors play a substantial effect in the success of information technology projects. Hoffman et al. (2014) define organizational culture as "output-oriented organizations, effective managerial abilities, and an appropriate work environment," whereas environmental influences include governmental regulations and competitive pressures. According to the article, top management accountability is also crucial, as initiatives that lack management backing are unlikely to be executed. Additionally, it argues for managers to foster and support new ideas.

7 Conclusion

Aligning IT and business strategy is crucial for the progress and achievement of an organization's strategic goals. IT has remained a critical instrument for improving an organization's competitiveness. Additionally, the globe has evolved into a competitive environment that includes parastatals. Organizations are impacted by the environment's demanding requirements and must ensure that their IT strategy and business strategy are in sync.

The literature suggested factors in the categories of technological and organizational factors, which were then examined in the setting of the study organization. Each component was found to have an effect on IT alignment process. The study's findings are intended to be used by organizations, particularly parastatals, when making technology

investment decisions. It will be interesting to repeat the study in the same organization, given that these were 2018 findings and the organization may have transformed in the aftermath of the COVID 19 epidemic.

The COVID-19 epidemic and South Africa's quarantine period necessitated that all organizations in the country convert to teleworking. Technology was now the driving factor behind the NRF and other government parastatals. They required technology in order to continue working with their partners. All face-to-face meetings were changed to online meetings, and all documents couriered to collaborators for signatures needed to be signed online and their forms digitized. This is the moment during which they were compelled to completely embrace technology in order to ensure business continuity. Management needed to establish structures, their plans needed to evolve, and IT strategies needed to be realigned with the business's to assure success.

References

Alghazi, A., Li, M., Shen, J., Fosso, S.: Misalignment between Business and IT Strategic Objectives in Saudi Arabia Public Sector Organisations, pp. 212–220 (2019)

Alghazi, A., Li, M., Shen, J., Fosso, S.: Aligning business strategy with IT strategy from business model to enterprise in Saudi Arabia public sector. In: Proceedings of the 21st Pacific Asia Conference on Information Systems: Societal Transformation Through IS/IT (2017)

Carpenter, A., Sanders, G., Harling, F.: Strategic Management: A Dynamic Perspective (2011)

Chen, Y.Y., Huang, H.L., Sung, S.F.: Alignment effect between electronic business strategy and information technology capabilities on value creation in employing industrial Internet of Things'. Sens. Mater. **33**(2), 657–669 (2021)

Chevez, N.V.A.: Unified strategic business and IT alignment model. Public Perform. Manag. Rev. (2010)

DPSA: Public Service Corporate Governance of Information and Communication Technology Policy Framework (2012)

Elmorshidy, A.: Aligning IT with business objectives: a critical survival and success factor in today's business. J. Appl. Bus. Res. **29**(3), 819–828 (2013)

Henderson, J.C., Venkatraman, N.: Strategic alignment: leveraging information technology for transforming organizations. IBM Syst. J. **38**(2), 472–484 (1999)

Krafzig, D., Banke, K., Slama, D.: Enterprise SOA - Service-Oriented Architecture Best Practices (2004)

Tornatzky, L.G., Fleischer, M.: The Processes of Technological Innovation. Lexington Books, Lexington (1990)

Luftman, J., Papp, R., Brier, T.: Enablers and Inhibitors of Business-IT Alignment, Communications of the Association for Information Systems (2002)

Mithas, S., Rust, R.: How information technology strategy and investments influence firm performance: conjecture and empirical evidence. Manag. Inf. Syst. Q. **40**(1), 223–246 (2016)

Mogase, R.C., Kalema, B.M.: E-Resources Usage in South African Higher Institutions of Learning, pp. 261–266 (2015)

Mongale, A., Kekwaletswe, P.R.M., Mogoale, P.D.: A Framework for the Alignment of Information Technology and Business in South African Development Banks (2021)

Motheogane, M., Pretorius, P.: Alignment of information technology and strategic business objectives. In: Proceedings of the International MultiConference of Engineers and Computer Scientists 2021, Lecture Notes in Engineering and Computer Science, 20–22 October 2021, Hong Kong, pp. 55–58 (2021)

Porter, M.E., Canada, M.M.: The Competitive Advantage: Creating and Sustaining Superior Performance. Free Press, New York (1985)
Porter, M., Siggelkow, N.: Contextuality within activity systems and sustainability of competitive advantage. Acad. Manag. Perspect. **22**(2), 34–56 (2008)
Roztocki, N., Soja, P., Weistroffer, H.R., Soja, P.: Information technology for development'. Inf. Technol. Dev. **25**(2), 171–183 (2019)
Reinartz, W., Wiegand, N., Imschloss, M.: The impact of digital transformation on the retailing value chain. Int. J. Res. Mark. **36**(3), 350–366 (2019)
Schwalbe, K.: Project Management Techniques (2004)
Serfontein, K., Hough, J.: Nature of the relationship between strategic leadership, operational strategy and organisational performance. S. Afr. J. Econ. Manag. Sci. **14**(4), 393–406 (2011)
Spósito, M.A.F., Neto, A.C.D., Da Silva Barreto, R.: Business-IT alignment research field: a systematic literature review. In: Proceedings of the 18th International Conference on Enterprise Information Systems, ICEIS 2016, 2 April 2016, pp. 549–558 (2016)
Strnadl, C.F.: Aligning business and IT: the process-driven architecture model Inf. Syst. Manag. **23**, 67–77 (2006)
Thackrah, J.: The alignment of business and information technology strategy in the financial services sector in South Africa (2008)
Treasury: Public Finance Management Act (1), pp. 1–82 (2010)
Yeh, C.H., Lee, G.G., Pai, J.C.: Using a technology-organization-environment framework to investigate the factors influencing e-business information technology capabilities (2014)

Enhanced Protection of Pseudonymized User Data via the Use of Multilayered Hardware Security

Mukuka Kangwa(✉), Charles S. Lubobya, and Jackson Phiri

University of Zambia, Great East Campus, P.O. Box 32379, Lusaka, Zambia
mukukakangwa@yahoo.com, cslubobya@unza.zm,
jackson.phiri@cs.unza.zm

Abstract. The use of Information Communication Technology (ICT) in various spheres of life has led to the need for service consumers to have an electronic presence in the cyber-space. The world has become a global village; one can buy goods and services thousands of miles away from the country thanks to technology. One can even access medical services remotely using telemedicine. ICT is also being used at Airports, Universities, and many other areas that require interaction. To be able to interact and use some of these online platforms, a user is normally required to provide some form of Identity for their electronic Id to be created and thereafter be allowed to participate. The Information harvested is stored on various databases across the globe owned by different Service providers. This has often led to personal data being inadvertently leaked thereby putting the owners at risk. Several software solutions have been proposed and implemented to address this challenge but the problem still exists. This paper proposes multilayered hardware solution to protect user data while allowing the user to enjoy the convenience and benefits of accessing electronic services. The prototype was developed using Arduino microcontrollers, Arduino IDE (programming platform) and the Proteus Simulation software. Results obtained from the experiments demonstrated that the protection of user data can be enhanced via the use of multilayered hardware Security.

Keywords: One time password · Data privacy · Pseudo data · Data protection · Hardware security · Microcontroller security

1 Background

Service providers offering online services usually require those who want to use their platforms to provide user data such as names, Phone numbers, and emails addresses before granting the user access to their platforms [1]. This has resulted in huge amounts of personal data being collected by several service providers. This accumulated data is deposited thereby being susceptible to leakage [2]. Leaked user data can be used to commit cyber-related crimes like financial fraud perpetrated using leaked user credentials. While numerous solutions have been designed and engaged to address this challenge,

S. Ao et al. (Eds.): IMECS 2021, *Transactions on Engineering Technologies*, pp. 103–115, 2023.
https://doi.org/10.1007/978-981-19-7138-9_9

the problem persists [3]. Incidents keep being reported where huge amounts of user data are being leaked inadvertently [4]. Data exposed to the internet is at risk of being leaked hence the need to provide more effective data protection methods [5].

