

# Battery Charging System Using Sun Tracker with Dual-Axis Based on the Internet of things.

I Wayan Eka Krisna Putra  
Department of Information  
Technology  
Udayana University  
Badung, Bali, Indonesia

I Nyoman Piarsa  
Department of Information  
Technology  
Udayana University  
Badung, Bali, Indonesia

Kadek Suar Wibawa  
Department of Information  
Technology  
Udayana University  
Badung, Bali, Indonesia

---

**Abstract:** Sun Tracker System is a device which developed to utilized the sunlight energy maximally by following the movement of sunlight. This sun tracker system is designed because the application of solar cells is still mostly placed statically so that the absorption of sunlight obtained is not optimal. In this system, we use four mini solar system cells that can accept 6 voltages. Solar cells obtain voltage values using a voltage sensor, this voltage sensor is only able to read a maximum voltage of 25 Voltage. The sun tracker system aims to make it easier to determine the voltage position based on the voltage value received by voltage sensor based on the Arduino Uno microcontroller. The result of this system is to be able to find the maximum voltage that will be displayed on a website using the ESP8266 module to determine the amount of voltage received by the solar cell and to know the location of the angular position when receiving voltage.

**Keywords:** Sun Tracker; Solar cell; Arduino; Motor Servo; ESP8266

---

## 1. INTRODUCTION

The necessary energy is never run out. In this cause, the utilization of fossil is able extant. In addition, fossil energy has adverse impacts such as the environmental pollution of water and land is counted 2,200 in Indonesia villages confront the environmental pollution of land and water. The utilization of energy is also increasing bad effects such as global warming, the alteration of extreme climate, etc. This cause, renewable energy is needed to produce enough energy and enough energy and doesn't have a negative impact on the surrounding environment.

This renewable energy is an energy that starts from the sustainable use of natural energy, for example, is sunlight. The utilization of sunlight has changed into the sun's rays into electrical energy. This statement is able to use solar cells. A solar cell is a tool that can be used as a tool for renewable energy.

The solar cell is a device that has several cells and various types of the component to convert sunlight into electrical energy. The utilization of solar cells has been used in several developing countries. the utilization of solar cells has been used for industrial purposes. But in the installation of this solar cell is manually that is by way of static or in a state of silence. This can be overcome by using a sun tracker system on solar cells.

Sun Tracker is a tool that will developed with the internet of things technology where the sun tracker can follow the sun's movements. This can be done to optimize the absorption of solar energy and can store the voltage into the battery to the maximum. The utilization of solar energy in its development will produce electrical energy and become a backup energy source for the future.

The working principle of this system is sunlight coming from one corner leads to the system then the sun tracker will follow the movement of sunlight. This system requires four mini solar cells and four voltage sensors that are used to read

voltage values obtained from sunlight. This value will be used as a reference to run the servo motor to a position that has an average or maximum voltage. Therefore, the sun tracker that has been designed can increase the maximum absorption of solar energy to become renewable electricity.

## 2. LITERATURE REVIEW

### 2.1 Internet of Things

Internet of things is a technology that uses the internet to control and communicate with various other devices. The Internet of things aims to get the benefits of internet connectivity that is connected continuously. The capabilities that are able using data sharing, remote control, and others.

In essence, the Internet of things leads to a component that can be uniquely identified to represent virtually in an Internet-based structure. The word Internet of things was originally proposed by Kevin Ashton in the 1999 era and began to be known through the Auto-ID Center research held at MIT. Based on analysis from the McKinsey Global Institute, the Internet of things is the latest technology that allows users to interact with machines, tools, and other electronic objects with the help of network sensors and actuators to enable machines to get the latest information that they get as desired.

The Internet of things works by using a programming code where each command of the programming code will produce an interaction between fellow actuators that are mechanically connected without any human intervention at any distance. An Internet network connection is a connection between the two interactions of the actuator, while the user only has the duty to be the maker and manager of the operation of the tool directly.

### 2.2 Sun Tracker

Sun Tracker is a tool that will be made with the internet of things technology where this tool will later do the efficiency of the voltage on the solar cell by following the movement of sunlight to maximize the sunlight obtained from the solar cell.

The sun tracker system is useful for accumulating the absorption of sunlight energy for sustainable processes, and with the most consistent balance as the Sun's position shifts with the seasons. Except, the greater the absorption rate used, the search for sunlight energy becomes more stable, because the proportion of energy coming from direct sunlight is higher, and areas that have stable energy will take precedence.

### 2.3 Arduino

Arduino is a single-board microcontroller that is open-source, derived from the wiring platform, which is structured to facilitate the use of actuators in various fields. The hardware has a VAR atmospheric processor and software that uses its own programming language.

Arduino is an open-source hardware platform aimed at anyone who wants to make prototypes of interactive electronic equipment based on hardware and software that are easy to learn and easy to use. The Arduino microcontroller is programmed using the Arduino programming language which has syntax similarities with the C programming language. Because it is open-source, users can download the Arduino hardware scheme and build it.

Arduino has a connection with the ATmega microcontroller formed by Atmel as a base, but there are individuals or companies that make Arduino imitations using other microcontrollers and remain compatible with Arduino at the hardware level. Using flexibility, the program code is uploaded via the bootloader although there are several options for bypassing the bootloader and using the downloader to program the microcontroller directly through the ISP port.

### 2.4 ESP8266

Esp8266 is a Smart on Chip (SoC) device that is formed in small size and uses a few external circuits. Esp8266 can communicate via wifi infrastructure that uses IPv4, TPC / IP, and HTTP protocols. The processor used on the Esp8266 is the Tensilica L106 diamond series, whose speed is 32-bit and has an on-chip SRAM. Esp8266 has wifi radio, CPU, Memory, Flash, and peripheral interfaces. Esp8266 also has the ability to be used stand-alone or as an access point for a microcontroller.

### 2.5 Servo Motor

Servo motor is a Direct Current motor device that can do feedback on the close position of the motor which will then be explained again into the control circuit contained in the servo motor. Servo motors consist of several gear circuits, potentiometers and control circuits. The potentiometer has a function used to ensure the angle limit of the servo motor movement. Although the angle of the axis of the servo motor is set based on the results of the pulse value sent through the microcontroller of the servo motor cable .

### 2.6 Voltage Sensors

The voltage sensor is a sensor produced by suppressing the resistance value of an electrical circuit. The voltage sensor takes a voltage value from an electrical circuit that is read using a microcontroller to determine the value of the voltage received by the electrical circuit.

### 3.7 Solar Cell

A solar cell is the role of electricity that can convert sunlight into electrical energy. Solar cells described as photovoltaic in

addition to electrical energy solar cells can also be used as a photodetector (infrared detector) or detect electromagnetic radiation. The characterize of Solar cells on absorbing electromagnetic radiation that is received as a semi-conductor that can receive photons from sunlight and convert them into electricity.

Most solar cells are made of pieces of silicon coated with special chemicals to form the basis of these solar cells. Solar cells basically have the smallest thickness of about 0.3 mm which is divided from several semiconductor materials, each of which is found as a positive type semiconductor and a negative type semiconductor as shown in the following Figure 1.

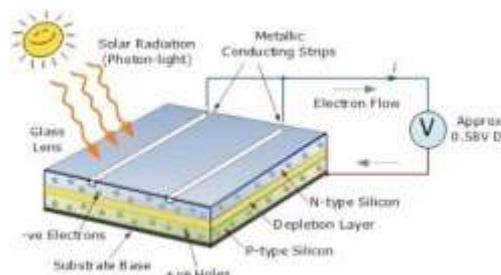


Figure 1 Solar cell

(Source:<https://i0.wp.com/www.electricalengineering123.com/wpcontent/uploads/2018/08/construction-of-solar-cell.jpg?resize=452%2C250&ssl=1>)

Figure 1 is a way of working solar cells that can convert directly to sunlight that is converted into electrical energy. Solar radiation is referred to like the main character to be able to maximize the amount of potential sunlight energy received by solar cells to the earth, in addition to being used to produce electrical energy, the energy produced by the sun can also optimize its thermal energy through the solar thermal system.

## 3. RESEARCH METHOD

In this study, a battery charging system uses a sun tracker with Dual-Axis based on the internet of things that were developed to be able to produce renewable energy. This system consists of a voltage sensor to detect the voltage value received by the solar cell and the servo motor is used to drive the solar cell. The system is designed using four mini solar cells that can receive a voltage of 6V. This system uses a battery as the storage of electricity generated by solar cells. The sun tracker system can be used to be able to monitor the voltage received by the solar cell and can find out the voltage position with a maximum voltage value which will then be displayed on the website. Following is an overview to build a battery charging system using a sun tracker with dual-axis based on internet of things as follows:

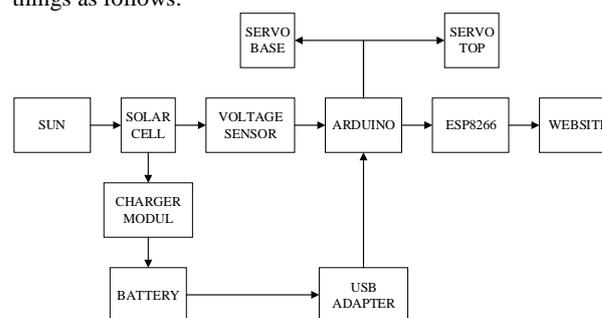


Figure.2 General Description

### 3.1 Concatenation Analysis

Concatenation analysis is a component design that begins by determining the tools is used based on existing theories in accordance with the literature review is used in this study. The following are tools and materials used in designing battery charging systems using a sun tracker with dual-axis internet-based things as follows:

**Table 1. Tools and Materials**

No	Tools and Materials	Function
1	1 Arduino Uno Module	a system controller module with the Arduino Uno microcontroller.
2	1 USB Adapated	a liaison with media that requires power.
3	2 Servo Motors	a solar cell drive.
4	4 Pieces of Solar Panel 6 V	a component that can convert sunlight into electrical energy.
5	1 Breadboard	an electronic circuit that connects with other components.
6	4 Voltage Sensors	Set the voltage from Arduino to solar cell.
7	1 Soldering iron	to installing the components.
8	Cable Jumper	a link between components.
9	2 Battery Holder	a battery co.
10	2 Battery 18650	a power of storage.
11	1 Module Wifi ESP8266 01	a link between solar panels and the website.
12	1 Module USB Charger	a connector of the solar panel to the battery.
13	4 Dioda	rectification of electricity.
14	3 LED (Red, Yellow, Orange)	indicators that indicate the system is running.
15	3 Resistor 330 Ohm	as current dividers in the circuit.

Table 1 is a tool and material used in designing this study which was implemented as an initial step to build a battery charging system using a sun tracker with Internet-based Dual-Axis of things.

### 3.2 Block Diagram

The results of the circuit analysis are translated into a circuit model design or layout scheme by designing a system block diagram. The following is the design of a block diagram to build a sun tracker with dual-axis based on the internet of things as follows.

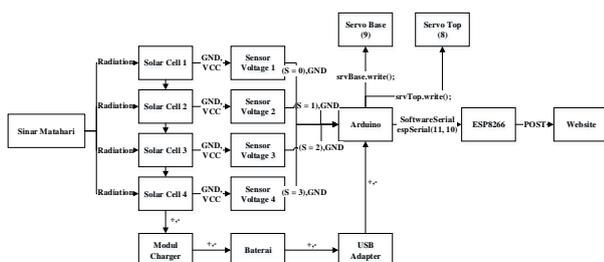


Figure.3 Block diagram of the system

Figure 3 is a block diagram of the system, the solar cell receives radiation from the sun and then produces a voltage

which will be received by the voltage sensor. Voltage sensor will accept the voltage generated by the solar cell to determine the value of the voltage received by the solar cell, the results of the sensor are used as a reference to drive the servo motor that is connected to Arduino, then the results of the voltage sensor and servo motor will be displayed on the website using the WiFi internet network.

### 3.3 Hardware Design

Hardware design is the stage of the shape of hardware that is used both in gluing to soldering of the components to be used when assembling the entire system. Here is a hardware design on a battery charging system using a sun tracker with Dual-Axis based on internet of things is as follows.

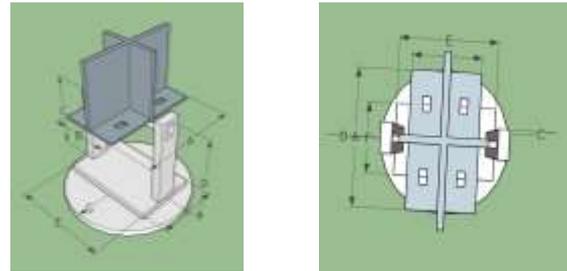


Figure. 4 (a) Side View, (b) Top View

The hardware design used was designed using the SketchUp 2018 application. This prototype has a size to meet the needs needed in making this research system for A in the picture has a size of 23 cm, B with a size of 13 cm, C with a size of 20 cm, D with a size 24 cm, E 23.5 cm, F 7 cm, and G 31.2 cm in diameter.

### 3.4 Schematic Design

The results of the hardware design that has been done then the system are translated into a schematic design that functions to connect the device to other devices. The following is a schematic or wiring design on a battery charging system using a sun tracker with Internet-based Dual-Axis of things is as follows.

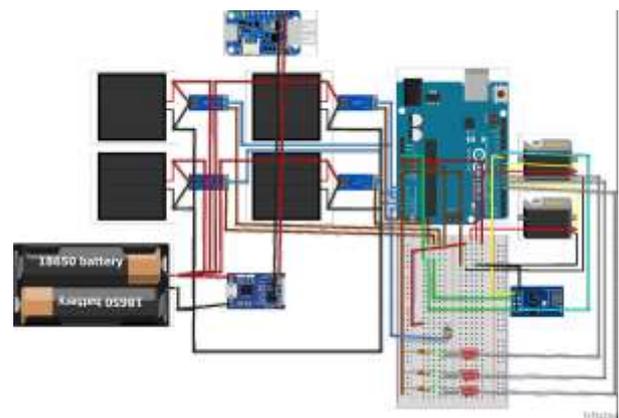


Figure. 5 Schematic Design

### 3.5 Mockup Design

The design of a mockup on a website will not be separated by the design of the User Interface (UI) mockup that can provide a real picture of a website that, if later has been applied as a

program for monitoring. The following is a form of a mockup in this research system as follows.

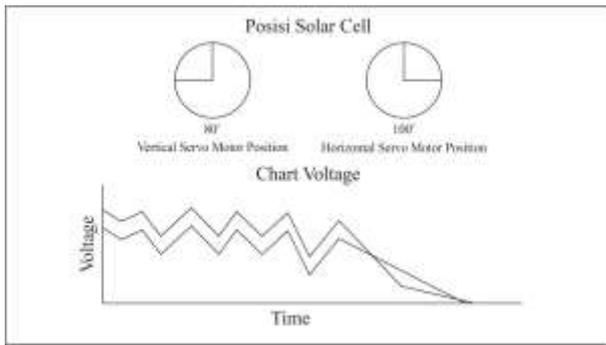


Figure. 6 Mockup System

Figure 6 is a mockup website design that will later display information about the graph of the amount of voltage, vertical servo motor position, horizontal servo motor position and based on the time the data is obtained.