Frank and Michael proposed a model based on having a Trusted Party that would provide static User Identities (ID). The user, having previously registered with the ID service provider, obtains an ID from the digital ID provider and submits it to the e-commerce platform service provider as proof of identification. The e-commerce platform provider confirms with the ID provider if the user can be trusted as being a traceable person through the ID service providers if need arose. The response the ID provider returns determines whether or not a service will be rendered to the user. Pseudonymization was used and not anonymization to make the user traceable when the need arose [2].

The use of static electronic IDs is not effective in providing user privacy as the static ID can be profiled and hence privacy getting violated [1]. Further, Blockchain technology proposed in this solution is currently resource-intensive and hence might not be suitable for a worldwide solution [6]. The less energy used, the better. Another challenge is the lack of scalability of blockchain. For any transaction to be confirmed as having taken place, there is a need for several participating nodes to confirm it as such hence the more nodes the slower the process [7]. The complication of the technology makes it cost-prohibitive. The platform can be designed and developed as a shared service to minimize associated costs [8].

Most of the user information is stored online by e-commerce platforms thereby exposing that information to potential online leakage [9]. Several cyber-attacks have been carried out on e-services like e-bay whose databases holding user personal data were accessed and customer information got compromised [10].

Fanghan and his team proposed the use of a cloud server to store user data. The user (owner of data) encrypts their data before sharing and decides who can be given access by sharing their decryption keys with the desired users [9]. This means several users can be given access to the data hence increasing the possibility of leakage. Despite data being encrypted, data is most likely to be in plain text when being processed making it susceptible to leakage [11].

The advancement in technology has resulted in the need for more effective techniques to protect personal and other sensitive information. Existing solutions might not be adequate to meet the levels of privacy and security required by regulations such as the GDPR [12]. Techniques such as pseudonymization and anonymization can be used. For commerce user data, pseudonymization is favored over anonymization as there might be a need to trace users such as when there is fraud involved.

2 Hardware Security

Hardware Security has become very paramount as more Internet of Things (IoT) devices are being developed and deployed on the internet. These IoT devices are mostly controlled by Microcontrollers. A microcontroller is a small programmable processor specifically designed for embedded systems because of their low power consumption and low cost [13].

Secure microcontrollers are designed to protect confidential internal information [14]. IoT devices driven by microcontrollers are being deployed to monitor and report

about the world around them. They are interconnected via the internet. However, their security has become a great concern [13]. These devices can be attacked in various ways such as eavesdropping, IP spoofing, and denial of service [15]. Further, even micro probing can be used to attack microcontrollers despite data in memory being encrypted [14]. The vulnerabilities being discovered can compromise their security [13]. Vulnerabilities, once discovered, can be patched [16]. However, as long as they have not yet been discovered, there will be no effort to address them. Even so, despite being discovered and patches being developed to address the security vulnerabilities, there remains several users who do not make an effort to patch their devices. It could be due to a lack of knowledge about the vulnerabilities or not being concerned with the issues that come with the risks. There is a need to provide a solution to enhance hardware security. Patching might not be sufficient as not everyone patches their devices. Moreover, even if patches are installed, not all vulnerabilities might have been discovered and hence patched. A more effective solution is needed.

Some devices are deliberately infected with Hardware Trojans (HT) when being manufactured. This act aims to later control them and activate the HTs so that critical data can be leaked [17]. The data leak due to insecure firmware in devices can harm user privacy as well as discourage the adoption of internet-based services [16]. Some malicious code can be installed in the firmware and run when the hardware is booted. It would be very difficult to detect such malicious code [18]. There is a need to provide a more effective way of detecting and, more importantly, preventing damage by such trojans.

It might be very difficult for a layperson, or even an expert, to notice, in good time, that there is an HT in the firmware of their devices that is leaking sensitive data. Most of these devices are deployed as single entities. This means that once one has access to this device and has managed to compromise it, then they are able to access all the data this device accesses. There is a need to provide a solution that would help protect the hardware devices.

Our research is proposing the use of multilayered hardware to protect confidential data. The multilayered approach will help prevent unauthorized access by ensuring that data can only flow from the offline system holding sensitive data to the internet-facing system. Multilayered Hardware security means that even if an attacker accessed the internet-facing device, the attacker would still need to access the offline devices to be able to gain access to the most sensitive data. The sections that follow explain in detail how our solution is going to achieve enhanced user data protection.

3 Design of the Multilayered Hardware Security Solution

One of the best ways to protect data from leakage is to make it inaccessible from online sources [19]. The design in Fig. 1 below seeks to provide security using multilayered hardware.

The approach is designed on two principles; the first being that the user will utilize Random Pseudo IDs and the second being that the user data with identifying information will be kept off the internet using the Hardware Security solution.

A Know Your Customer Agency will be set up in a given country to register and issue IDs to its citizens. The Agency can be the institution that issues National IDs such

as Passports and National ID cards. This will entail that online platform providers will be prohibited from collecting and aggregating personal user information.

The first step a person who desires to use e-platforms will take is to register with the Agency in that country. They will need to physically visit the agency and share their personally identifying information such as their physical and postal addresses, Mobile and office contact numbers, passport size photos, and any other personal data that will be deemed necessary for differentiating that particular individual from any other.

The Agency will then create a unique record with the information collected from the user. This record of data will be stored on the offline system with the actual personal electronic identifier sitting on the database. What makes the offline system inaccessible is the physical separation created by the Hardware Security Solution between the offline and online system. This will be discussed in detail shortly. The Figure below shows the operations of the proposed solution.

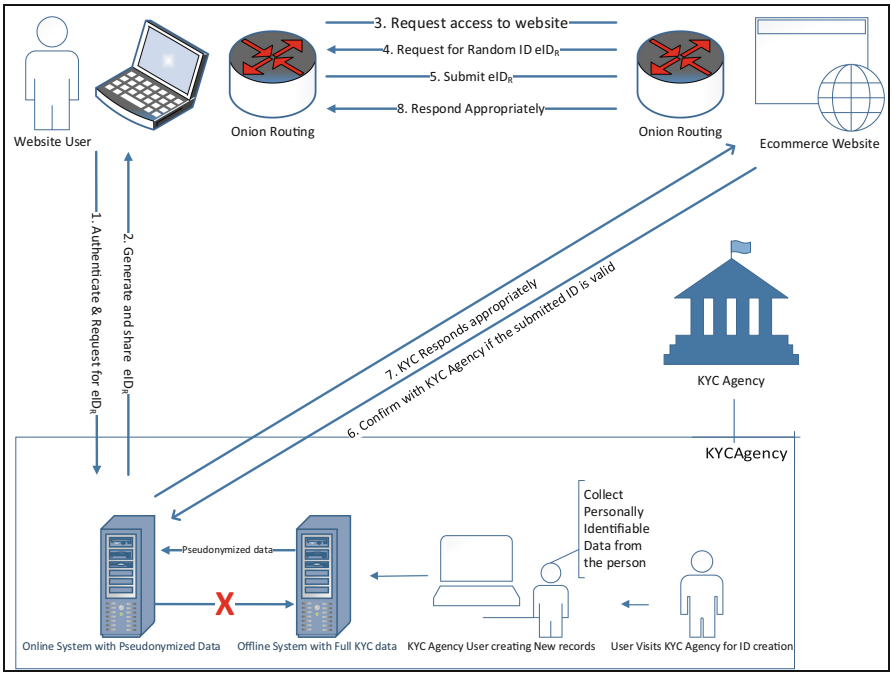


Fig. 1. Know your customer agency operation [20]

To mitigate the risk of sensitive user information being leaked, the original user data created on the offline system is not shared with the online system that interacts with the internet. Instead, data pseudonymization is performed on the user data to transform it into a form that cannot be used to identify the actual owner if that pseudo data were to be leaked online. In addition, data minimization is also carried out on the data to sit online to ensure only necessary data is kept for transaction processing. A predetermined function is used to achieve this. The offline system generates a Pseudo electronic ID,

eID, which is then propagated to the online system. This Pseudo ID is not recorded on the offline system to avoid one inferring the actual owner of data if they happened to have access to both sets of data.