## 4. RESULTS AND DISCUSSION

### 4.1 System Implementation

The system implementation is carried out to implement the design of the system that has been designed to build a real prototype of the system so that the system is ready to use. The following is the hardware design of the prototype battery charging system using the Sun tracker with Dual-Axis based on the internet of things as follows.



Figure. 7 Result of hardware design

Figure 7 is the result of hardware design on the battery charging system using a sun tracker with Dual-Axis based on the internet of things. This prototype uses four mini solar cells, four voltage sensors, 18650 battery type, Arduino Uno microcontroller, two servo motors, and an ESP8266 wifi module. To process the system charger using a charging module, this charging module is equipped with an indicator light to indicate the battery is charging by indicating a red LED (Figure 8) and the battery is fully charged marked with a blue LED (Figure 8).

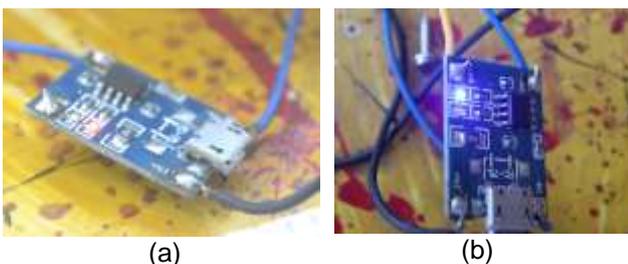


Figure. 8 (a) Charging the battery, (b) The battery has been charged

### 4.2 Testing Tool

Testing the tool is the most important part to find out whether the system has been running according to predetermined standards. Testing the battery charging system using a sun tracker with Dual-Axis based on the internet of things is done by testing the voltage sensor, testing the movement of the servo motor, and testing the wifi with the ESP8266 module. The following detailed explanation of the testing in this study is as follows.

#### 4.2.1 Voltage Sensor Testing

Voltage sensor testing is performed to determine the amount of voltage received by the solar cell by using a voltage sensor. This test uses four solar cells and a voltage sensor to determine the voltage value received by the solar cell. Here is how the voltage sensor testing works as shown in Figure 10 below.

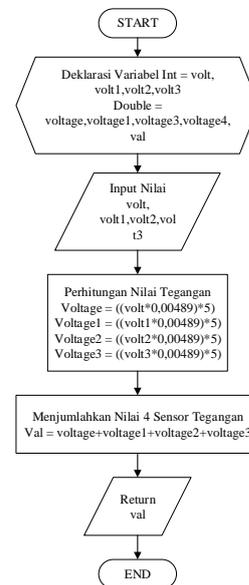


Figure 9 How the Voltage Sensor Test Works

The results of voltage testing using a voltage sensor can be seen using the Arduino serial monitor application in Figure 10.

```
Sudut Base (34) : (0.34V ,0.29V ,0.83V ,0.49 V ) = 1.96
Sudut Base (35) : (0.32V ,0.29V ,0.56V ,0.44 V ) = 1.61
Sudut Base (36) : (0.32V ,0.27V ,0.17V ,0.46 V ) = 1.32
Sudut Base (37) : (0.34V ,0.29V ,0.27V ,0.46 V ) = 1.37
Sudut Base (38) : (0.32V ,0.27V ,0.42V ,0.46 V ) = 1.47
```

Figure 10 Test Results of Voltage Sensors

#### 4.2.2 Servo Motor Movement Testing

Servo motor movement testing is done to determine the movement of the solar cell. Movement testing is used by using 2 servo motors as vertical servo motors and horizontal servo motors. The steps to test this movement are done by determining the initial position, namely the vertical position and horizontal position, which then moves to follow the direction of the sun's rays. Following the results of testing the movement of horizontal servo motors and vertical servo motors can be seen in the Arduino serial monitor in Figure 11 and Figure 12.

```
Sudut Base (96) : (4.69V ,4.72V ,5.33V ,5.23 V ) = 19.98
Sudut Base (97) : (5.09V ,5.06V ,5.55V ,5.16 V ) = 20.95
Sudut Base (98) : (4.77V ,4.82V ,5.28V ,5.01 V ) = 19.88
Sudut Base (99) : (5.01V ,4.79V ,5.28V ,5.01 V ) = 20.10
Sudut Base (100) : (4.89V ,4.79V ,5.28V ,5.01 V ) = 19.98
Base Position Choice(36) : 22.57
```

Figure 11 Results of horizontal servo motor movements

```
Sudut Top (176) : (5.04V ,5.31V ,5.92V ,5.82 V ) = 22.08
Sudut Top (177) : (5.38V ,5.60V ,6.01V ,5.87 V ) = 22.86
Sudut Top (178) : (5.06V ,5.31V ,5.67V ,5.57 V ) = 21.61
Sudut Top (179) : (5.04V ,5.33V ,5.99V ,5.82 V ) = 22.27
Sudut Top (180) : (4.99V ,5.60V ,5.94V ,5.57 V ) = 22.10
Top Position Choice(61) : 26.23
```

Figure 12 Results of vertical servo motor movements

Figure 11 and Figure 12 is the result of horizontal servo motor testing and vertical servo motor testing which will move each servo motor to the position that has the highest voltage.

After determining the initial position of the system testing the servo motor movement by sun tracker or following the movement of the direction of sunlight. The results of the motion test follow the movement of sunlight which will later divide the voltage to drive the vertical servo motor and horizontal servo motor in an effort to optimize the intensity of sunlight received by the solar cell. The following results from the movement of the sun tracker system can be seen in the Arduino serial monitor application such as Figure 13.

```
Value Top: 2.18V
Value Down: 2.11V
Value Left: 2.46V
Value Right: 1.83V
Move to Position Vertical: 114 Degree, Position Horizontal: 100 Degree
```

Figure 13 Test Results of the Sun tracker system

#### 4.2.3 Testing Wifi

Wifi testing is used to send the voltage data received by the voltage sensor and send the servo motor position data to the website for monitoring. This Wi-Fi testing uses the Wi-Fi module ESP8266 to connect to the internet and send data to the server. Voltage sensor that is at tolerance (0.2V) which means that if all values of the voltage divider with a tolerance difference (0.2V), the ESP8266 module will send the voltage value and position of the servo motor to the website. The results of the wifi module testing can be seen on the website <https://pandan-sari.com/solar/> which will later display the position of the servo motor (Figure 14) and will display a graph with the amount of voltage per hour (Figure 15).

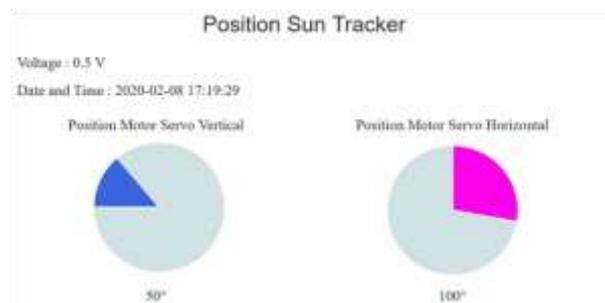


Figure 14 Sun tracker position

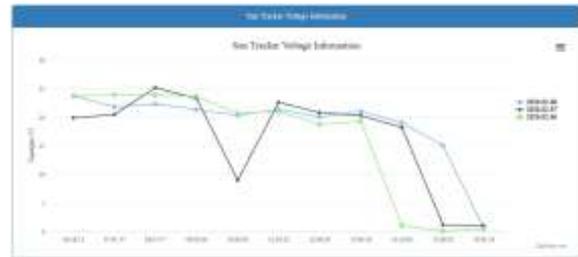


Figure 15 Voltage chart of the sun tracker system

## 5. CONCLUSION

The results of these experiments are able to move the sun tracker system by tracking sunlight with Dual-Axis or two-way. This system can also be monitored by using a website to display voltage data received by a voltage sensor and can display the location of the angular position carried out by a servo motor.

## 6. REFERENCES

- [1] J. Bambang Hari Purwoto, Muhamad Alimul F, Ilham Fahmi Huda, "EFISIENSI PENGGUNAAN PANEL SURYA SEBAGAI SUMBER ENERGI ALTERNATIF," Emitter, vol. 18 No. 01.
- [2] G. W. Wendryanto, W Sutaya, "PENGEMBANGAN PENGERAK SOLAR PANEL DUA SUMBU UNTUK MENINGKATKAN DAYA PADA SOLAR PANEL TIPE POLIKRISTAL," Jurnal Jurusan Pendidikan Teknik Mesin (JJPTM), vol. Volume : 8 No : 2, 2017.
- [3] N. K. Rizki Waloyo Tricahyono, "SISTEM MONITORING INTENSITAS CAHAYA DAN DAYA PADA DUAL AXIS SOLAR TRACKING SYSTEM BERBASIS IOT," Jurusan Teknik Elektro, vol. Volume 7 Nomor 4, 2018.
- [4] I. N. P. I Wayan Pande Agustiana Putra, Kadek Suar Wibawa, "Sistem Pendeteksi Kebakaran Menggunakan Raspberry Pi Berbasis Android," MERPATI, vol. 6, Desember 2018 2018.
- [5] I. K. G. D. P. Putri Isma Oktawiani, Kadek Suar Wibawa, "Sistem Penjemur Pakaian Otomatis Menggunakan Raspberry Pi Berbasis Android," MERPATI, vol. 6, 3 DESEMBER 2018 2018.
- [6] D. Prihatmoko, "PENERAPAN INTERNET OF THINGS ( IoT ) DALAM PEMBELAJARAN DI UNISNU JEPARA," Jurnal SIMETRIS., vol. 7, pp. 567-574, 2016.
- [7] F. Djuandi, in PENGENALAN ARDUINO, 2011.
- [8] B. T. Ima Maysha, Hasbullah, "PEMANFAATAN TENAGA SURYA MENGGUNAKAN RANCANGAN PANEL SURYA BERBASIS TRANSISTOR 2N3055 DAN THERMOELECTRIC COOLER," ELECTRANS., vol. VOL.12., pp. 89 - 96, SEPTEMBER 2013 2013.
- [9] S. P. Elmeki Satria, S. P. Elmeki Satria, Ed. MODUL PEMBELAJARAN MOTOR SERVO. Komplek Kementerian Pendidikan dan Kebudayaan, Gedung E, Lantai 13 Jalan Jenderal Sudirman, Senayan, Jakarta 10270: Direktorat Pembinaan Sekolah Menengah Kejuruan, 2017.
- [10] E. I. Team, ESP8266EX Datasheet. www.espressif.com, 2019..

# Whitening Processing of Seismic Signal Spectrum Based on SSGST

Shulin Zheng

School of Communication Engineering  
Chengdu University of Information Technology  
Sichuan, China

**Abstract:** In order to avoid the problem that single-domain spectral whitening cannot improve the time-frequency accuracy of seismic signals at the same time, this paper proposes a time-frequency domain spectral whitening based on synchrosqueezing generalized S transform(SSGST). First, the SSGST is used to perform time-frequency analysis on the seismic signal to obtain the time-frequency spectrum of the seismic signal. Then the whitening filter is used to perform spectral whitening on the time-frequency spectrum to obtain a seismic signal with higher time-frequency resolution. The verification of synthetic seismic signals and actual seismic signals proves that spectral whitening based on SSGST can simultaneously improve the time-frequency resolution of seismic signals, make the stratum information clearer, and have great significance for high-resolution processing of seismic signals.

**Keywords:** synchrosqueezing generalized S transform; spectral whitening; high resolution; seismic signal; data processing

## 1. INTRODUCTION

In seismic exploration technology, data interpretation occupies a very important position. Before interpreting seismic data, we must perform appropriate high-resolution processing on seismic data to further improve the accuracy of formation information estimation. Spectral whitening is one of the high-resolution processing methods. Its basic idea is to perform regional energy equalization of signals in the time domain or frequency domain to achieve the purpose of frequency compensation. In 1986, Bian Guozhu proposed spectral whitening of seismic data, and briefly described the principle of spectral whitening and its advantages in seismic signal processing<sup>[1]</sup>. In 2000, Chen Chuanren and others proposed to combine spectral whitening processing and wavelet transform. First, wavelet decomposition was used to decompose seismic signals to obtain signals of different scales, and then the signals were frequency compensated. Finally, the compensated signals were subjected to wavelet inverse transform. To obtain the high-resolution signal after spectral whitening<sup>[2]</sup>. In 2016, Lu Yiqi and others proposed to combine HHT with spectral whitening. First, use Hilbert-Huang transform to perform time-frequency analysis on seismic signals, and then whiten filter on time-frequency spectrum to obtain higher resolution seismic signals<sup>[3]</sup>. In 2018, Yan Zhonghui and others proposed to apply the Hilbert spectral whitening method to marine seismic data processing, which proved that the method is superior to conventional methods in local time-frequency characterization of signals and relative amplitude fidelity. Based on reading related literatures, this paper proposes a point spectrum whitening based on synchrosqueezing generalized S transform (SSGST). Compared with Fourier transform, synchrosqueezing generalized S transform can process non-stationary signals; compared with wavelet transform, synchrosqueezing generalized S transform can obtain time frequency spectrum instead of time scale spectrum, which can more intuitively analyze seismic signals. Time domain and frequency domain. First, use SSGST to perform time-frequency analysis on the seismic

signal to obtain the time-frequency spectrum of the seismic signal. Each point on the time-frequency spectrum has time and frequency characteristics. Then, using the whitening filter to whiten the time-frequency spectrum, we can obtain the high whitened spectrum. Resolution record.

## 2. PRINCIPLE

### 2.1 SSGST

The synchrosqueezing generalized S transform (SSGST)<sup>[4]</sup> is derived from the generalized S transform<sup>[5]</sup>. It combines the advantages of the synchronous extrusion algorithm and the traditional GST algorithm. It can flexibly adjust the window function and adapt to various frequencies. Seismic signals can also make the time-frequency high-focus distribution and obtain a higher-resolution time-frequency spectrum map.

Assuming the original signal is  $h(t)$ , its generalized S transform is expressed as:

$$\int_{-\infty}^{+\infty} h(t) \frac{|f|^{rgs}}{\sqrt{2\pi\rho}} \exp\left[-\frac{(\tau-t)^2 |f|^{2rgs}}{2\rho^2}\right] \exp(-i2\pi ft) dt$$

Where  $\tau$  is the time change,  $f$  is the frequency,  $rgs$  and  $\rho$  are the adjustment parameters of the window function. When  $rgs$  and  $\rho$  are different, different window functions can be obtained, so that the generalized S transform is transformed into Gabor transform or standard S transform.

According to the generalized S transform expression, we can derive the instantaneous frequency expression of the signal:

$$f_h(\tau, f) = \sum_{n=1}^N \left\{ \delta(f - f_n) \left[ f_n + (i2\pi GST_n(\tau, f_n))^{-1} \frac{\partial GST_n(\tau, f_n)}{\partial \tau} \right] \right\}$$

Where  $\delta(f)$  is the impulse function,  $f_h$  is the instantaneous frequency, and  $f_n$  is the frequency.