When the Offline System receives the Pseudo ID, it will then create a user record for online transaction processing. Therefore, the user will only have a Pseudo ID kept online but not exposed to everyone. However, even if this ID is leaked inadvertently, the record will not have identifying information to compromise the privacy of the data owner. The diagram below shows the detailed registration process (Fig. 2).

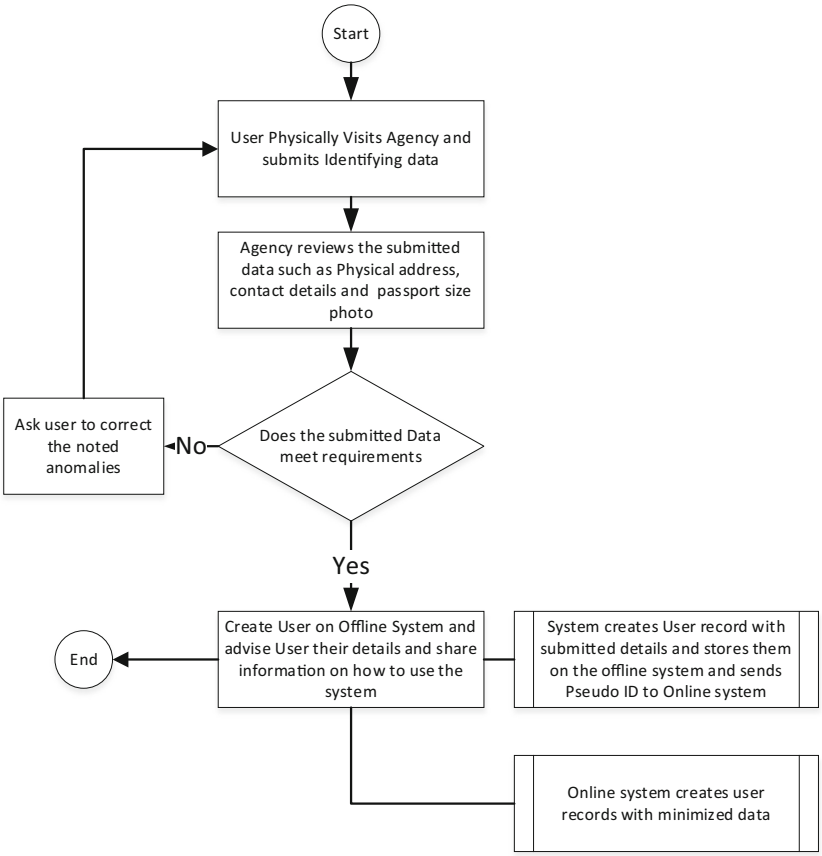


Fig. 2. User registration with KPA

The solution is premised on one key foundation; keeping sensitive data offline to prevent online leakage and unauthorized access to user information. To achieve this, a Hardware solution was designed to operate as follows:

A physical separation between the offline and online systems is the key protection buffer of the sensitive offline data. The Hardware Security System called Data Protector or RMS is the physical separator between the two cyber environments. One end connects

to the internet-facing world while the other connects to the offline Agency's world. The system was designed to only allow one-way-information-flow; from the offline system to the online system. This approach achieves two main objectives; to prevent an unauthorized person from the internet accessing sensitive user data and to prevent malware from being injected into the offline system from the internet [1] (Fig. 3).

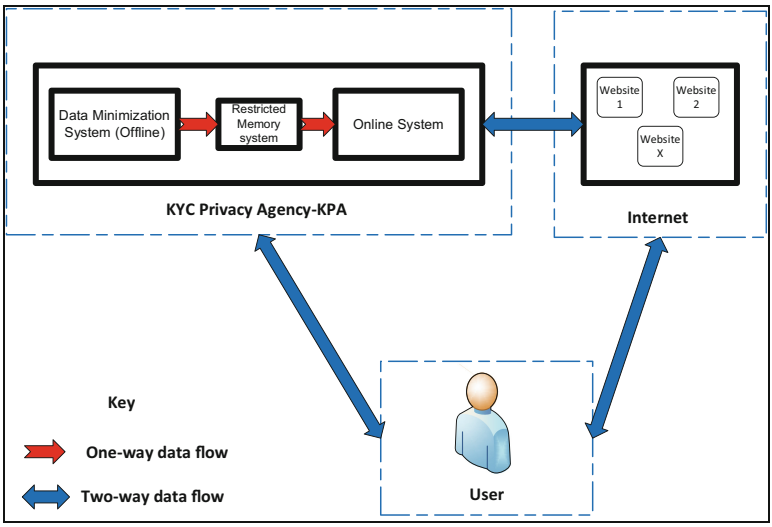


Fig. 3. Hardware security unit

Hardware is susceptible to Trojan infection. This can be planted into the hardware at the point of manufacture and activated when the hardware is in use to leak sensitive user data. To address this challenge, a physical security buffer had to be created using the hardware security unit using two microcontrollers.

The challenge that comes with this solution is data being able to flow from the offline system where sensitive data is stored to the online system that is easier to compromise. To resolve this problem, that is, to prevent huge amounts of data from being leaked online, serial communication was used to connect the two microcontrollers. Moreover, bandwidth was also kept to a minimum only sufficient enough to pass small bursts of data when sending Pseudo IDs to the online system for the creation of online minimized data for transaction purposes. For example, setting a speed of 9600bps would require 10,000 days to send 1terabyte of data. This is very prohibitive. The system can be further enhanced by setting up regular resets of the connection to ensure such attempts are disrupted and the process restarts thereby frustrating the unauthorized users.

To ensure that users are not profiled while online, the Pseudo IDs are not to be used for online commerce. Instead, Random IDs will be generated whenever the user wants to transact online. In other words, users will be anonymous while online to provide privacy. The Agency generated random IDs using an enhanced TOTP algorithm based on the RFC 6238 standard.

The user will generate the required random ID either from the Agency website or using an app. The generated ID will be used to access e-platforms. The platform providers will contact the Agency to confirm the authenticity of the provided random ID before granting the potential user access. This way Service providers won't need to collect identifying data from the user but will draw comfort from the Agency which would have done the necessary KYC due diligence before granting the user access to their system for electronic ID provision.

Despite using Random IDs, user privacy can be compromised to some extent via the profiling of their network devices. To prevent such a compromise, Onion Routing (TOR) can be employed and it hides identifying information about the user [21].

4 Implementation Method

Only the Multilayered Hardware Security Module was put together to implement the proposed method of data protection. The other components used existing approaches.

The Hardware Security Module (HSM) was assembled from power conductors, Virtual and physical ports, emulation monitors, Arduino microcontrollers, and software for Arduino configuration. Simplex serial communication was adopted for simplicity and reliability. In addition, serial communication enabled the use of multiple microcontrollers to create a physical divide between online data and offline data thereby helping to achieve the objective of this solution. The receiver terminal, Tr, of one device would connect to the sender terminal, Ts, of the other. In this case, the Ts, of the Master controller connected to the offline component, was connected to the Tr, of the Slave controller connected to the online component. There was no cable connecting the Tr of the Master to the Ts of the slave controller. Therefore any attempt to send any form of data such as malware to the offline component would fail as there was no provision due to the physical separation by not providing any cable that would enable data to be sent in the reverse direction.

HTs can be implanted into the microcontrollers when being mass-produced to intentionally leak data when deployed for use in various devices. To activate HTs, one will require physical or remote access to the controller [17]. The use of multiple, in this case, two controllers would require one to make multiple hopes before gaining access to the other side of the system. That's, if one wanted to access sensitive data sitting on the offline system, they would need to comprise the controller on the online system and then breach the physical divide before they can access the offline system. For this to be achieved, the cable we intentionally removed to set up a one-way data flow would need to be in place for that to be possible. Meaning one needs physical access to the controllers. This is what makes the solution effective. Figure 4 below shows the setup of the Hardware Multilayer Security Module.

To ascertain the effectiveness of the solution, several Tests scenarios were conducted. Information was sent from the Offline system to the online system to determine whether or not data could be successfully passed from the offline system to the online system via the HSM. Besides, data flow bandwidth was restricted to about 9600bps between the offline and online systems to minimize the amount of data that can be allowed to cross the HSM. Further, data was transmitted from the online system to the offline one to find out if the system could allow data to flow from the online side to the offline side.

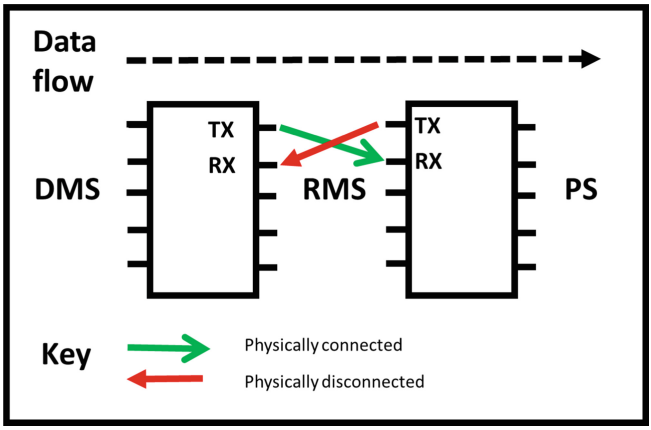


Fig. 4. Multilayered hardware configuration

Additionally, random ID was generated successfully using the modified algorithm. Moreover, an attempt was successfully made to decipher the original user Id from the generated random ID. This is critical in case of fraud by the user.

5 Random Electronic IDs

To help enhance privacy, the Hardware Security solution was coupled with random electronic IDs. These were generated using an algorithm derived from modifying the TOTP standard RFC6238. This is a standard put together by the Internet Engineering Task Force (IETF) to govern the generation of Time-sensitive One-Time-Passwords. To achieve time sensitivity, One of the inputs is the Unix epoch time [22]. Unix Epoch time keeps changing hence the generated OTPs also keep changing. The OTP formulated using the standard is then manipulated with a function using the Pseudo ID as one of the inputs as shown in Fig. 5 below. The idea is to protect user privacy by preventing the online profiling of that user via their static IDs.

The positive of this approach is that if the user commits fraud while transacting anonymously, they can be traced successfully by using reversal algorithms based on the RFC6238 standard.

Another benefit of this approach is that, it will be possible to trace who generated which Random ID at any given time without keeping voluminous records of logs indicating the Random ID generated by who and at what time. By just examining the Random ID, the unique user ID can be deciphered by the Agency which will have a secret Algorithm. The Algorithm used here is only for proof of concept. Each implementation can have its own suitable algorithm.

To further enhance user privacy, the use of Onion Routing (OR) as a complementary security layer is encouraged. Without OR, identifying data such as IP addresses, Mac addresses, and many more can be traced and used to profile a user while they are online thereby compromising their privacy [23].

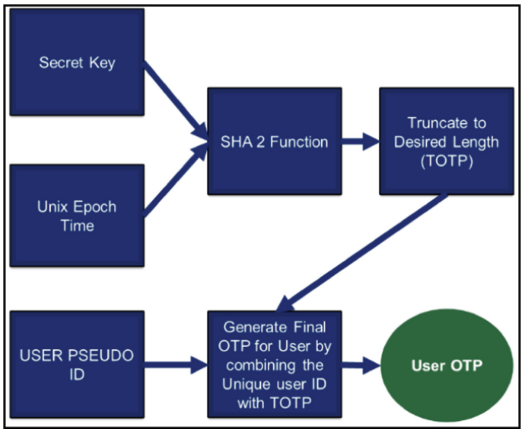


Fig. 5. Generation of random IDs (User OTP) [20]

6 Test Results and Discussions

Table 1. Experiment results

No	Test | Scenario	Results	Desired Result	Overall Result	Status
1	Data Transmission from Offline System to Online System	Data was successfully sent and reached destination	Data should only flow in one direction	Pass	
2	Data Transmission from online to Offline System	Data could not be successfully sent	Data should only flow in one direction	Pass	
3	Bandwidth Limitation Effect	Sending 10GB estimated at more than 100days	Estimated time should be frustrating. Sending 10GB should take less than 1 hour under normal circumstances	Pass	
4	Generation of Random IDs	Generated Random ID with ebededed User ID	Always produce unique Random ID	Possible	
5	Decipher original static user ID	Generate Static User ID kept offline by KYC agency	Always produce unique static ID	Possible	
6	Connect to Master microntroller (Offline) from Slave controller (Online)	Not able to connect	Should not be able to connect	Pass	

Scenario 1 above indicates that data was able to flow from the offline system to the online system. This is the desired state as data should be able to flow in this direction to enable the transmission of Pseudo IDs generated for newly added users. Without this result being achieved, the proposed solution would be a failure as manual transfer of Pseudo IDs is not desired to prevent malware transmission and physical exchange of data media. The figure below shows the screenshot of the results from the experiment (Table 1 and Fig. 6):