According to the synchrosqueezing transform algorithm, we can get the spectrum of the original signal's synchrosqueezing generalized S transform:

$$SSGST_x(f_i, \tau) = L_f^{-1} \sum_{f_k: |f_k - f_i| \leq \Delta f / 2} GST(\tau, f) \exp(i2\pi f_k \tau) f_k^{-1} \Delta f_k$$

Where  $\tau$  is time variable,  $f_i$  is the frequency of results,  $f$  is frequency.  $f_k$  is half the length of the frequency range.  $f_k$  is the discrete frequency points.

## 2.2 Spectral Whitening

Different from the conventional time-domain spectral whitening or frequency-domain spectral whitening, in the time-frequency domain spectral whitening algorithm, we first need to perform time-frequency analysis on the original signal, and we design a whitening filter based on the obtained time-frequency spectrum.

$$H(f, T_i) = \frac{\max SSGST(f, T_i)}{SSGST(f, T_i) + \varepsilon \max SSGST(f, T_i)}$$

Where the numerator is the maximum value of the amplitude spectrum at all times and is the white noise coefficient,  $\varepsilon$  is used to adjust the signal-to-noise ratio and resolution of the signal after spectrum whitening. When it increases, the signal-to-noise ratio of the signal becomes larger, and the resolution decreases. When it decreases, the signal-to-noise ratio of the signal becomes smaller, and the resolution increases. Then, the generalized S inverse transform of the spectrum-whitened signal is simultaneously squeezed to obtain a high-resolution seismic record.

## 3. APPLICATION

We validate the above method using actual seismic data. Intercept the 1250th signal of 5000 actual seismic signals, and collect 2,000 sampling points per channel with a sampling frequency of 1ms.

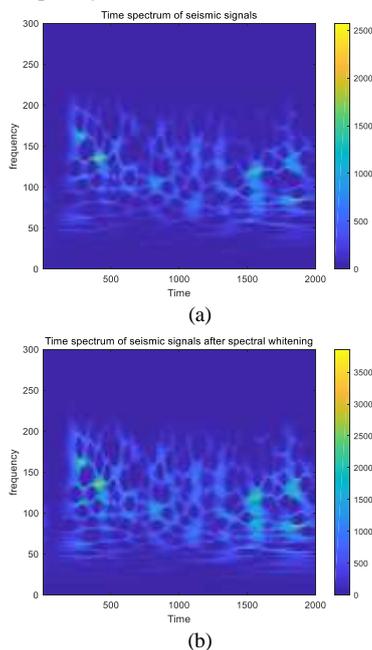


Figure1.Example of SSGST

Figure 1 (a) is the time-frequency diagram of the original single-channel seismic signal after SSGST processing. Figure. 1 (b) is a time-frequency spectrum diagram after spectrum whitening processing. By comparing (a) and (b), we can see that after spectral whitening, the amplitude of the signal divided in the high frequency range becomes larger, and the resolution is clearer. In the time-spectrum diagram, the detailed information of the signal after spectral whitening is

supplemented, the width of the effective frequency band is increased, and the information of the high-frequency part is compensated, which can more effectively reflect the detailed information of the seismic signal.

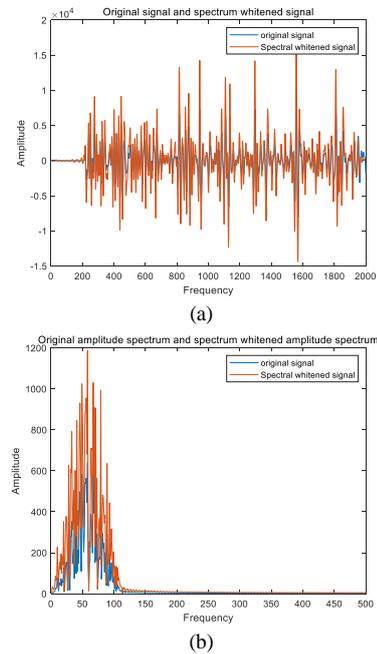


Figure2.example of Single-channel seismic signal

Figure 2 (a) is the comparison between the original signal and the spectrum-whitened signal. By performing ISSGST on the spectrum-whitened time-spectrum, we can reconstruct the original signal after spectrum-whitening. Figure 2 (b) is a comparison chart of the amplitude spectrum of the signal. It can be seen from the figure that the main frequency part of the seismic signal is compensated more and more detailed information is obtained.

Next, we perform spectral whitening on multiple seismic signals:

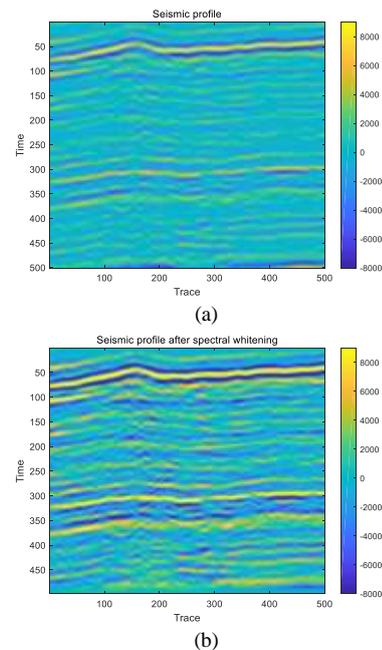


Figure3.Example of Multi-channel seismic signal

Figure 3 (a) is a cross-sectional view of the seismic signal. In the figure, 500 actual seismic signals are intercepted, and each

channel collects 500 sampling points with a sampling frequency of 1 ms. Figure. 3 (b) is a seismic cross-sectional view after spectral whitening. Comparing the two pictures (a) and (b), it can be seen that the seismic profile after spectral whitening processing has more obvious layering information. The low-amplitude part of the high-frequency part in the original image is compensated and the resolution is improved. Stratigraphic conditions and their distribution are clearer.

#### 4. CONCLUSION

The SSGST-based spectral whitening method proposed in this paper can effectively compensate seismic signals in the time and frequency domains. Time-frequency analysis is performed on the original signal, and then point spectrum whitening is performed on the time spectrum. Finally, ISSGST is performed on the spectrum-whitened signal to obtain a high-resolution seismic signal. This method overcomes the limitation that traditional spectral whitening methods can only improve time resolution or frequency resolution, compensates signals in the time-frequency domain, and has the characteristics that the SSGST algorithm can flexibly adjust the time window to adapt to different frequency signals. It is practical in high-resolution processing methods.

#### 5. REFERENCES

- [1] Guozhu, Bian. And Liqun, Zhang, J. 1986. Spectral Whitening of Seismic Data. Geophysical Prospecting for Petroleum.
- [2] Chuanren, Chen.and Xirang, Zhou. J. 2000. Wavelet Spectrum Whitening Method Improves Resolution of Seismic Data. Petroleum Geophysical Exploration.
- [3] Linzhi, Wu. And Hui, Chen. C. 2016. High Resolution Processing of Spectral Whitening Seismic Data Based on HHT Method. The 15th National Symposium on Mathematical Geology and Geoscience Information.
- [4] Chen, H. And Lu, L. J. 2017. The Synchrosqueezing Algorithm Based on Generalized S-transform for High-Precision Time-Frequency Analysis.
- [5] Handing, H. And Na F. J. 2014. Research on Generalized S-Transform Seismic Resolution Method. Petroleum Geophysical Exploration

# The Challenges of Contactless Payment Models using Near Field Communication Enabled Devices

<sup>1</sup>Charles O. Okunbor

<sup>2</sup>Yinka A. Adekunle

<sup>3</sup>Adewale O. Adebayo

<sup>4</sup>Alao, O.D.

<sup>5</sup>Eze, M.O.

<sup>12345</sup>Babcock University,  
Department of Computer Science, School of Computing and Engineering Sciences  
Ilishan-Remo, Ogun State,  
Nigeria.

---

**Abstract** – The goal of financial institutions is to improve customers experience by making transactions safer, faster, easier, and more convenient. The use of electronics and computing has been a reliable and fast way of achieving these goals. One of such improvement is contactless payment. Contactless payments is the use of mobile phones, electronic cards, and other devices with Near Field Communication(NFC) technology to conduct transactions that do not require a physical connection between the Point of Sales (POS) terminal and the device of the consumer. NFC is the wireless technology that is used to transfer card payment data from the device to the payment terminal via radio waves. Electronic cards has been the most popular means of payment around the world but financial institution want these cards to be emulated by NFC enabled mobile devices which is perceived to be more fashionable, convenient and more secure due to their computing capabilities. Despite the auspicious features of contactless payment using NFC enabled device, its adoption has been slow due to some challenges associated with their models. Host Card Emulation(HCE) and Secure Element(SE) models for contactless payment on NFC enabled devices were reviewed in this paper and its challenges were highlighted.

**Keywords:** Contactless payment, Host card emulation, Near field communication, Secure element.

---

## 1.0 Introduction

Contactless payment is the process of making secure payments using a short-range wireless technology between a contactless card or payment enabled device and a contactless enabled checkout terminal [1]. Payment information is sent or transmitted for authorization when a contactless card or an enabled contactless device is tapped or placed close to a contactless point of sale(POS) terminal. Transactions involving contactless payments are about twice as fast as transactions involving cash, debit, or credit cards. This technology give retailers a way to encourage more frequent visits by customers who prefer this payment method, as well as to potentially speed up transactions and reduce operational costs. In general, some transactions using contactless payment, especially those that involve little amount require no PIN verification or signature and hence makes it beneficial for users that value speed [2], [3]. Security has been a concern for contactless payments and because of that, financial institutions and countries have policies placing limits on this form of payment to reduce risky transactions [4], [3]. Due to this reason, it has been

challenging for people to adopt this technology and its growth has been slow. With the recent trends and advancement in computing, the world is moving towards a cashless economy and contactless payment will be a big contributor to achieving this [4]. According to a research carried out by [1], it is predicted that by 2025, 65% of all transactions will be by smartphones and 50.2% of the transactions will be contactless. Contactless payments are used on a range of devices including pre-paid, debit and credit cards; wearable devices, such as watches and wristbands; and mobile devices, such as smartphones and tablets [5]. Contactless devices used for contactless transaction uses a short-range radio frequency identification chip (RFID) known as Near Field Communication (NFC) technology. This technology is used to transfer payment information and other forms of data via radio waves when a user places a smart device or card within 4 inches or 10 centimeters of a reader or just waves it near a reader [6]. This process is also referred to as “tap and go” or “wave and pay” transactions.

## 2. REVIEW OF LITERATURE

### 2.1 Contactless Payment

Contactless payments refer to mobile phones, credit cards, and other devices that do not require a physical connection between the Point of Sales (POS) terminal and the payment device of the consumer. Transactions involving contactless payments are about twice as fast as transactions involving cash, debit, or credit cards [4]. This technology give retailers a way to encourage more frequent visits by customers who prefer this payment method, as well as to potentially speed up transactions because they do not require PIN verification or signature at the POS to authorize a transaction [2], [3]. Despite the convenience presented by this form of payment, users have questioned the security implications of using this technology [2].

### 2.2 Near Field Communication

NFC enables the mobile phone to act as a means of identification and a credit card for customers. NFC is a communication protocol that enables contactless transaction by establishing a short range wireless communication between two technical devices using frequency of 13.56 MHz, for instance between a mobile phone and a point of sales (POS) terminal [7]. NFC tags communication and data exchanges are based on standards like ISO 14443 A, MIFARE and FeliCa. It provides high comfort level and ease of use as there are no further configuration steps required to initiate a session to share data [8]. NFC is similar to WiFi, Bluetooth and other forms of wireless signals because they work on the principal of transmitting information using radio waves but NFC uses a different standard for wireless data communication which means that devices adhere to some specifications in order for them to properly communicate with each other [9].

NFC as a subset of RFID, was developed to provide a more secure, short-distance, and implicit paired communication capability. A good important aspect of NFC technology is its inherent security due to its very short communication range which makes it suitable for contactless payment. In NFC communication, bringing two devices very close to each other starts communication and separating the devices beyond a certain limit terminates the communication immediately [10].

Mobile payment is the driving force behind NFC technology over the past years, it is mostly used in contactless mobile payment. VISA estimates that mobile payment via NFC will replace the bank card in the coming years and most manufacturers of smartphones like Samsung, Apple have equipped their devices with this technology [11]. To enable mobile contactless payments, the NFC-enabled mobile device operates in card emulation mode and appears to an external reader to be a traditional contactless smart card [12].

### 2.3 Europay, Mastercard, and Visa (EMV)

EMV stands for “Europay, Mastercard and Visa”. EMV is an open-standard set of specifications for smart card payments and acceptance devices. Globally, financial institutions have migrated from magnetic stripe bank cards and infrastructure to EMV chip cards and infrastructure. [13] data shows that majority of POS terminals have been converted to EMV-enabled because of the shift of blames that occurs when fraudulent transactions take place. NFC mobile contactless payment transactions between a mobile phone and a POS terminal use the standard ISO/IEC 14443 communication protocol which is currently used by contactless EMV credit and debit cards [14].

### 3.0 CONTACTLESS PAYMENT MODELS

Two major architectures are used for mobile phones to store and communicate sensitive information such as card number, primary account number and other payment information. They are either by a hardware with Secure Element (SE) or a software with Host Card Emulation(HCE) [15], [16]. When card emulation is performed using an NFC mobile phone with a secure element, the interface to the payment reader (e.g. a point-of-sale or POS) is the same as for a traditional payment Credit/Debit card. This is similar with NFC mobile gadgets using the HCE, a POS or reader sees an application hosted in the mobile phones operating system as a standard EMV card (Andersson, 2016).

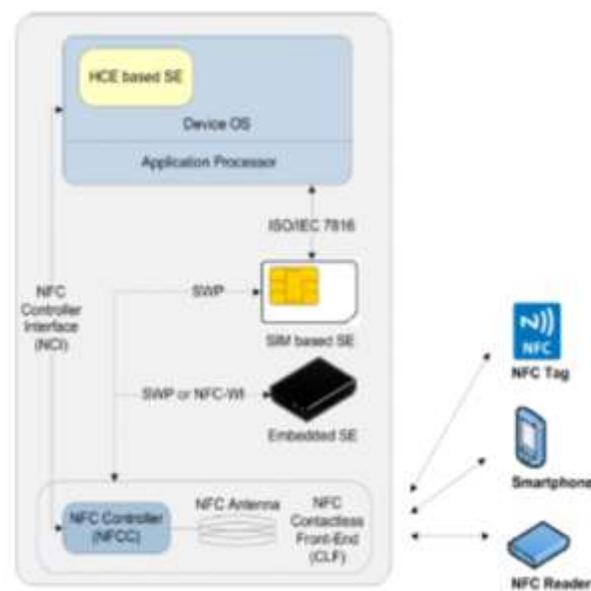


Figure 2.1. General architecture of an NFC smartphone. Image source: Coskun, Ozdenizci & Ok (2015).

### 3.1 Secure Element (SE) Model

When mobile devices with NFC are used to emulate smart cards, credentials like secret cryptographic keys used by payment applications are stored in a tamper resistant hardware module known as the Secure Element (SE) in accordance with the security requirements set forth by a known and trusted authorities [18]. The SE which is a tamper resistant hardware used to store sensitive credentials, has a direct connection with the NFC controller/antenna [16].

To make simple the idea of SE, [15] describes SE as a smart card in mobile devices. SE is known to have the highest level of security for applications residing on it. The level of security provided by SE is the same as the security level of classic smart cards [19]. One of the key advantages of SE is that it is a standalone component that creates a tough security against malicious sophisticated attacks. For SEs to offer a good level of interoperability and unparalleled rich portfolio of vital services, they are supported by mature ETSI, 3GPP, GlobalPlatform and Java Card standards.



Figure 2.2. A mobile phone with a Secure Element. (Image Source; Swaminathan, 2017)

The introduction of SE led to the development of new business models and partnerships regarding to the ownership and management issues of SE. NFC ecosystem actors such as mobile network operators (MNOs), mobile handset manufacturers, financial institutions like banks and transport institutions have tried to impose an alternative to SE using a specific business model from which they could benefit most [20]. The SE was modeled in three forms that could benefit these actors. According to [18], the secure element can reside in an embedded secure smart card chip on the handset, on the Subscriber Identity Module (SIM) or Universal Integrated Circuit Card (UICC), or on a secure digital (SD) card that can be inserted into the mobile phone.

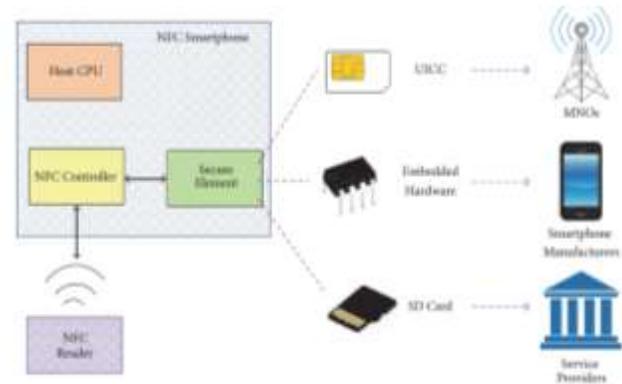


Figure 2.3. Diagram of different model of SE for card emulation. Image source; Ozdenizci, Ok., and Coskun (2016).

### 3.2 Host Card Emulation (HCE) Model

Host Card Emulation (HCE) was proposed as a short-cut for mobile NFC payments as it could allow financial institutions to launch mobile NFC products without the need of MNOs UICC/SIM card, mobile manufacturers embedded SE chip or other forms of Secure Element (SE) that will need a business agreement between NFC actors. HCE model allows the mobile device operating system (OS) to communicate directly over the NFC interface in card emulation mode. This would allow financial institutions like banks to offer mobile NFC services to customers over the top (OTP), bypassing the need to cooperate with mobile operators, phone manufacturers and other actors in the ecosystem, with the aim of reducing cost and complexity [19]. It simplifies the ecosystem by providing OTT technology for applications used for NFC contactless transactions at the expense of increasing payment transaction risk management.

#### 3.2.1 Tokenization

Storing payment credentials and cryptographic keys in the mobile device OS instead of the SE is considered less secure as discussed in HCE model, which is why additional security measures like tokenization is needed for HCE (Pandy & Crowe, 2016). Wadii, Boutahar and Ghazi (2017) defined tokenization as a process by which the primary account number (PAN) is replaced by a substitution value referred to as a Token.

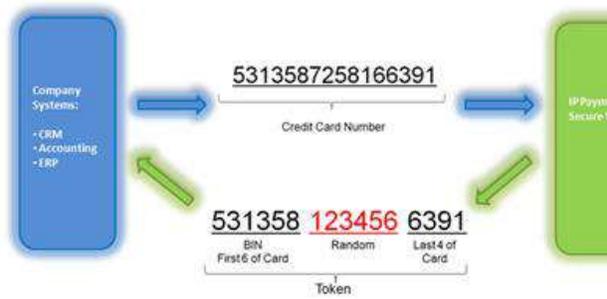


Figure 2.4. A token been generated in a banking transaction. Image Source; *Wadii, Boutahar and Ghazi (2017)*

## 4.0 IDENTIFIED CHALLENGES THE MODELS

In payment system, trust is a primary factor that cannot be overlooked because an accepted payment system must be perceived as being secure so that in the event of fraudulent transaction, a user is certain to be protected and refunded if need be (Smart Payment Association, 2015). With the introduction of NFC contactless payment, users and Issuers have questioned the safety of transactions and the storage of payment information. The following are the challenges faced by contactless payments models using NFC enabled devices.

### 4.1 Challenges of SE Model

SE is seen to satisfy the fundamental security paradigm, that a sensitive application must only be run in a secure computing platform, certified as tamper resistant [21]. Despite SE certified as a temper resistant, it has some challenges that has contributed to the slow growth of this technology as highlighted below;

- As pointed out earlier, SE was model into forms that could benefit various actors within the ecosystem. These alternative models brought disagreement among actors because each party wanted to hold a favourable business positions were they could benefit the most [19]. Unfortunately, a satisfactory agreement could not be reached by these actors. These disagreements was one of the shortfalls of SE which contributed to limiting the development of NFC contactless payment systems and other services in card emulation mode [20].
- UICC-based SE model created an advantage and opportunity for MNOs because they are issued and managed by them. A major challenge faced by financial institution for this model is that before a UICC/SIM can be used to host any financial application, a business agreement must be made with MNO before access and host space can be granted [19]. For customers, only UICC of MNO's

who have an agreement with their financial institution can be used. This means that users will need to go through the inconvenience of swapping UICC's or exiting a MNO in order to have access to contactless payment using the UICC of another MNO who has an agreement with their financial institution.

- SE has a problem of limited storage. It cannot accommodate so many applications for different Issuers and payment networks due to inadequate storage capacity on MNO's UICC's and smart phone embedded chips used for SE.
- Beyond the complexity and cost of establishing a relationship with a third party, only the SE provider determines who and what can access the SE. Some mobile carriers have a vested interest in limiting access to the secure element because they offer their own mobile wallets. For example, according to [22], Google wanted to install its application for contactless payment on UICC's. Major mobile phone operators such as Verizon, AT&T and T-Mobile declined their cooperation, instead promoted their own application which was initially called Isis Wallet but was later renamed Softcard.

### 4.2 Challenges of HCE Model

In other to find a more independent solution for SE led to the discovery of cloud based SE known as Host Card Emulation. HCE model have helped to remove the dependence of SEs owned and managed by third parties by allowing the mobile devices operating system (OS) to communicate directly over the NFC interface in card emulation mode. However, the relaxed security of HCE is still an important obstacle in its way [23]. Below are some of the challenges faced by HCE.

- Isolation and sandboxing provided by mobile OS is regularly broken, and consumers often root or jail break their device which unknowingly makes them risk sensitive data leakage. This makes access to users sensitive information for transaction such as payment credentials held by the HCE application hosted on the device's OS exposed and can be extracted and used by criminals for fraudulent transactions.
- The HCE runs on a non-secure platform, meaning that other applications resident in the mobile device, malicious or not, may compromise the integrity of payment applications. Malwares, spywares, viruses and other malicious programs can find themselves into users mobile OS and end up compromising the integrity of HCE applications. They can cause a Denial of Service (DoS) by maliciously modifying routing table from Android OS domain table or even saturating it by declaring a lot of AIDs (Application Identifiers). These malicious programs can steal,

- expose and transmit sensitive information from applications without the users knowledge.
- c) When a device having the HCE application is lost or stolen and falls into the hands of criminally minded persons, they could connect to all the information stored within the application and use them to make fraudulent payments.
  - d) HCE depends on a network connectivity to retrieve payment credentials from the cloud. This service becomes inaccessible if devices cannot connect to their service providers due to network failures.
  - e) Tokenization was introduced to minimize the risk of financial institutions and their customers using HCE by substituting payment credentials with temporary tokenized pseudo data. Tokenization comes with its challenges:
    - i. It increases the cost of processing transactions because a fee needs to be paid to tokenize and detokenize the card information of a customer for every transaction.
    - ii. The process of tokenization and detokenization before transactions can be processed and approved reduces transaction speed.
    - iii. Good tokens do not give room for data to be reconstituted, hence data analysis cannot be performed on tokenized transactions on HCE.

## 5.0 CONCLUSION

This paper has been able to highlight the challenges facing contactless payment using NFC enabled devices. Most of the challenges faced with this form of payment are security issues relating to secure storage of payment credentials and lack of cooperation between actors within the contactless payment ecosystem. In order to mitigate the identified challenges faced with contactless payment models, cooperation is needed between financial institutions, MNO's, device manufacturers and other actors within the ecosystem. Cooperation is needed because aside the renting of SE, other services such as network connectivity, NFC enabled devices, supporting operating system and so many others provided by different actors is also needed to make the payment system successful. HCE should be seen as a viable alternative to SE and minimum security requirements should be set for the implementation of HCE through standardization.

## REFERENCES

- [1] VISA (2018). Contactless Payments. Retrieved on 24th October, 2018 from; <https://usa.visa.com/pay-with-visa/contactless-payments/contactless-payments.html>
- [2] Fuller, A. (2017). Contactless Payments – The pros and cons of tap and go. Retrieved on 24th October, 2018 from; <https://www.finder.com/contactless-payment>

- [3] Heropay, (2018). Contactless Card Payment. Retrieved from: <https://www.heropay.com/glossary/contactless-card-payment/>

- [4] Banka, L., & Ambre S. (2016). Contactless Payments. Retrieved from: <https://www.arx.cfa/up/post/2461/Contactless%20payments.pdf>

- [5] Phair, N. (2016). The Truth about Contactless Payments. Center for Internet Security. Retrieved from; <http://www.canberra.edu.au/cis/storage/Contactless%20payments.pdf>

- [6] Quibria, N. (2008). The Contactless Wave: A Case Study in Transit Payments. Federal Reserve Bank of Boston: Emerging Payments Industry Briefing.

- [7] Chae, J. S. U., & Hedman, J. (2015). Business Models for NFC based mobile payments. *Journal of Business Models*, 3(1).

- [8] Trivedi, D. (2015). Near field communication. Masters degree thesis, Nirma University, Ahmedabad-382481.

- [9] Triggs, R., (2017). NFC Tags - How Do They Work. Retrieved from; <https://www.androidauthority.com/nfc-tags-explained-271872/>

- [10] Coskun V., Ozdenizci B., & Ok K. (2015). The Survey on Near Field Communication. *Sensors* 2015. Basel, Switzerland. 1424-8220

- [11] Wadii, H.E., Boutahar, J., & Ghazi, E.S. (2017). NFC Technology for Contactless Payment Ecosystems. *International Journal of Advanced Computer Science and Applications*,

- [12] Noh, S. K., Lee, S. R., & Choi, D. (2014). Proposed m-payment system using near-field communication and based on WSN-enabled location-based services for m-commerce. *International Journal of Distributed Sensor Networks*, 10(4), 865172.

- [13] NCR (2015). The Road to Contactless Payments, EMV, Apple Play and Tokenization Patents Pending 15FIN3279A-0615.

- [14] Smart Card Alliance (2011). The Mobile Payments and NFC Landscape: A U.S. Perspective. A Smart Card Alliance Payments Council White Paper. Publication Number: PC-11002.

- [15] Lepojevic, B. (2012). Management of secure element in systems based on NFC technology. (M. Sc. thesis), Faculty of organizational sciences, University of Belgrade, Serbia.

- [16] Pannifer, S., Clark, D., & Birch, D. (2014). HCE and SIM Secure Element: It's not black and white. A Discussion Paper from Consult Hyperion.

- [17] Andersson, D. (2016). A survey on contactless payment methods for smartphones. Faculty of Computing, Blekinge Institute of Technology, Sweden. Thesis no: BCS-2016-05

[18] Smart Card Alliance (2014). Host Card Emulation. A Smart Card Alliance Mobile & NFC Council White Paper. Retrieved from; <http://www.smartcardalliance.org>.

[19] Lepojevic, B., Pavlovic, B., & Radulovic, A. (2014). Implementing NFC service security – SE VS TEE VS HCE. Ministry of Defence Republic of Serbia. Retrieved from; [https://www.researchgate.net/publication/263506976\\_Implementing\\_NFC\\_service\\_security\\_-\\_SE\\_VS\\_TEE\\_VS\\_HCE](https://www.researchgate.net/publication/263506976_Implementing_NFC_service_security_-_SE_VS_TEE_VS_HCE)

[20] Ozdenizci, B., Ok, K., & Coskun, V. (2016). A Tokenization-Based Communication Architecture for HCE-Enabled NFC Services. Hindawi Publishing Corporation Mobile Information Systems. 2016 (5046284).

[21] Smart Payment Association (2015). Is host card emulation (hce) the big enabler for mobile contactless payments? *An SPA Position paper*.

[22] Grustniy, L. (2019). Secure Element — securing contactless payments in smartphones. Kaspersky daily. Retrieved from; <https://www.kaspersky.co.uk/blog/secureelement/13687/>

[23] Alattar, M., & Achemlal, M. (2014). Host-based card emulation: development, security and ecosystem impact analysis. Proceedings of the IEEE International Conference on High Performance Computing and Communications

# SQL Injection Attacks Predictive Analytics Using Supervised Machine Learning Techniques

<sup>1</sup>Akinsola, Jide E. T.  
Department of Computer  
Science, Babcock University,  
Ilisan-Remo, Ogun State,  
Nigeria

<sup>2</sup>Awodele, Oludele  
Department of Computer  
Science, Babcock University,  
Ilisan-Remo, Ogun State,  
Nigeria

<sup>3</sup>Idowu, Sunday A.  
Department of Computer  
Science, Babcock University,  
Ilisan-Remo, Ogun State,  
Nigeria

<sup>4</sup>Kuyoro, Shade O.  
Department of Computer Science  
Babcock University,  
Ilisan-Remo, Ogun State

**Abstract:** Structured Query Language Injection Attack (SQLIA) is one of the most prevalent cyber attacks against web-based application vulnerabilities; that are manipulated through injection techniques to gain access to restricted data, bypass authentication mechanisms, and execute unauthorized data manipulation language. There are several solutions and approaches for identification and prevention of SQLIA, such as Cryptography, Extensible Markup Language (XML), Pattern Matching, Parsing and Machine Learning. Machine Learning (ML) approach has been found to be profound for SQLIA mitigation, which is implemented through defensive coding approach. Machine Learning Approach requires a lot of data for efficient model training with capability for using several attack patterns. ML approach can be used to mitigate a very hard blind SQL injection attack. An experimental analysis was performed in Waikato Environment for Knowledge Analysis on Logistic Regression (LRN), Stochastic Gradient Descent (SDG), Sequential Minimal Optimization (SMO), Bayes Network (BNK), Instance Based Learner (IBK), Multilayer Perceptron (MLP), Naive Bayes (NBS), and J48. Hold-Out (70%) and 10-fold Cross Validation evaluation techniques were used to evaluate the performance of the supervised learning classification algorithms to choose the best algorithm. The results of Cross Validation technique showed that SMO, IBK and J48 had Accuracy of 99.982%, 99.995% and 99.999% respectively; while Hold-Out technique showed that SMO, IBK and J48 had Accuracy of 99.986 %, 99.989 % and 100 respectively. On the other hand, in Cross Validation technique SMO, IBK and J48 had time to build model value of 10.15sec, 0.06sec, and 14.12sec respectively while in Hold-Out technique SMO, IBK and J48 had time to build model value of 9.71sec, 0.16sec and 14.28sec respectively. From the findings, IBK had the minimum time to build model in Cross Validation technique in addition to better performance in Accuracy, Sensitivity as well as Specificity and was chosen as the classifier for SQLIA detection and prevention. Therefore, beyond Accuracy, other performance evaluation metrics are critical for optimal algorithm selection for predictive analytics.