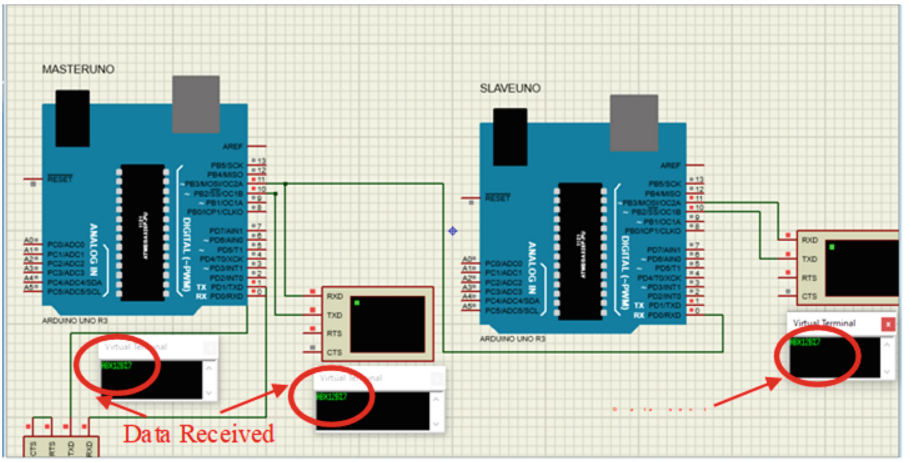
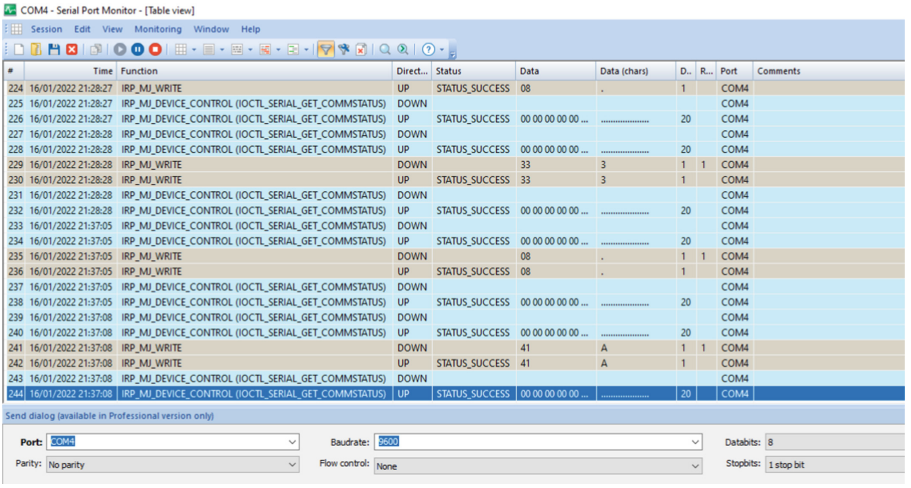


Fig. 6. Test 1 result

Figure 7 below shows data monitored on the serial port being received after transmission. The port monitored was for the online system to determine if data from the offline would reach the intended destination.



#	Time	Function	Direct...	Status	Data	Data (chars)	D..	R...	Port	Comments
224	16/01/2022 21:28:27	IRP_MJ_WRITE	UP	STATUS_SUCCESS	08	.	1		COM4	
225	16/01/2022 21:28:27	IRP_MJ_DEVICE_CONTROL (IOCTL_SERIAL_GET_COMMSTATUS)	DOWN						COM4	
226	16/01/2022 21:28:27	IRP_MJ_DEVICE_CONTROL (IOCTL_SERIAL_GET_COMMSTATUS)	UP	STATUS_SUCCESS	00 00 00 00 00 ...	.....	20		COM4	
227	16/01/2022 21:28:28	IRP_MJ_DEVICE_CONTROL (IOCTL_SERIAL_GET_COMMSTATUS)	DOWN						COM4	
228	16/01/2022 21:28:28	IRP_MJ_DEVICE_CONTROL (IOCTL_SERIAL_GET_COMMSTATUS)	UP	STATUS_SUCCESS	00 00 00 00 00 ...	.....	20		COM4	
229	16/01/2022 21:28:28	IRP_MJ_WRITE	DOWN		33	3	1	1	COM4	
230	16/01/2022 21:28:28	IRP_MJ_WRITE	UP	STATUS_SUCCESS	33	3	1		COM4	
231	16/01/2022 21:28:28	IRP_MJ_DEVICE_CONTROL (IOCTL_SERIAL_GET_COMMSTATUS)	DOWN						COM4	
232	16/01/2022 21:28:28	IRP_MJ_DEVICE_CONTROL (IOCTL_SERIAL_GET_COMMSTATUS)	UP	STATUS_SUCCESS	00 00 00 00 00 ...	.....	20		COM4	
233	16/01/2022 21:37:05	IRP_MJ_DEVICE_CONTROL (IOCTL_SERIAL_GET_COMMSTATUS)	DOWN						COM4	
234	16/01/2022 21:37:05	IRP_MJ_DEVICE_CONTROL (IOCTL_SERIAL_GET_COMMSTATUS)	UP	STATUS_SUCCESS	00 00 00 00 00 ...	.....	20		COM4	
235	16/01/2022 21:37:05	IRP_MJ_WRITE	DOWN		08	.	1	1	COM4	
236	16/01/2022 21:37:05	IRP_MJ_WRITE	UP	STATUS_SUCCESS	08	.	1		COM4	
237	16/01/2022 21:37:05	IRP_MJ_DEVICE_CONTROL (IOCTL_SERIAL_GET_COMMSTATUS)	DOWN						COM4	
238	16/01/2022 21:37:05	IRP_MJ_DEVICE_CONTROL (IOCTL_SERIAL_GET_COMMSTATUS)	UP	STATUS_SUCCESS	00 00 00 00 00 ...	.....	20		COM4	
239	16/01/2022 21:37:08	IRP_MJ_DEVICE_CONTROL (IOCTL_SERIAL_GET_COMMSTATUS)	DOWN						COM4	
240	16/01/2022 21:37:08	IRP_MJ_DEVICE_CONTROL (IOCTL_SERIAL_GET_COMMSTATUS)	UP	STATUS_SUCCESS	00 00 00 00 00 ...	.....	20		COM4	
241	16/01/2022 21:37:08	IRP_MJ_WRITE	DOWN		41	A	1	1	COM4	
242	16/01/2022 21:37:08	IRP_MJ_WRITE	UP	STATUS_SUCCESS	41	A	1		COM4	
243	16/01/2022 21:37:08	IRP_MJ_DEVICE_CONTROL (IOCTL_SERIAL_GET_COMMSTATUS)	DOWN						COM4	
244	16/01/2022 21:37:08	IRP_MJ_DEVICE_CONTROL (IOCTL_SERIAL_GET_COMMSTATUS)	UP	STATUS_SUCCESS	00 00 00 00 00 ...	.....	20		COM4	

Fig. 7. Serial data monitored on serial port

Scenario 2 indicates that data could not be sent from the online system to the offline system. This is the desired result. No form of communication should take place from the internet to the offline system storing sensitive user data. This result enables our physical Hardware security system to protect data from unauthorized access by online users as well as malware attack such as ransomware. These attacks are normally orchestrated by hackers who would be miles away launching the attack using their computers or other

botnets. Preventing access to data from the internet would help stop such attacks from being successful. The figure below shows that the data sent did not reach the intended target (Fig. 8).

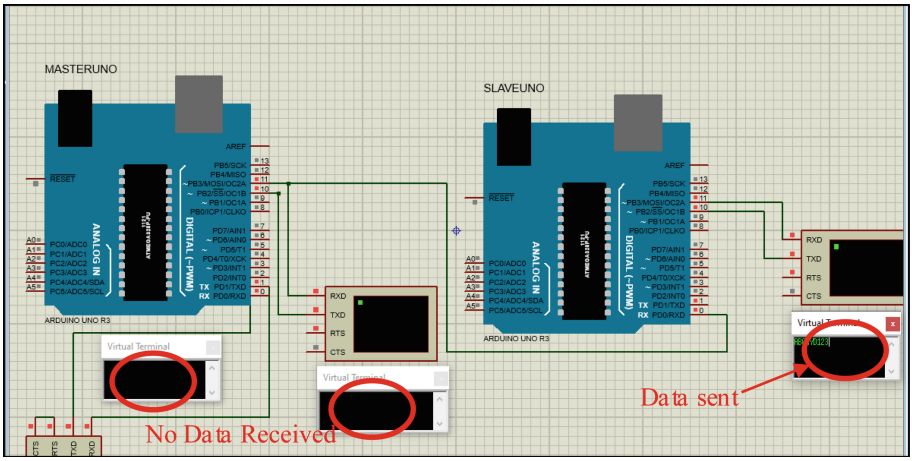


Fig. 8. Test 2 result

Result 3 shows that restricting bandwidth would adversely affect the speed at which data can be transmitted. This is desired to prevent the transfer of huge amounts of data from the offline system to the online system. It would be frustrating for a hacker to wait for 100 days to transmit 10 GB of data. The graph below gives a feel of how long it would take to transmit various amounts of data across the HSM at the given speed (Fig. 9).