**Keywords:** Cyber security, Injection attack, Injection vulnerability, Predictive analytics, SQLIA, Machine learning, Web application

## 1. INTRODUCTION

Structured Query Language Injection Attack (SQLIA) is a forceful code manipulation insertion attack targeted against database through vulnerable web applications, which is a source of assault into the database. According to [1], SQLIA is a deliberate query manipulation insertion assaults lunched against backend database through compromised web applications, which exposes the database schema, circumvent confidentiality, integrity and availability of the sensitive information in the database. The prime purpose of which is to steal sensitive information.

Structured Query Language (SQL) injection vulnerability is the one of the most common web-based application vulnerabilities that can be exploited to gain access to restricted data, bypass authentication mechanism, and execute unauthorized data manipulation language [2]. Web application vulnerabilities are cyber security problems. Practically, there have been serious apprehensions on Cyber Security issues in the entire industries. These issues affect organizations and transverse practically all industries, but not limited to

distribution, monetary, investment, transmission, transportation and communications. Specific of the most efficient defenses against cyber infringements and disruption are big data and analytics [3]. Web applications have SQL injection vulnerabilities because there is no sanitization of the inputs used in constructing structured output [4]. Vulnerability is an application implementation vulnerability or defect that enables an intruder to trigger unwanted activities or obtain unlawful access. The presence of vulnerability introduce a threat to the application as it may result in a compromise of stored information [5].

SQLIA is a cyber attack aimed at database, which uses manipulation of query language called SQL. The SQL is a language intended to support database information which are stored (being at rest) and on motion. However, it is susceptible to various assaults. Developing methods to detect and avert assaults on database is highly essential while allowing web applications to be as user-friendly as possible [6]. Injection vulnerabilities for example SQL injection and Cross Site Scripting (XSS), rank among the top two out of the top 10 from the analysis conducted by Open Web Application

Security Project (OWASP). SQLIA and XSS are the two most prominent attacks affecting applications running on the web with security vulnerabilities or assaults. Code injection attacks such as SQLIA account for 40.8% and Cross Site Scripting (XSS) attacks account for 11.3% of total attacks in 2018 according to OWASP [7]. Consequent upon this alarming outrage impact on online applications, vulnerability assessment of enterprise Internet-based applications are crucial. There has been upsurge in availability of information and device connectivity have brought about increase in application of machine learning (which is a sub-domain Artificial Intelligence (AI) in diverse areas and data extraction research activities into more prominence to tackle the menace of cyber attacks.

There are different types of approaches for initiating SQL injections such as Tautology, Inference, System Stored Procedure, Piggy-backed Query, Union Query, Logically Query, and Alternate Encodings. However, for effective mitigation of all SQLIA, these approaches can be further characterized into eighteen for effective SQLIA detection and prevention, which are Time-based error, Database Fingerprinting, are Stored procedure, Buffer Overflow, Second Order, Deep Blind, Out of band, Alternate Encoding, Conditional Error, Union, Double blind, Conditional response, Illegal / Invalid / Logical Incorrect, Piggy Back, Error based (blind), Database Mapping, Literal, and Tautology with the non-malicious class referred to as Benign.

There are several solutions and techniques for identification and avoidance of SQLIA, such as Cryptography, Extensible Markup Language (XML), Pattern Matching, Parsing and Machine Learning. Machine learning approach have been found to be profound for SQLIA mitigation. It can handle input type checking, pattern matching and encoding input categories of injection attacks to address login, URL and search vulnerabilities mechanisms. Machine Learning Approach for SQLIA detection and prevention utilizes a lot of data for training using several attack patterns. ML approach can be used to mitigate a very hard blind SQL injection attack. The performance of ML technique is dependent on dataset reliability to meet the intended purpose and elimination of bias in choosing the best classifier for testing and training the model.

Hence, this study implements predictive analytics to detect as well as prevent web application exposures with focus on various SQLIA types / classes. Therefore, mitigation of escalating security breaches using supervised machine learning are addressed from Static and dynamic analysis approaches.

## 2. LITETRATURE REVIEW

Structured Query Language Injection (SQLI) was first used openly in 2000 and came into existence way back to 1998 [8]. SQLIA has since become one of the most frequent Internet attacks [9]. It happens when the mischievous user changes the allowable or genuine query syntax with the introduction of new SQL keywords or operators that result in unforeseen outcomes not intended for web applications [10]. Code injection is used generally to mean injecting code attacks that are consequently performed by a vulnerable application [11]. SQL Injection may be used to cause serious problems in a variety of ways. An intruder can bypass authentication, gain entrance, modification, and deletion of information within a database by using SQL Injection. SQL Injecting can even be implemented in some instances to execute controls on operating system, which could enable an attacker to scale into

a network behind a firewall to further commit devastating attacks [12].

### 2.1 SQL Injection (SQLI) Types

SQL injections and XSS are the two main security risks with un-sanitized user input. SQLIA is divided into three main arrangements such as In-band SQLI (also referred to as Classic SQLI), Inferential SQLI (also called Blind SQLI) as well as Out-of-band SQLI.

#### 2.1.1 Classic In-band SQLI

The utmost prevalent and straightforward SQLIA is in-band SQLI. In-Band-based SQLI takes place after an intruder both initiate the assault and collect outcomes using the same interaction route. The two most prevalent kinds of SQL injection that are in- band based are SQLI based on errors and SQLI based on Union [12].

#### 2.1.2 Inferential Blind SQLI

No actual transmitted of data through the web application when a SQLI attack is inferential, and the attacker cannot view the consequence of an in-band form of attack. The intruder can redefine the database structure of the database by placing payloads in Inferential SQL injection to observe web application reaction and the resultant database server behaviour. Boolean-based (Content-based) Blind SQLI in addition to time-based blind SQLI are the two kinds of inferential SQL injection [12].

#### 2.1.3 SQLI Based On Out-of-Band

Out - of-band SQLI is not quite prevalent, mainly since it will depend on the functionality of the web application that is used on the database server. SQL Out - of-band Injection happens if an invader cannot start the attack by using the same route and gathering outcomes that will be inimical [12].

### 2.2 SQL Injection Attack (SQLIA) Types

[13] Opined that the effect of SQL injection assaults could range from delicate data collection to file manipulation, from system-level command execution to application Denial of Service (DoS). The effect relies also on the database of the destination computer and the SQL Statement's functions and preferences. If DOS attack is lunched, it could have devastating effect of the entire system. Thus, rendering the web application in-accessible. This could also be in a coordinated fashion referred to as Distributed Denial of Service (DDoS). SQLIA can generally be divided into three classifications:

#### 2.2.1 Attack of First Order

A malicious string can basically be entered and the modified code implemented immediately.

#### 2.2.2 Attack of Second Order

The invader inserts into persistent trusted source storage (for example a table row). The hackers then utilized another activity later to perform an attack.

#### 2.2.3 Injection Based on Literals

The attacker may change the To\_Char() implicit function through altering of the environment variables, NLS\_Date\_Format or NLS\_Numeric\_Characters values.

### 2.3 SQL Injection Attack Techniques

The SQLIA techniques focuses on the attack mechanism, which hackers can attempt to carry out the hacking. These seven majorly used SQLIA techniques are: Tautology, System Stored Procedure, Inference, Illegal / Logically Incorrect Query, Alternate Encodings, Union Query and Piggy-backed Query. These mechanisms require effective taxonomical

formulation in order to handle the various forms of SQLIA efficiently. Therefore, classification based on these common seven mechanisms only can create loopholes for the intruders to gain access to database schema, evading detection and hence circumvent confidentiality, integrity and availability of the sensitive information in the database. Thus, the characterization into eighteen classes for effective SQLIA detection and prevention, which are Time-based error, Database Fingerprinting, are Stored procedure, Buffer Overflow, Second Order, Deep Blind, Out of band, Alternate Encoding, Conditional Error, Union, Double blind, Conditional response, Illegal / Invalid / Logical Incorrect, Piggy Back, Error based (blind), Database Mapping, Literal, and Tautology as malicious attacks with the non-malicious class referred to as Benign:

### 2.3.1 Tautology

This is concerned with one or more conditional statements used to inject code so as to always validate the true statements. This method occurs when the input data to the database is not checked. An instance of such a vibrant SQL statement is the code given thus; query= "SELECT details FROM customer WHERE name=' name' AND pwd=' pwd;'" attackers may use tautologies to make use of this software balance by providing an entry parameter number(' OR '1'='1') with the significance. An intruder could enter customer data without a relevant consideration because the situation of the WHERE clause becomes the same (which makes the system validates the outcome to be true and terminates the remaining query using (--)). (WHERE='x' Or '1'='1'--;).

Example of Tautology query attack: **SELECT \* FROM employee WHERE name = ' ' or 1=1 -- ' AND password = '12345';**

### 2.3.2 Piggy-Backed

The hackers will insert additional queries to be performed by the database in this scenario to extract, input or alter information, service performance denial or carry out commands from distance [14]. Attackers do not attempt to change the initial request in that situation. They actually attempt to attach an additional and different entry to the initial request using personal SQL-based phrases such as OR, AND, INSERT, UPDATE, DROP or DELETE to permit various SQL queries to the database [15].

Example of Piggy-backed query attack: **SELECT \* FROM employee WHERE name = 'guest' and password = '1234'; DROP TABLE employee; -- ;**

### 2.3.3 Alternate Encoding

Hackers mainly aim to avoid identification when using this technique. In fact, this sort of attack is used to encode the attack strings to avoid the filtering from the programmer (e.g. by using hexadecimal, ASCII and Unicode character set). In reality, additional encodings are generally applied in relation to other attack methods and target dissimilar application levels [16]. The usage of quote (') in the SQL statement declaration that can be used in the creation of different form of malicious database query request is prohibited for most of SQL injection mechanism that uses filters.. In place of a single quote which can easily be detected as bad character, for instance, the intruder uses char (44). This attacks combines char () function and ASCII hexadecimal encryption. Real characters(s) are returned when char () function is used to convert to hexadecimal character(s) encoding equivalent.

Example of alternate encoding query attack: **SELECT accounts FROM login WHERE username="" AND password=0; exec (char (0x736875746466j776e))**

### 2.3.4 Illegal / Logical Incorrect Query

In that assault, attacker attempts injecting declarations which cause the application servers to return a syntax error page to identify injectable parameters, it applies finger-printing and extract data from the web application's back-end databases [17]. In reality, error page gives hackers information about few details of tables' name in the databases, such as instances, or discloses vulnerable / injectable parameters for an intruder and such details will be used in carrying out the next attack phase [18].

Example of Illegal / Logical Incorrect query attack: **SELECT \* FROM employee WHERE name = ' ' UNION SELECT SUM(username) from users -- ' and password= ' ' ;**

### 2.3.5 Union Query

In this attack technique, the malicious query is added to the initial request via the UNION keyword to obtain information concerning additional database tables. An intruder can pull out column data or type of data details from this sort of attack [19]. By rule, most of the SQL conforming databases, including SQL Server stores metadata with sysobject numbers, syscolumns, sysindexes, and so on, in a set of system tables. This allows a hacker to use the information about the database table to identify schema information for a database in order to help hackers to launch assaults to the database further.

Example of Union query attack: **SELECT emp\_id FROM employee WHERE name = ' ' UNION SELECT cardNo FROM creditCard WHERE accNo = 10032 -- AND password = ' ' ;**

### 2.3.6 Stored Procedures

This method uses vicious SQL codes to execute integrated built-in functions, which further escalate privilege, ensures service denial or to execute remote controls. Indeed, most database providers develop database solutions with standard stored procedures and features to enhance the database functionality and brings interactivity with the operating system. Therefore SQLIAs may be created to perform stored procedures on this particular database once an attacker has known the backend database [19] [20].

Example of stored procedure query attack: **CREATE PROCEDURE DBO @userName varchar2, @pass varchar2, AS EXEC ("SELECT \* FROM user WHERE id= ' "+@userName+" and password= ' "+@pass+""); GO**

### 2.3.7 Inference

An intruder draws logical conclusion from a response to a right / wrong enquiry about database server answer. Two Blind injection and time injection input methods are used to launch this attack [21]. In-Blind injection, hackers obtain database information by submitting a server's true / false questions and the answers from this page gives leading information that will be exploited further. If the response is accurate, the request is correct and if the response is wrong, then an error will be triggered. An intruder can therefore obtain implicit response from the database [22]. Part of Inference attack can be classified into Blind SQL injection and Timing Attack.

Example of inference (blind) SQL injection attack: **SELECT \* FROM emp\_name, emp\_address, gender, from employee where 1=0; drop employee**

## 2.4 Defensive Coding Approaches for SQLIA

Defensive Coding (DC) is one of the SQLIA detection and prevention approaches. It is employed to execute safe queries so that it is compatible with unforeseen inputs or user behaviour in a timely way regardless of the kind of inputs supplied or actions exhibited by the user. The concept being taken advantage of is that every module of the program is exclusively independent. DC approaches being used for web application vulnerabilities mitigation are Cryptography, Input Type Checking, Pattern Matching, Extensible Markup Language (XML), Encoding Input, Parameterized Query and Stored Procedure which can be implemented using Parse Tree Approach or Machine Learning Approach. Parse Tree Approach encompasses input type checking and pattern matching while machine learning approach encompasses input type checking, pattern matching and encoding input. Pattern matching is capable of mitigating SQLIA requests through login, URL and search. Figure 1 shows the defensive coding approaches for SQLIA.

### 2.4.1 Parsing Approach

Parsing Approach is also known as parse tree approach. This is a technique to detect and avoid a SQLIA on the application's URL, was suggested by [23]. In this technique, the SQL\_statement\_safe query model was developed as a library with a SQL statement syntax grammar. This grammar syntax was based on two viewpoints, one for a fixed query and the other for a stacked request. It also includes the SQL query tree structure. The query will first be tested on SQL\_statement\_safe when a user requests SQL query from a website URL to check if the query is single and is consistent with the semantics of a genuine SQL statement declaration.



Figure 1. Defensive Coding Approaches for SQLIA

### 2.4.2 Machine Learning Approach

In order to identify and prevent SQLIAs, [24] suggested an automation method using the Bayesian algorithm. The monitor intercepts the SQL query, breaks it into numerous keywords based on blank space in a dynamic query, and calculates the length of the SQL dynamic query in the URL from a website when the user sends a dynamic SQL query. Furthermore, amount of keywords is calculated and numerical values and dynamic query keywords are sent to the classifier.

The classification algorithm then calculates the probability of SQL injection in a dynamic query based on the result received

from a converter, and then compares the likelihood of SQLI with a user threshold defined as a data set that helps calculate the probability of legitimate query and the likelihood of malicious query. If dynamic SQL query likelihood is calculated by classification algorithm, the query is permitted if there is a match with the likelihood of legitimate query computed in training dataset; otherwise, the query is blocked. In this method, one essential approach or mechanism is to simulate numerous attack patterns in training data, along with a very difficult blind SQL injection attack.