Result 4 shows that Random IDs can be generated. This is very important for this solution to achieve its intended objectives. Using static IDs would easily give away the privacy of the user. A continuously changing ID would help hide the profile of the user and hence guard their privacy while online. It would be very difficult to profile a user who keeps changing their IDs when transacting online.

Result 5 is equally important. The main reason Pseudonymization was preferred to Anonymization, was to make it possible to decipher the original ID of the user from the Random ID if the need arises. Some users might take advantage of the fact that they are anonymous while online. They might engage in fraudulent activities believing they cannot be traced. This result shows that despite being anonymous while online, the Agency can decipher the original ID if the need arose. This can help cab fraud by users knowing they would be traced if they misbehaved.

Last, but not the least, result 6 shows a very critical part of the solution. One cannot connect from the online microcontroller to the offline one. This is very vital as one of the main objectives of the solution is to prevent unauthorized access from internet users. Usually, hackers attack from the internet. This goes to show that the HSM can protect our offline system despite the hardware security solution facing the internet being compromised. Multilayered hardware comes into play.

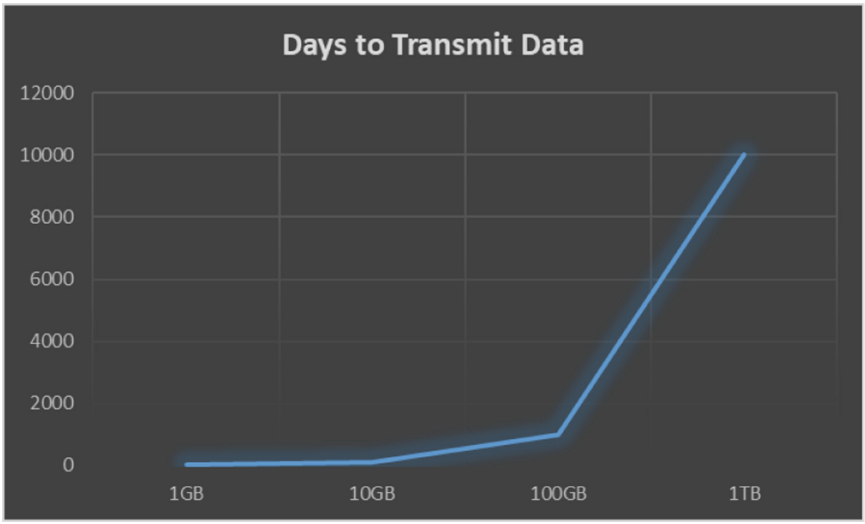


Fig. 9. Data transmission duration

7 Conclusion

Results from the experiments conducted provide reasonable assurance that the Multilayered Hardware Security approach protects user data. Keeping user data offline is the best way to prevent leakage. The use of multilayered Hardware security provides improved protection as the offline component remains safe from attack by internet users even if the online component was compromised. Furthermore, the ability to generate Random IDs gives assurance and comfort to users about their privacy. They can transact anonymously thereby protecting their privacy. In addition, The successful deciphering of the original ID from a random ID gives confidence to the e-platform providers that even if they do not collect user-identifying information before granting users access to their platforms, the users can still be traced successfully if they misbehaved while using their platforms. Further, the ability to decipher IDs from random IDs eliminates the need to store huge volumes of log data about who generated a given random ID. This removes the need to have huge storage for logs as thousands of transactions take place every day. In a nutshell, a combination of existing solutions and new approaches can help enhance how user data can be protected from unauthorized access and online leakage.

References

1. Kangwa, M., Lubobya, C.S., Phiri, J.: Prevention of personally identifiable information leakage in E-commerce via offline data minimisation and pseudonymisation. Int. J. Innov. Sci. Res. Technol. **6**(1), 209–212 (2021)
2. Cona, F.A.: Digital Identity "Personal Dato" US 2019/0333054 A1 (2019)
3. Pawar, H.R., Harkut, D.G.: Classical and quantum cryptography for image encryption decryption. In: Proceedings of 2018 3rd IEEE International Conference on Research in Intelligent and Computing in Engineering, RICE 2018, pp. 1–4 (2018)

4. Hauer, B.: Data and information leakage prevention within the scope of information security. IEEE Access **3**, 2554–2565 (2015)
5. Coppolino, L., D'Antonio, S., Mazzeo, G., Romano, L.: A comprehensive survey of hardware-assisted security: from the edge to the cloud. Internet Things **6**, 100055 (2019)
6. Tan, C.D., Min, F., Wang, M., Zhang, H.R., Zhang, Z.H.: Discovering patterns with weak-wildcard gaps. IEEE Access **4**, 4922–4932 (2016)
7. II-Agure, Z., Belsam, A., Yun-ke, C.: The semantics of anomalies in IoT integrated BlockChain network, pp. 144–146. IEEE (2019)
8. Zhang, P., Alkubati, M., Bao, Y., Yu, G.: Research advances on blockchain-as-a-service: architectures, applications and challenges. Digit. Commun. Netw. (2021)
9. Ye, F., Dong, X., Shen, J., Cao, Z., Zhao, W.: A verifiable dynamic multi-user searchable encryption scheme without trusted third parties. In: Proceedings of International Conference on Parallel and Distributed Systems, ICPADS, vol. 2019-December, pp. 896–900 (2019)
10. Innab, N., Alamri, A.: The impact of DDoS on E-commerce. In: 21st Saudi Computer Society National Computer Conference, NCC 2018, pp. 1–4 (2018)
11. Zhan, J., Fan, X., Cai, L., Gao, Y., Zhuang, J.: TPTVer: a trusted third party based trusted verifier for multi-layered outsourced big data system in cloud environment. China Commun. **15**(2), 122–137 (2018)
12. Ribeiro, S.L., Nakamura, E.T.: Privacy protection with pseudonymization and anonymization in a health IoT system: results from OCARIoT. In: Proceedings - 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering, BIBE 2019, pp. 904–908 (2019)
13. Gao, C., Luo, L., Zhang, Y., Pearson, B., Fu, X.: Microcontroller based IoT system firmware security: case studies. In: Proceedings - IEEE International Conference on Industrial Internet Cloud, ICII 2019, no. ICII, pp. 200–209 (2019)
14. Skorobogatov, S.: How microprobing can attack encrypted memory. In: Proceedings - 20th Euromicro Conference on Digital System Design, DSD 2017, pp. 244–251 (2017)
15. Ande, R., Adebisi, B., Hammoudeh, M., Saleem, J.: Internet of things: evolution and technologies from a security perspective. Sustain. Cities Soc. **54**, 101728 (2020)
16. Bettayeb, M., Nasir, Q., Talib, M.A.: Firmware update attacks and security for IoT devices survey. In: ACM International Conference Proceeding Series (2019)
17. Ali, S.S., Chakraborty, R.S., Mukhopadhyay, D., Bhunia, S.: Multi-level attacks: an emerging security concern for cryptographic hardware. In: Proceedings - Design, Automation Test in Europe, DATE, pp. 1176–1179 (2011)
18. Adelstein, F., Stillerman, M., Kozen, D.: Malicious code detection for open firmware. In: Proceedings - Annual Computer Security Applications Conference, ACSAC, vol. 2002-January, pp. 403–412 (2002)
19. Kangwa, M., Lubobya, C., Phiri, J.: Enhanced protection of ecommerce users' personal data and privacy using the trusted third party model. In: Proceedings of the 18th International Conference on e-Business - ICE-B, pp. 116–126 (2021)
20. Kangwa, M., Lubobya, C., Phiri, J.: Protection of personally identifiable information and privacy via the use of hardware and software. In: Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2021, 20–22 October 2021, Hong Kong, pp. 75–81 (2021)
21. Jardine, E., Lindner, A.M., Owenson, G.: The potential harms of the Tor anonymity network cluster disproportionately in free countries. Proc. Natl. Acad. Sci. U. S. A. **117**(50), 31716–31721 (2020)
22. M'Raihi, D., Machani, S., Pei, M., Rydell, J.: Internet Engineering Task Force (IETF): Request for Comments: 6238 (2011)
23. Winter, P.: Enhancing Censorship Resistance in the Tor Anonymity Network Enhancing Censorship Resistance in the Tor Anonymity Network. Karlstad University (2014)

Alternative Approach to Rounding Issues in Precision Computing with Accumulators, with Less Memory Consumption: A Proposal for Implementation

Roy P. Gulla(✉)

555 Highland Dr, Mesquite, NV 89027, USA
rgullape@gmail.com

Abstract. Recent development efforts with the newest version of Unum formatting, posits, have coincided with the continued progress of the new P1788 standard which includes dot products as one of the five fundamental arithmetic operations. These trends have led to increased interest in large scale implementation of accumulators for numerical computation [10], and the creation of encoding and decoding schemes between IEEE float types and posits [5]. Here an alternative in light of these new developments is presented in a way to incorporate the minposis set of Gustafson's Posits formatting, which implements an alternative weighted circuit for fractional bits in his system, capitalizing on inherent memory latency in newer hardware circuitry. A model program template is also touched upon in the new expanded outline of this formatting introduced in [6], proposing the incorporation of fine-grained parallelism, at instruction level.

Keywords: Number representations · Roundoff error · Complex instruction set computer (CISC) · Instruction level parallelism · Fine grained parallelism

1 Introduction to the Approach

1.1 Introduction

Low level hardware designs are rapidly undergoing a major sea level change in response to the new emerging technology surrounding quantum computing. In order to control for some of the inherent volatility in these new products, new formatting for precision computing is taking shape with several performance constraints enacted in order to limit the nearly unlimited bounds for memory consumption in these new devices.

Newer models of hardware and system architectures are paving the way to new models of memory management, which in turn have lead to multilayered computational pathways in hardware circuits. And even given the current trend towards offloading of memory in these newer heterogeneous architectures, in the case of commonly stored precise scientific constants, it is the loading and storing of these large constants which can prove expensive on processor performance.

S. Ao et al. (Eds.): IMECS 2021, *Transactions on Engineering Technologies*, pp. 116–121, 2023.
https://doi.org/10.1007/978-981-19-7138-9_10

Very recently, John Gustafson's POSIT formatting has found some implementations in a square root and divide circuit described in [3], and a new standard creation is underway to directly implement the direct dot product as a fifth arithmetic operation.

In the first of these developments, power consumption is highlighted as a primary concern. And among the differences of the posits formatting than that of the floating point type, the one which might be most attractive to memory management is that of the optional fractional bitfield. So while the IEEE float type can spend the majority of time in determining the right exponent range for a scientific constant, the posits option to exclude the fractional bitfield allows for some more efficient pathways in a high precision computing circuit, so long as the traditional encoding of the floating point standard is abandoned.

In the first portion of the paper the author will describe this proposed arithmetic circuit in light of the new quantum dot cellular automata application implementation with one-bit full adders described in [1], in a way that the instruction set is not in any way bloated with extraneous operands (i.e. store and/or load) (Fig. 1).

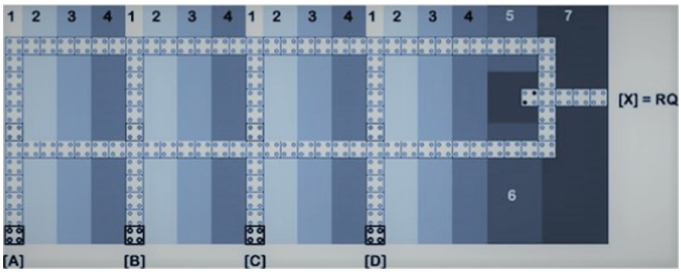


Fig. 1. A parallel to serial QCA w/ majority gate design, which is implemented in [1].

In the final section of this paper a model prototype with minimally accessible instruction sets for global registers will be introduced for implementation as a complex instruction set computer architecture.

1.2 The Problem of Rounding with Variable and Large Word Sizes

The proposed alternative to the typical fractional bitfields of floating point types in the proposed formatting bypasses the issue of loading and storing of intermediate values in memory, which largely contribute to both precision and performance issues in the rounding stages of floating point computation. These are costly operands for comparisons which result in an intermediate difference term for each comparison performed, and it is this comparison difference term which amazingly can require a higher level of precision than the final representation of the number [5].

A simple example of a fundamental issue with binary numbering systems shows how memory constraints can seriously impede pursuits of high precision levels. In the floating point representation of .1, by the tenth subunitary bit, the represented number is still .0008 away from the desired value. And eight bits later, it has overshot this number.

One common example of a non-exactly representable number is presented here in order to introduce the proposed formatting approach of this paper intended to correct for this unpredictable behavior in floating types. In the formatting representation, a repeating minposis, or subunitary, term used in the calculation of the fraction is weighted with integer values. (In the final portion of this paper these weights are encoded as repeated read and shift instructions.)

Decimal Value	**.076923076923..(1/13)**
Alternate Expanded form	**1/13 = (1/16)(1+3/13)** **=(1/16)(1+3(1/16) (1+3/13))** **=(1/16) (1+3(1/16)(1+3(1/16) (1+3/13)))** **etc.**
Minposis Fractional Bits Binary Representation	**00010**

Fig. 2. It is worth noting here that the repeating digits in the lower table's rows hint at a costly string representation, which is avoided here. But the use of weighted bits is important in the sections to follow.

1.3 The Number Rings of These Non-exactly Representable Numbers

Here we intend to introduce a formatting which more efficiently deals with these fractional bit fields that vary in width for each implementation. Particularly, those numbers which can be represented in the following geometric series form fit the proposed formatting:

$$\frac{1/2^{i}}{1-1/2^{i}}, i \in N$$

The approach taken with the these numbers here hopes to present one of the very pronounced differences between the posit's minposis bitfield and the subunitary digits of IEEE floating point types, namely the optional fractional bitfield, and to expand on a possible implementation of these bits.

With a repeating bit pattern as that shown in the above Fig. 2, the floating point fractional bits are both a source of memory and precision issues. The issue is that the repeating remainder in the binary numbering scheme cannot be transformed into binary form, so an iterative expansion approach is taken as shown in Fig. 2 above and Table 1 below, utilizing a feedback adder circuit such as the one in Fig. 3.