## 3. MACHINE LEARNING

A specialized area of artificial intelligence (AI), referred to as Machine Learning (ML), focuses on allowing computing systems to learn from data how to automatically perform the desired task. Machine learning is a key technology in the use of data and large data mining technology in diverse fields of healthcare, science, engineering, business and finance, and includes decision making, forecasting, or prediction [25].

### 3.1 Types of Machine Learning

There are different types of machine learning such as Supervised Learning (SL) and Unsupervised Learning (UL), Semi-Supervised Learning (SSL) in addition to Reinforcement Learning (RL) and Evolutionary Learning (EL) and Deep Learning (DL) [26].

#### 3.1.1 Supervised Learning (SL) Technique

SL trains a system from known input and output data to predict future outputs. The predictive model is developed based on the data input and output. Classification and regression are examples of supervised learning from two different categories. It is used mostly for the prediction and classification of numerical values such as regression and predicting the corresponding class respectively [27].

#### 3.1.2 Unsupervised Learning (UL) Technique

UL technique aims at finding underlying data structures and hidden patterns in data. The datasets consisting of input data without labelled responses are used for drawing inferences. Clustering is a type as well as the utmost prevailing unsupervised method of learning. Unsupervised learning is implemented to locate unknown pattern in turn data grouping. It is mostly applied to market research, object detection, predicting heart attack (medical) and so on. Fuzzy C-Means and k-Medoids, Self-Organized Maps, Gaussian Mixing Models, Hidden Markov Models, Hierarchical Clustering, K-Means and Subtractive Clustering are all algorithms for performing clustering.

#### 3.1.3 Semi Supervised Learning (SSL) Technique

SSL lies between UL as well as SL techniques. It is part or class of the machine learning, including unlabelled training data (e.g. a tiny number of data labelled with a lot of unlabelled data) which includes methods and tasks of the learning. SSL may also be known to be either transductive learning or preparative form of learning [28].

#### 3.1.4 Reinforcement Learning (RL) Technique

RL is one of the Machine Learning techniques that deals on how software agents in an environment should take action to optimize a notion of aggregate reward. In order to maximize recompense in a specific situation, reinforcement is about taking appropriate steps. The best possible conduct or approach to a particular situation is sought through various machines and software. For instance RL is widely used in

Personal Computer games and Robotics for industrial automation.

### 3.1.5 Evolutionary Learning (EL) Technique

EL is an Evolutionary Computation sub-set, a specific Meta Heuristic Optimization Algorithm based on population. An Evolutionary Algorithm (EA) utilizes biological evolutionary processes, including reproduction, modification, recombination as well as selection.

### 3.1.6 Deep Learning (DL) Technique

DL is an aspect of ML relying on data depictions rather than algorithms for specific tasks. ML can be supervised form of learning, semi-supervised form of learning [29]. It is applicable in the areas of computer vision, Natural Language Processing (NLP), voice identification, sound identification, machine translation, bioinformatics, drug design, filtering of social networks, analysis of medical images, product inspections and board game programs where Deep Learning frameworks like Deep Neural Networks, Deep Beliefs Networks in addition to Recurrent Neural Networks were created to generate some results similar to and in some instances of higher quality in comparison to human specialists [30].

## 3.2 Machine Learning Algorithms

Machine learning classification algorithms for evaluation of performance metrics belong to the following four categories / classes of classifiers accordingly such as function (Logistic Regression), Bayes (Bayes Network (BNK) and Naive Bayes (NBS)), Sequential Minimal Optimization (SMO) and Multilayer Perceptron (MLP)), Tree (J48) as well as Lazy (Instance Based Learner (IBK)). It is to be noted that SMO is a variant of Support Vector Machine (SVM). Also MLP is a variant of Artificial Neural Networks (ANNs).

### 3.2.1 Bayes classifiers

BNK and NBS are examples of Bayes classifiers. Bayes classifiers are probabilistic classifiers relying on the fundamental probability law of Thomas Bayes known as the Bayes Theorem as depicted in equation 1.

$$P(B/A) = \frac{P(B/A) \times P(A)}{P(B)} \quad (1)$$

Equation 1 shows the connection between A and B conditional likelihoods and the likelihoods. A classifier called Naïve Bayes as a classifier is an uncomplicated algorithm having autonomous characteristics that implies that an algorithm believes that the characteristics are not mutually likely. Bayesian networks are comparatively advanced algorithms that evaluate the likelihood of ambiguity and thus allow more complicated information from the analyzed data.

### 3.2.2 Function Classifiers

The function classifiers are Logistic Regression (LRN), Sequential Minimal Optimization (SMO) and Multilayer Perceptron (MLP).

#### 3.2.2.1 Logistic Regression

This is a classification function using a building class and a single logistic regression multinomial model with a single estimator. Logistics generally specifies where the class border lies. The class probabilities are also determined in a specific approach depending on the distance from the boundary [31]. When dataset is bigger, it passes to ends which are (0 and 1). These probability statements do not just make logistic

regression a classifier, but an efficient classifier. It makes stronger, more detailed forecasts and can fit in another way; however, these strong predictions can go wrong sometimes.

#### 3.2.2.2 Sequential Minimal Optimization (SMO)

This is an SVM variant. Classical Mufti- Layer Perceptron Neural Networks are strongly linked to SVM algorithms. SVMs revolve on both sides of the hyperplanes around the concept of a gap that distinguishes two categories of data [32]. It has been shown that the maximization of the margin and hence the maximum possible distance between the separating hyperplanes and the instances on both sides reduces the upper limit of the expected generalization error [33]. The SVM classification accepts the data of several classes after this now creates vectors for the best hyperplanes to be separated into a feature space or parameter space. The hyperplane that is placed to the closest data points nodes at the highest range is described as ideal [34].

#### 3.2.2.3 Multi-Layer Perceptron (MLP)

The ANN version is MLP. MLP is a categorizing element which determines the weights of the network, not by creating a non-convex, uncompromising minimization issue as in conventional Neural Network training but by addressing a quadratic programming problem with linear limitations [35]. ANN is an algorithm of learning that solves problem of classification. An ANN model consists of several neuron network systems, which are parallel, dynamic and inter-linked. A neuron is used by a defined mathematical processor to generate outputs using inputs [36].

### 3.2.3 Tree Classifiers

J48 is the Iterative Dichotomiser 3 (ID3) expansion. J48 also contains features for missing values remedy, pruning of decision-trees, continuous value collections for attributes, rules derivation, etc. It's an algorithm for the decision tree. The algorithm named decision tree is used to determine the behaviour of the attributes / vector in several instances. The classes for the recently produced instances are also discovered on the basis of the teaching instances [37].

### 3.2.4 Lazy Classifiers

Learner based on instances for example IBK is classified as Lazy classifier. It is an algorithm of k-Nearest Neighbour (k-NN). The method is a straightforward and simple method of classifying a certain dataset with fixed apriori K-means algorithm clusters (suppose k clusters). When labelled data are not accessible, K-means algorithms are used [38]. It utilizes a particular way to transform rough thumb rules into extremely precise forecast rule. As a result of weak learning algorithms, classifications (thumb rules) can continuously be at least slightly reliable than random, with about 55 percent accuracy. But a boosting algorithm can likely build one classifier with increased accuracy, say 99 per cent [39].

## 4. METHODOLOGY

The experimental analysis of the machine learning algorithms was performed using Waikato Environment for Knowledge Analysis (WEKA). A model that can be used for better classification of SQLIA dataset into attack classes effectively was developed using the algorithm with optimal performance. Hold-out (70%) and 10-fold cross-validation evaluation techniques were used to evaluate the performance of the classification algorithms (supervised learning) to choose the best algorithms. This was carried out in relation to evaluating

performance metrics which comprises of criteria such as Kappa Statistic, True Positive (TP) Rate, Accuracy, True Negative (TN) and Training Time (time to build model (TTB)), for each of the machine learning algorithm.

#### 4.1 Performance Evaluation metrics

In evaluating the performance of the classification algorithms, the model was built in WEKA 3.8.0 using the hold-out (70% training data) and 10-fold cross-validation evaluation methods on Logistic Regression (LRN), SMO, Bayes Network (BNK), IBK, Multilayer Perceptron (MLP), Naive Bayes (NBS), and J48. After the training process, the values of benefit criteria such as correctly classified instances (accuracy), Kappa Statistic, True Positive (TP) Rate, True Negative (TN) Rate and Training Time (i.e. Time to Build) were compared.

### 5. RESULTS

The performance of the machine learning algorithms were measured based on ten (10) existing performance benchmarks: Accuracy, Kappa Statistic, True Positive (TP) Rate, True Negative (TN) Rate and Training Time (i.e. time to build). Tables 1 to 7 and Figures 2 to 6 depict the comparison of the results of the algorithms implemented in WEKA. The choice of algorithm selection for building a model is an important aspect of machine learning problems. The selection of the optimal algorithm should not be based on a singular metric such as accuracy that is mostly chosen by researchers.

#### 5.1 Comparison Based on Accuracy (Correctly Classified Instances)

The result of both Holdout and Cross-Validation methods for Binary Classification showed that the Accuracy outcome for the algorithms are closely related. In Hold-out, SDG and J48 performed excellent equally with 100% Accuracy, followed by LRN. On the other hand, in 10-F C-V, J48 has the best performance followed by SMO then SDG. However, LRN performance dwindled in relation to others. Table 1 shows the comparison of accuracy results.

Hence, from comparison in Table 1 according to correctly classified instances, the results shows that SDG, J48 and LRN can be used on one hand in Hold-Out as candidate algorithms for model building. Similarly, J48, SDG and IBK can be used on the other hand in 10-F C-V as candidate algorithms for identifying SQL Injection signatures in SQL query strings for effective mitigation as depicted in Figure 2.

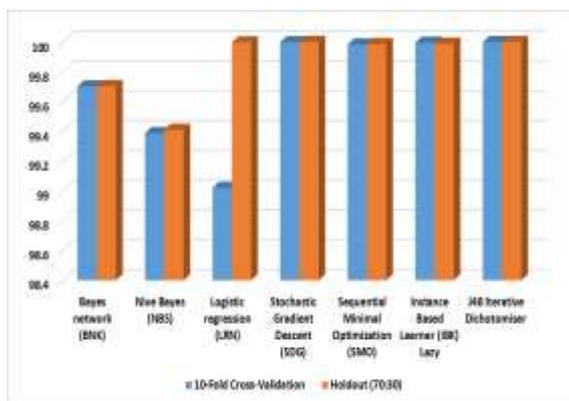


Figure 2. Comparison based on Accuracy of Hold-Out and Cross-Validation Methods

Beyond accuracy, AUC should be keenly considered in algorithm selection and model building [40]. This is highly essential as there could be a lot of false positives as a result of over-fitting and noise in the dataset.

#### 5.2 Comparison Based on Sensitivity (True Positive Rate)

The result of both Holdout and Cross-Validation methods for algorithm classification showed that the Sensitivity outcome for four algorithms such as SDG, SMO, IBK and J48 are the same in both Hold-Out and 10-F C-V methods, thus choosing the best classifier might be biased for model building. Equally, LRN had 100% Sensitivity in Hold-Out method whereas had 99% in 10-F as the least performance. Meanwhile, NBS performed woefully in comparison to other MLAs in Hold-Out method with 99.4% as shown in Table 2.

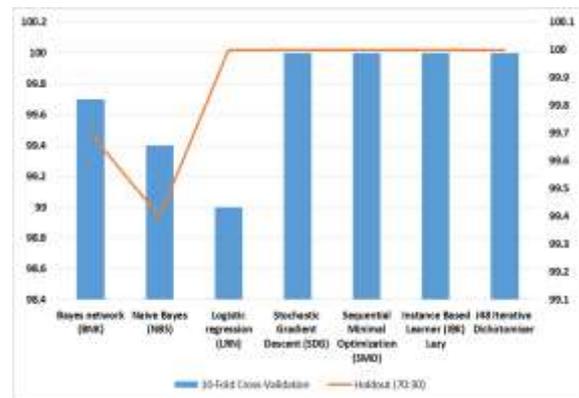


Figure 3. Comparison based on Sensitivity of Hold-Out and Cross-Validation Methods

Figure 3 shows the relationship among various algorithms in terms of sensitivity both in 10 fold Cross validation and Hold-Out techniques.

Therefore, from comparison in Table 2, according to sensitivity, the results shows that J48, SMO, SDG, IBK and LRN can be used on one hand in Hold-Out as candidate algorithms for model building. Likewise, J48, SDG, SMO and IBK can be used on the other hand in 10-Fold Cross Validation as candidate algorithms for detecting SQL Injection signatures in SQL query strings for effective mitigation as depicted in Figure 3. This shows that sensitivity cannot be used in isolation in choosing an optimal algorithm for building the model.

#### 5.3 Comparison Based on Specificity (True Negative Rate)

The outcome of both Holdout and Cross Validation methods for Binary Classification showed that the Specificity outcome for five algorithms such as SDG, SMO, IBK, LRN and J48 are the same in Hold-Out method and four algorithms except LRN, BNK and NBS are the same in 10-Fold Cross Validation method, thus choosing the best classifier might be confusing for model building without taking into consideration the various metrics concerned. Similarly, LRN had the least Sensitivity value in 10-Fold Cross Validation method whereas BNK had 99.3% as the least Sensitivity value in hold-out method as shown in Table 3.

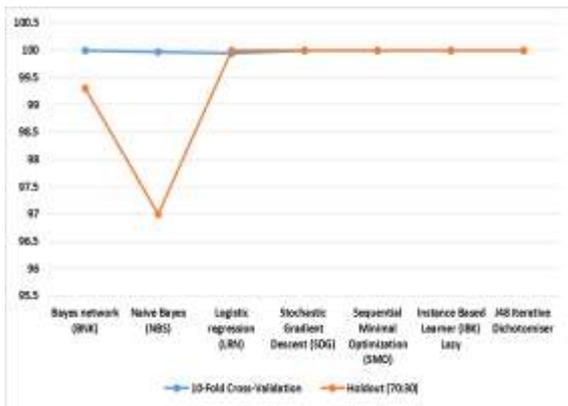


Figure 4. Comparison based on Specificity of Hold-Out and Cross-Validation Methods

Therefore, from comparison in Table 3, based on Specificity, the results revealed that J48, SMO, SDG, IBK and LRN can be used in building the model when Hold-Out is of concern. Equally, J48, SDG, SMO and IBK can be used for model building when 10-Fold Cross Validation is of importance in detecting SQL Injection signatures in SQL query strings for effective prevention of attacks as depicted in Figure 4.

#### 5.4 Comparison Based on Kappa-Statistic

The result of Binary Classification for the algorithms showed that SDG and J48 had the same Kappa-Statistic value of 100% for Hold-Out method. Likewise for 10-Fold Cross Validation with the same value of 99.99%. The least performed algorithm in Hold-Out was NBS and LRN in 10-Fold Cross Validation as revealed in Table 4.