1.4 A Common Floating Point Constant with a Weighted Bit Representation

An alternative representation of the natural log base is demonstrated here as a final example for the proposed alternative formatting, incorporating a weighted representation of bits to calculate its value in the accumulator, as opposed to the traditional and costly approach of loading its stored value. A simple description of the iterative scheme used in the computation of its value in real time is given here:

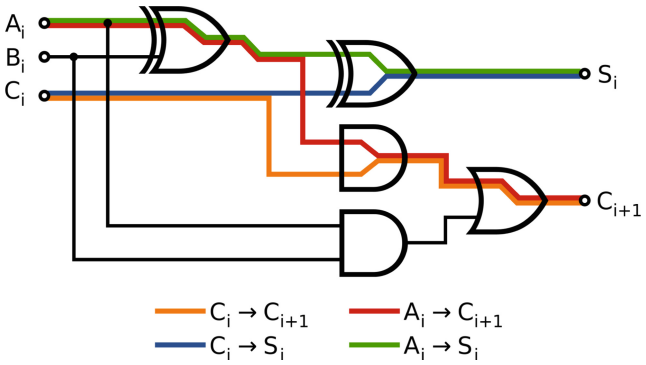


Fig. 3. The design of the serial full adder circuit architecture. In this paper A_i and B_i represent the same bit operand

- Iterate through the subunitary bits, and,
- Repeat the adding of the current bit until the value overshoots, at which point discard the previous summand, and continue the iteration.

This weighted bit representation of the natural log base is shown in Table 1 to machine level precision, and includes the extra 49^{th}, 52^{nd}, and 54^{th} subunitary bits, which can more easily be implemented when no fractional bits are included in the format.

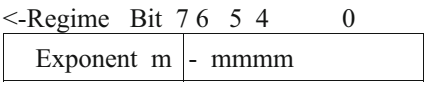
Table 1. The natural log base, with the weighted bit formatting

SET SIGNIFICAND BITS	2,4,5,7,8,9,10,11,12,17,20,21,22,23, 24,25,26,27,29,30,31,32,33,34,35,3 6,37,38,39,40,42,43,44,45,46,47,49, 52,54
CORRESPONDING BIT WEIGHTS	2,2,2,2,2,2,2,2,2,2,2,2,4,2,2,2,4,4,2,2,2,2,4,2,2,2,8,4,4,8,8,8,4,4,4,2,2,2,2,1
SUPERUNITARY BIT	1

2 Proposal for Implementation of the Approach

In [5], the difficult case of a non-exactly representable number in a formatting system with large storage constraints is clearly demonstrated. Summarizing the point here, these constraints place limitations on allocating enough memory for both wide exponent bit fields and wide fractional bit fields. The solution proposed here is to sidestep the wide (and in Gustafson's system, optional) fractional bit field by utilizing a single set bit in a QCA feedback loop to compute our minposis values.

Implementing these feedback loops, much as feedback loops are in [11], the intention is to wrap pipelined arithmetic circuits into posit decoders/encoders, particularly for those numbers where the fractional portion of the posits number is a significant enough part of the representation. Since this portion of the posit number could be computed



m-mmmm (fractional bits alternate encoding):
0- 0010:
0- 0010:
0- 0010:
0- 0010:

Fig. 4. An illustration of copies of the proposed alternative to the fractional bitfield (with a word size of one half byte) necessary for our offloaded compute instructions is shown. Note the size of the exponent or the regime fields is not dealt with here, as the format is proposed to be implemented in a vector processor. Here the weight value for our arithmetic circuit is 4. The multilayered QCA architecture could implement each copy.

simultaneously with the regime and/or the exponent in the QCA architecture described in [11], no additional performance costs would be incurred with this alternative to the posits' fractional field. And a suggested implementation of these arithmetic circuits utilized in place of the fractional bits would be with a complex instruction set as that of the arm neon processor.

As implementations of these numerical types go through layers of architecture to higher levels, what once was a double type can become mixed with integer and decimal types, and precision suffers under these race conditions. And when these global (general purpose) registers are implemented in high performance compute nodes being shared on a large cluster of systems, either data corruption, or performance fall-off becomes an issue.

Therefore the design choice for the implementation here is to avoid altogether the use of global general purpose registers, and in particular to avoid the implementation of the proposed instruction sets in any higher level language wrappers, and instead pursue the course of instruction level parallelism. This both enables the avoidance of global general purpose registers, and implements more local and effective usage of memory, making global memory registers unnecessary. And as architectures trend more and more towards multithreaded ones, the larger blocks of instructions which are mapped into ISBs should be grouped together at the earliest point in the control flow graph anyhow (i.e. in the closest sequential proximity to each other in the code block). In the case of dataflow architectures where token matching is a crucial part of the compilation process, this can prove vital to processor performance. Therefore we attempt not to branch anywhere outside our current block of compute instructions.

In order to do so, we suggest a code design already in place in some legacy and most newer arm neon processors, with recursive C macro calls to generate the product of sums circuitry.

```
#define D2 (X) D1 (D1 (X) D(D(D() ) ) )
```

3 Suggestions for Further Development of the Model

However fine grained we design our proposed parallelism for the arithmetic circuits used in these computations, and restrict our memory space consumption, this limiting of memory consumption will in no way damper our QCA architecture's performance.

For the purpose of ease in the testing of the model, we suggest a word in our vectorized system with one set bit (as shown in Fig. 2 and Fig. 4) in the initialization stages. Following this line of thought, in-wire processing described in [11] is a suggested way of incorporating one bit blocks of data in the circuits, with multiple fetch cycles (by means of a feedback circuit as described in [11]) being used to "weight" the bits.

Complex instruction set computers such as those produced by Intel have shown how instruction level parallelism can be successfully implemented to maximize the performance of their supercomputing architectures. As branching is already an available mechanism engineered into modern arithmetic circuits, the proposal is to implement such instruction level parallelism at the most localized level, by means of the branching mechanism alluded to in Sect. 2 above.

References

1. Angizi, S.F.: A structured ultra dense QCA one-bit full adder cell. Quantum Matter **4**, 1–6 (2015)
2. Bian, T.D.: Context aware quantum simulation of a matrix stored in quantum memory. Quantum Inf. Process. **18**(12), 1–12 (2019)
3. Cheng, S.L.: Posit arithmetic hardware implementations with the minimum cost divider and square root. Electronics **2020**(9), 1622 (2020)
4. Enbody, R.J.: Performance monitoring in advanced computing architecture. In: WCAE 1998: Proceedings of the 1998 Workshop on Computer Architecture Education, Barcelona, pp. 17-es. ACM, 1998 June
5. Florent de Dinechin, L.F.-M.: Posits: the good, the bad and the ugly. hal-01959581v3 (2019)
6. Gulla, R.P.: Two alternative approaches to rounding issues in precision computing with accumulators, with less memory consumption. In: Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2021, 20–22 October 2021, Hong Kong, pp. 240–243 (2021)
7. Gustafson, J.: The End of Error: Unum Computing. CRC Press, Boca Raton (2015)
8. Hittinger, J.L.: Universal coding of the reals: alternatives to IEEE float. In: CoNGA 2018: Proceedings of the Conference for Next Generation Arithmetic: March 2018, Singapore, pp. 1–14. ACM (2018)
9. Johnson, J.: Rethinking Floating Point for Deep Learning. arXiv:1811.01721 (2018)
10. Koenig, J.: A hardware accelerator for computing an exact dot product. In: 2017 ARITH Symposium on Computer Arithmetic, London, pp. 114–121. IEEE (2017)
11. Kogge, P.N.: Exploring and exploiting wire-level pipelining in emerging technologies. In: ACM SIGARCH Computer Architecture News. ACM SIGARCH Computer Architecture News, pp. 166–177 (2001)
12. Kumar, R.P.: A new compiler for space-time scheduling of ILP processors. Int. J. Comput. Electr. Eng. **4**(3), 536–543 (2011)

Author Index

S. Ao et al. (Eds.): IMECS 2021, *Transactions on Engineering Technologies*, p. 123, 2023.
https://doi.org/10.1007/978-981-19-7138-9