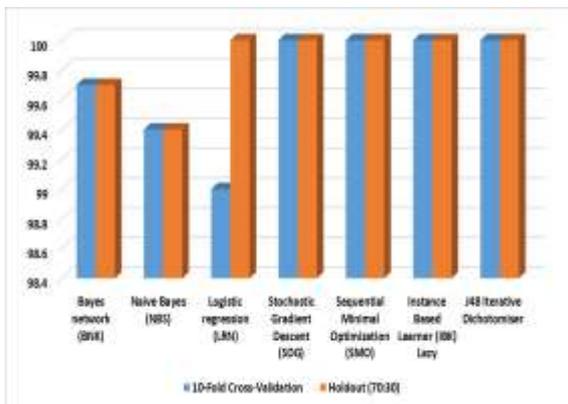


Figure 5. Comparison based on Kappa Statistic of Hold-Out and Cross-Validation Methods.

Therefore, based on comparison in Table 4 in relation to Kappa-Statistic, the results revealed that J48 and SDG can be used in building the model for both Hold-Out and 10-Fold Cross Validation methods. It is to be noted that, Kappa Statistic is a classifier performance measure that estimates the similarity between the members of an ensemble in a multi classifiers systems.

#### 5.5 Comparison Based on Time to Build (Time To Build (TTB))

The end result of algorithm classification for the algorithms showed that IBK had the least conceivable running time to build both at Hold-Out and 10-Fold Cross Validation with

values of 0.16second and 0.06 seconds, next in TTB is NBS with 4.09second and 4.95 respectively as shown in Table 5

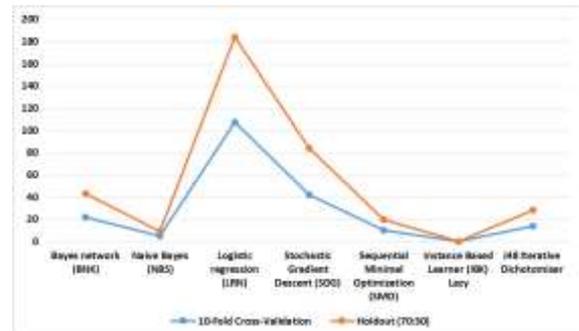


Figure 6. Comparison based on TTB of Hold-Out and Cross-Validation Methods

Therefore, based on comparison in Table 5 according to time to build the algorithms, the results showed that IBK and NBS were those with the least TTB and this does not connote their choice for building the algorithm. The TTB as shown in Figure 6 require the effective decision for optimal algorithm selection since these two algorithms do not have the same specificity and sensitivity values.

Table 6 and Table 7 shows the comparative analysis for all the metrics used in this study for both cross validation and hold-out ML algorithms performance evaluation techniques.

## 6. CONCLUSION

The results of the performance evaluation of the model for detection and classification of the SQLIA showed a good performance in terms of accuracy, true-positive rate, false-positive rate as well as time to build model. Pattern matching is capable of mitigating SQLIA requests through login, URL and search which can be implemented using machine learning paradigm. Machine learning algorithm selection for model building can be subjective and biased if necessary precautions are not put in place. Therefore, every performance metric must be considered holistically before choosing an optimal algorithm for predictive analytics.

According to [41] to have supervised predictive machine learning, ML algorithms require precise accuracy and minimum errors in addition to putting several factors into consideration. Also, it may be difficult or impossible to find a single classifier doing as well as a good group of classifiers if the only performance metric being utilized is best possible classification accuracy. [42] Opined that Multi-Criteria Decision Method (MCDM) methods can be used to find the optimal classification and regression models in relation to supervised machine learning algorithms.

## 7. REFERENCES

- [1] Sheykhkanloo, N. N. (2017). A Learning-based Neural Network Model for the Detection and Classification of SQL Injection Attacks. International Journal of Cyber Warfare and Terrorism. Available at: <https://dl.acm.org/doi/10.4018/IJCWT.2017040102>
- [2] Aliero, M. S., Ardo, A. A. Ghani, I. & Atiku, M. (2016). Classification of Sql Injection Detection and Prevention Measure. IOSR Journal of Engineering (IOSRJEN) [www.iosrjen.org](http://www.iosrjen.org), ISSN (e): 2250-3021, ISSN (p): 2278-8719 Vol. 6, Issue 2, Pp. 6-17. Available at:

- [https://www.iosrjen.org/Papers/vol6\\_issue2%20\(part-1\)/B06210617.pdf](https://www.iosrjen.org/Papers/vol6_issue2%20(part-1)/B06210617.pdf)
- [3] Teymourlouei, H. & Jackson, L. (2017). How Big Data Can Improve Cyber Security. International Conference on Advances in Big Data Analytics (ABDA).
- [4] Jagdish, H. (2008). "SQL Injection analysis, Detection and Prevention" (2008). Master's Projects. Available at: [http://scholarworks.sjsu.edu/etd\\_projects/82](http://scholarworks.sjsu.edu/etd_projects/82)
- [5] Thosar, S., Mane, A. Raykar, S., Jain, R., Khude, P. & Guru, S. (2016). Vulnerability Assessment using Logs as BIG DATA. International Journal of Engineering Science and Computing (IJESC) Available at: [https://doi: 10.4010/2016.1156, ISSN 2321 3361, 6 \(5\), 4644 – 4647](https://doi: 10.4010/2016.1156, ISSN 2321 3361, 6 (5), 4644 – 4647)
- [6] Gaurav, K. T. and Gaurav, O. (2012). Enhanced Query based Layered Approach towards detection and prevention of Web Attacks. Procedia Technology Vol. 4, Published by Elsevier Ltd. Pp. 500 – 505
- [7] OWASP (2018). Open Web Application Security Project (OWASP), OWASP Top Ten Project. [http://www.owasp.org/index.php/Category:OWASP Top Ten Project](http://www.owasp.org/index.php/Category:OWASP_Top_Ten_Project).
- [8] Fu, X., Lu, X., Peltsverger, B., Chen, S., Qian, K. & Tao, L. (2007). A static analysis framework for detecting SQL injection vulnerabilities. In Proceedings of the 31st Annual International Computer Software and Applications Conference – Volume 01, COMPSAC, 87–96, Washington, DC, USA, 2007. IEEE Computer Society.
- [9] Friedl, S. (2005). SQL injection attacks by example
- [10] Ali, S, Rauf, A. & Javed, H. (2009). SQLIpa: An authentication mechanism against sql
- [11] Stasinopoulos, A., Ntantogiany, C. & Xenakisz, C. (2018). Commix: Automating Evaluation and Exploitation of Command Injection Vulnerabilities in Web Applications
- [12] Acunetix (2018). Types of SQL Injection (SQLI). Available at: <https://www.acunetix.com/websecurity/sql-injection2/>
- [13] Oracle (2018). Types of SQL Injection Attacks. Available at: [http://download.oracle.com/oll/tutorials/SQLInjection/html/lesson1/les01\\_tm\\_attacks.htm](http://download.oracle.com/oll/tutorials/SQLInjection/html/lesson1/les01_tm_attacks.htm)
- [14] Halfond, W.G.J., J. Viegas, & A. Orso, A. (2006). A classification of SQL injection attacks and countermeasures. In Proceedings of the IEEE International Symposium
- [15] Khari, M. & Kumar, N. (2013). SQLIA Detection And Prevention Approaches A Survey. International Journal of Computer Science & Information Technology, 3(5)
- [16] Kindy, D.A. & Pathan, A. S. K. (2013). A Detailed Survey on various aspects of SQL Injection in Web Applications: Vulnerabilities, Innovative Attacks and Remedies. International Journal of Communication Networks & Information Security, 5(2).
- [17] Khari, M. & Kumar, N. (2013). SQLIA Detection And Prevention Approaches A Survey. International Journal of Computer Science & Information Technology, 3(5)
- [18] Vinod, K. K & Jatin, D. D. (2013). Advanced Detecting and Defensive Coding Techniques to prevent SQLIAs in Web Applications A Survey. International Journal of Science and Modern Engineering (IJISME), 1(6), 26 - 31 Available at: <https://docplayer.net/2925830-Advanced-detecting-and-defensive-coding-techniques-to-prevent->
- [19] Roy, S., Singh, A.K. & Sairam, A. S. (2011). Detecting and Defeating SQL Injection Attacks. International Journal of Information and Electronics Engineering, 1(1).
- [20] Srivastava, S., (2012). A Survey On: Attacks due to SQL injection and their prevention method for web application.
- [21] Borade, M.R. & Deshpande, N.A. (2013). Extensive Review of SQLIA's Detection and Prevention Techniques. International Journal of Emerging Technology and Advanced Engineering, Vol. 3 No.10
- [22] Khochare, N., Chalurkar, S., Kakade, S. & Meshram, B. B. (2011). Survey on SQL Injection attacks and their Countermeasures. International Journal of Computational Engineering & Management (IJCEM). 14, 111 – 114 Available at: [https://www.ijcem.org/papers102011/ijcem\\_102011\\_19.pdf](https://www.ijcem.org/papers102011/ijcem_102011_19.pdf)
- [23] Narayanan, S. N., Alwyn, R. P., & Mohandas, R. (2011). "Detection and Prevention of SQL Injection Attacks Using Semantic Equivalence." Computer Networks and Intelligent Computing. 103-112, Springer Berlin Heidelberg,
- [24] Cheon, E. H., Zhongyue, H. & Yon S. L. (2013). Preventing SQL Injection Attack Based on Machine Learning. International Journal of Advancements in Computing Technology 5(9)
- [25] Anwaar, A., Junaid, Q., Raihan, R., Arjuna, S., Andrej, Z. & Jon C. (2016). Big data for development: applications and techniques. Big Data Analytics, 1 / 2, 1- 24. ISSN: 2058-6345. Available at: <https://doi: 10.1186/s41044-016-0002-4>
- [26] Fatima, M., & Pasha, M. (2017). Survey of Machine Learning Algorithms for Disease Diagnostic. Journal of Intelligent Learning Systems and Applications. <https://doi.org/10.4236/jilsa.2017.91001>
- [27] Brownlee, J. (2016). Supervised and Unsupervised Machine Learning Algorithms. BSD (2013). 3-clause

- BSD license Available at:  
<https://github.com/rapid7/metasploit-framework/blob/master/LICENSE>
- [28] Zhu, X., & Goldberg, A. B. (2007). Semi-Supervised Learning Tutorial. ICML. <https://doi.org/10.2200/S00196ED1V01Y200906AIM006>
- [29] Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: a review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8), 1798-1828. <https://doi.org/10.1109/TPAMI.2013.50>
- [30] Krizhevsky, A., Sutskever, I., Hinton, G. E., Levine, S., Finn, C., Darrell, T., ... Szegedy, C. (2012). ImageNet Classification with Deep Convolutional Neural Networks Alex. *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*. <https://doi.org/10.1007/s13398-014-0173-7.2>
- [31] Logistic Regression (LR) pp. 223 – 237. Available at: <https://www.stat.cmu.edu/~cshalizi/uADA/12/lectures/ch12.pdf>
- [32] Ali, S. & Smith, K. A. (2006). On learning algorithm selection for classification. *Applied Soft Computing*, 6, 119–138.
- [33] Kotsiantis, S. B. (2007). Supervised Machine Learning: A Review of Classification Techniques. *Informatica* 31 (2007). 249 – 268. Retrieved from IJS website: <http://wen.ijs.si/ojs-2.4.3/index.php/informatica/article/download/148/140>. pp. 249 – 268
- [34] Kecman, V. (2005). Support vector machines: An introduction in support vector machines: Theory and applications. In L. Wang (Ed.), 1–47. Berlin: Springer-Verlag. Chapter 1
- [35] Taiwo, O. A. (2010). Types of Machine Learning Algorithms, *New Advances in Machine Learning*, Yagang Zhang (Ed.), ISBN: 978-953-307-034-6, InTech, University of Portsmouth United Kingdom, 3 – 31. Available at InTech open website: <http://www.intechopen.com/books/new-advances-in-machine-learning/types-of-machine-learning-algorithms>
- [36] Ko, M., Twari, A., & Mehmen, J. (2010). A review of soft computing applications in supply chain management. *Applied Soft Computing*, 10, 661–664.
- [37] Korting, T. S. (n.d)."C4. 5 algorithm and Multivariate Decision Trees." Image Processing Division, National Institute for Space Research--INPE.
- [38] Alex S. & Vishwanathan, S.V.N. (2008). *Introduction to Machine Learning*. Published by the press syndicate of the University of Cambridge, Cambridge, United Kingdom. Copyright © Cambridge University Press 2008. ISBN: 0-521-82583-0.
- [39] Schapire, R. (n.d). *Machine Learning Algorithms for Classification*.
- [40] Akinsola, Jide E. T., Adeagbo, Moruf A., Awoseyi, Ayomikun A. Breast Cancer Predictive Analytics Using Supervised Machine Learning Techniques. *International Journal of Advanced Trends in Computer Science and Engineering* 8 (6), 3095- 3104, ISSN 2278-3091, November – December 2019. Available Online at <http://www.warse.org/IJATCSE/static/pdf/file/ijatcse70862019.pdf>. DOI: <https://doi.org/10.30534/ijatcse/2019/70862019>
- [41] F. Y. Osisanwo, J. E. T. Akinsola, O. Awodele, J. O. Hinmikaiye, O. Olakanmi and J. Akinjobi. Supervised Machine Learning Algorithms: Classification and Comparison. *International Journal of Computer Trends and Technology (IJCTT) – Volume 48 Number 3*, 2017, <https://doi.org/10.14445/22312803/IJCTT-V48P126>
- [42] J. E. T. Akinsola, S. O. Kuyoro, O. Awodele & F. A. Kasali. Performance Evaluation of Supervised Machine Learning Algorithms Using Multi-Criteria Decision Making Techniques. *International Conference on Information Technology in Education and Development (ITED) Proceedings*, 17 – 34, 2019.

**Table 1. Comparison based on Accuracy for Hold-Out and Cross-Validation Methods**

Machine Learning Algorithms	10-Fold Cross-Validation	Holdout (70:30)
Bayes network (BNK)	99.7015	99.7035
Naive Bayes (NBS)	99.3874	99.41
Logistic regression (LRN)	99.0204	99.9971
Stochastic Gradient Descent (SDG)	99.998	100
Sequential Minimal Optimization (SMO)	99.9824	99.9856
Instance Based Learner (IBK) Lazy	99.9951	99.9885
J48 Iterative Dichotomiser	99.999	100

**Table 2. Comparison based on Sensitivity for Hold-Out and Cross-Validation Methods**

Machine Learning Algorithms	10-Fold Cross-Validation	Holdout (70:30)
Bayes network (BNK)	99.7	99.7
Naive Bayes (NBS)	99.4	99.4
Logistic Regression (LRN)	99	100
Stochastic Gradient Descent (SDG)	100	100
Sequential Minimal Optimization (SMO)	100	100
Instance Based Learner (IBK) Lazy	100	100
J48 Iterative Dichotomiser	100	100

**Table 3. Comparison based on Specificity for Hold-Out and Cross-Validation Methods**

Machine Learning Algorithms	10-Fold Cross-Validation	Holdout (70:30)
Bayes network (BNK)	99.993	99.3
Naive Bayes (NBS)	99.969	97
Logistic Regression (LRN)	99.945	100
Stochastic Gradient Descent (SDG)	100	100
Sequential Minimal Optimization (SMO)	100	100
Instance Based Learner (IBK) Lazy	100	100
J48 Iterative Dichotomiser	100	100

**Table 4. Comparison based on Kappa-Statistic for Hold-Out and Cross-Validation Methods**

Machine Learning Algorithms	10-Fold Cross-Validation	Holdout (70:30)
Bayes network (BNK)	98.39	98.37
Naive Bayes (NBS)	96.67	96.72
Logistic Regression (LRN)	94.63	99.98
Stochastic Gradient Descent (SDG)	99.99	100
Sequential Minimal Optimization (SMO)	99.9	99.92
Instance Based Learner (IBK) Lazy	99.97	99.94
J48 Iterative Dichotomiser	99.99	100

**Table 5. Comparison Based on TTB for Hold-Out and Cross-Validation Methods**

Machine Learning Algorithms	10-Fold Cross-Validation	Holdout (70:30)
Bayes network (BNK)	22.06	21.1
Naive Bayes (NBS)	4.95	4.09
Logistic Regression (LRN)	107.6	76.56
Stochastic Gradient Descent (SDG)	42.25	41.96
Sequential Minimal Optimization (SMO)	10.15	9.71
Instance Based Learner (IBK) Lazy	0.06	0.16
J48 Iterative Dichotomiser	14.12	14.28

**Table 6. Summary of the Model Performance in Cross Validation Method**

Machine Learning Algorithms	ACC	TP_R	TN_R	Kappa Statistics	TTB
BNK	99.7015	99.7	99.993	98.39	22.06
NBS	99.3874	99.4	99.969	96.67	4.95
LRN	99.0204	99	99.945	94.63	107.6
SDG	99.998	100	100	99.99	42.25
SMO	99.9824	100	100	99.9	10.15
IBK	99.9951	100	100	99.97	0.06
J48	99.999	100	100	99.99	14.12

**Table 7. Summary of the Model Performance in Hold-Out Method**

Machine Learning Algorithms	ACC	TP_R	TN_R	Kappa Statistics	TTB
BNK	99.7035	99.7	99.3	98.37	21.1
NBS	99.41	99.4	97	96.72	4.09
LRN	99.9971	100	100	99.98	76.56
SDG	100	100	100	100	41.96
SMO	99.9856	100	100	99.92	9.71
IBK	99.9885	100	100	99.94	0.16
J48	100	100	100	100	14.28

# War against Coronavirus (COVID – 19) in Tanzania: Designing a Low Cost Automatic Water Tap

Herbert Wang  
Department of Information  
Technology  
University of Iringa (UoI)  
Iringa, Tanzania

Thobius Joseph  
Department of Information  
Technology  
University of Iringa (UoI)  
Iringa, Tanzania

Mauna Belius Chuma  
Department of Information  
Technology  
University of Iringa (UoI)  
Iringa, Tanzania

---

**Abstract:** In December 2019, an outbreak of severe acute respiratory syndrome coronavirus 2 (SARSCoV-2) infection occurred in Wuhan, Hubei Province, China and spread across China and beyond. On February 12, 2020, WHO officially named the disease caused by the novel coronavirus as Coronavirus Disease 2019 (COVID-19). On January 30, 2020, WHO declared COVID-19 as the sixth public health emergency of international concern. One of the major preventions is frequent hand washing. An attempt has been made to design a low-cost automatic water tap as a measure to fight the spread of Coronavirus (Covid-19). The designed system uses simple, easy to get and low cost instruments. The system does not require a user to touch the water tap. It uses microcontroller to implement intelligence. The solar panel is used as a source of power. The system design uses green solution based automatic water tap. The system will be used in gatherings to avoid Coronavirus spreading. Such gathering places include hospitals, markets, bus stops, public transports, churches, mosques, and restaurants.

**Keywords:** coronavirus; covid-19; automatic water tap; Wuhan pneumonia; Arduino

---

## 1. INTRODUCTION

### 1.1 Background

Coronavirus disease (COVID-19) is caused by SARS-COV2 and represents the causative agent of a potentially fatal disease that is of great global public health concern [1].

Since December 2019, an outbreak of corona virus disease 2019 (COVID-19) occurred in Wuhan, and rapidly spread to almost all parts of China [2]. An outbreak of a novel coronavirus (COVID-19 or 2019-CoV) infection has posed significant threats to international health and the economy [3]. The outbreak of COVID-19 raised intense attention not only within China but internationally [2].

On January 30, 2020, WHO has declared COVID-19 as the sixth public health emergency of international concern [4]. As of February 9, 2020, data from the World Health Organization (WHO) has shown that more than 37,000 confirmed cases have been identified in 28 countries, with more than 99% of the cases being detected in China [4]. As of 3 March 2020, a total of 80,151 cases have been detected and confirmed in Mainland China. Internationally, there are more than 10,566 additional cases detected and confirmed in 72 countries [5]. On 23rd March 2020, WHO rolling update showed Coronavirus disease (COVID – 19) outbreak situation of 294, 110 confirmed cases, 12, 944 confirmed deaths, and 187 countries, areas or territories with cases.

The most common symptoms at onset of COVID-19 illness are fever, cough, and fatigue, while other symptoms include sputum production, headache, haemoptysis, diarrhea, dyspnea, and lymphopenia [1]. Clinical features of COVID-19 include dry cough, fever, diarrhea, vomiting, and myalgia [6].

It is spread by human-to-human transmission via droplets or direct contact, and infection has been estimated to have mean incubation period of 6.4 days and a basic reproduction number of 2.24 - 3.58 [4]. Person-to-person transmission occurs

primarily via direct contact or through droplets spread by coughing or sneezing from an infected individual [1]

### 1.2 Health Risk in Tanzania

According to [7] Tanzania is sparsely populated with population density of 51 persons per square kilometer with variation across regions. Persons are concentrated in Dar es Salaam and *Mjini Magharibi* regions with population densities of 3,133 and 2,581 persons per square kilometer respectively. This population density increases health risk.

To fight Covid-19, Tanzania has banned all academic and unnecessary social gatherings. Academic gatherings closed include kindergarten, primary and secondary schools, middle level colleges and Universities. Socially, Tanzania has banned unnecessary gatherings such as sports, fiesta, music performances, theatres, until further notice.

Due to low or unavailability of automatic water taps, these gatherings depend on low cost mechanical water taps. One of the problems of using mechanical water taps for hand washing is in its method of operation. Mechanical water taps need user to open handle and wash hands, then close the handle again, causing contamination.

### 1.3 Problem statement

It is perhaps clear that quarantine alone may not be sufficient to prevent the spread of COVID-19, and the global impact of this viral infection is one of heightening concern [6]. According to [6], Basic hand hygiene measures are recommended, including frequent hand washing and the use of PPE such as face masks.

Water is an indispensable input for industrial production and human existence. Currently, people are increasingly concerned about their health, and their interest in the safety of water [8]. According to [9] Clean water is a critical resource, important for the health and well-being of all humans.

The challenge is, people use mechanical water taps. Just after washing their hands they touch the same taps, causing contamination.

The recent advances in technology have enabled the use of electronic gadgets in everyday life [10]. Everything is getting automatic and smarter with the passing of days. According to [11] automatic water faucets (taps) have become a staple in public places. They are convenient, energy efficient, hygienic and, according to manufacturers, water efficient.

Therefore, this paper comes up with designing a low cost automatic water tap which does not require a user to touch the water tap. The low cost, and green solution based automatic water tap will be used in gatherings to avoid coronavirus spreading. Such gathering places include hospitals, markets, bus stops, churches, mosques, schools, colleges, Universities, and restaurants.

#### 1.4 Added value of this study

- Instead of developing a complete basin for standard hand-wash, this paper aims to design an automatic hand wash tap, which can be used simply by replacing the standard tap with the developed one. It is designed for simpler solution of performing hand-wash standard procedure instead of buying the whole housing of hand wash machine, hence a low cost automatic water tap.
- Other automatic water taps use solenoid, Arduino, IR sensors in creation of such circuits. These projects were machine centered designed, in such that the projects objectives were mainly on producing automation functionality and enthusiasm with technology. Arduino is suitable for prototyping, but not for led to go design for industrial production.
- The designed system is Plug and play.
- No knowledge of electronics is necessary to set-up and use the sensor. Low cost
- Uses solar power instead of electricity.
- Portable, small size.
- Covered to protect against direct sunlight.
- Not affected by blockage. Other systems do not need any blockage like cockroaches.

#### 1.5 Paper organization

The remainder of the paper is organized as follows. Section 2 identifies related works. Section 3 presents designing the automatic water tap. Next, Section 4 offers some concluding remarks, finally section 5 gives acknowledgement.

### 2. RELATED WORKS

A study by [12] designed a system about Tap control system using the smart phones and the Arduino boards. This can be used to control the taps for plant watering and other taps in the home through internet by sitting anywhere in the world. Arduino 2560 board with Ethernet Shield is used in this project. Arduino is coded using Arduino codes and the app development is done using Android programming.

In their study [13] designed and built an automatic sensor-based water tap. The product can automatically close the tap when the user's hand is occupied and can't hold any water. However, the product is not widely adopted, since the product

needs electrical power, so it needs high initial investment to build the electrical infrastructure, which most of pesantren or mosque didn't possess.

Another study was conducted in Indonesia by [14]. They developed a prototype of solar-powered, automatic water ablation tap. The system uses ultrasonic sensors to detect human presence and employs an Arduino microcontroller. It sends the "on" and "off" instructions to the relay to activate the solenoid valve to allow water flow.

According to the study by [15] they designed and implemented an automatic faucet that is user-friendly and easy to plug in regular water pipe for standard hand-washing routine. The faucet utilized IR proximity sensor to detect presence of hands that would automatically commence the hand-washing process. First, it exited water and soap simultaneously to force users to use soap. The scrubbing duration was marked by buzzer's sound and LED's light.

The mentioned systems were designed in a form of integrated automatic basin. The machines were required to be assembled as a complete basin system. Instead of developing a complete basin for standard hand-wash, this paper aims to design an automatic hand wash tap, which can be used simply by replacing the standard tap with the developed one. It is designed for simpler solution of performing hand-wash standard procedure instead of buying the whole housing of hand wash machine, hence a low cost automatic water tap.

### 3. DESIGNING THE AUTOMATIC WATER TAP

#### 3.1 Selection of sensor

Three candidates were investigated. These were Infrared transceiver module, Slotted optocoupler/ photoelectric switch, and Ultrasonic Distance Sensor - HC-SR04 IR sensor.



Figure 1 Infrared transceiver module.



Figure 2 Slotted optocoupler/ photoelectric switch.

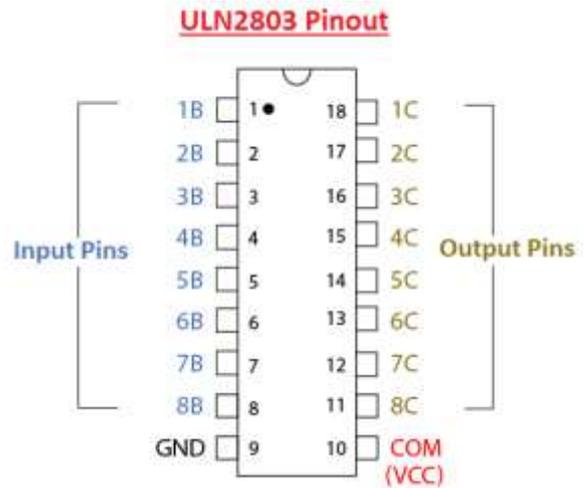


Figure 4 Attiny85V

To upload code to Attiny85 microcontrollers during industrial mass production you will need a programmer board. The Arduino Uno is selected as a programmer for this design. It will be connected direct to pins interfaces of Attiny during upload codes. Figure 5 below shows the connections for mass production of Attiny85 with code during production.



Figure 3 Ultrasonic Distance Sensor - HC-SR04 IR sensor.

The design uses Infrared transceiver module. It is small in size and it provides similar functionality as Ultrasonic Distance Sensor - HC-SR04 IR sensor.

The Slotted optocoupler/ photoelectric switch would be a best choice. However, its means of operation demands a placement of thin item about 3 mm inside the module. This would increase contact of shared items, hence unsuitable for the purpose.

### 3.2 Selection of microcontroller in processing unit

The selection of a microcontroller plays very important role in any embedded system. According to the need of the system a microcontroller is chosen [16]. Here in this system in order to design a low cost automatic water tap, Attiny85V microcontroller is used. The microcontroller has small size of 0.34 inches in length and 0.3 inches in width (ATtiny85V datasheet, 2020). It has 8 K Bytes of In-System Programmable Program Memory Flash which is enough to provide intelligence that is required to control solenoid.

In this paper we design a microcontroller based automatic water tap. The microcontroller allows implementation of intelligence in the applications.

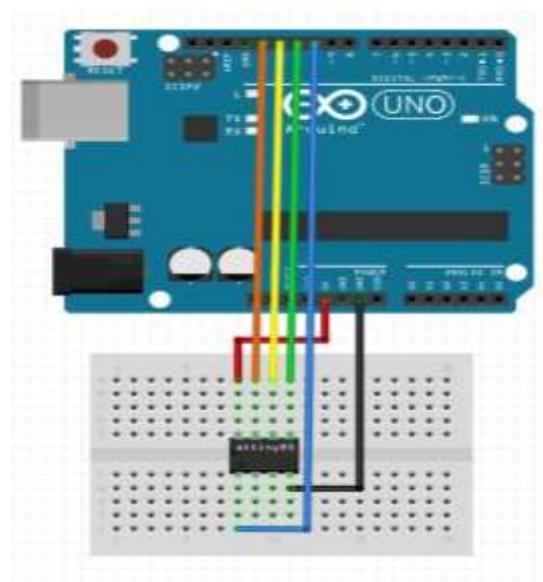


Figure 5 Attiny programming with Arduino Uno board.

The ULN2803 Darlington Transistor module will be used. This module has a capability to work with high voltage up to 50V. It has 8 input and 8 output connections respectively, in one to one mapping mode, meaning connection input one will produce output in output pin one. Each output pin can produce up to 500 milliAmpere. You can connect two input or more to produce additive output current of 1 Ampere or more. It has small size and fewer prices than relay and more capable than a TIP 120 transistor



Figure 6 ULN2803 transistor.

### 3.3 Selection of solenoid valve

DC6-12V MINI Aquarium Water Pump is selected for this design. The pump can push even hot water less than 60 Celsius. Its advantages include; relatively cheaper than mini servo motors, and you do not have to worry for positive and negative pole of connections.



Figure 7 Mini water pump.

### 3.4 Designed Automatic Water Tap

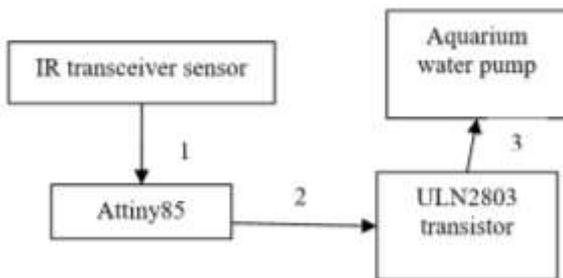


Figure 8 Block diagram of the system.

The block design shows communication flow of integrated equipment of a system. The Attiny85 will be embedded with

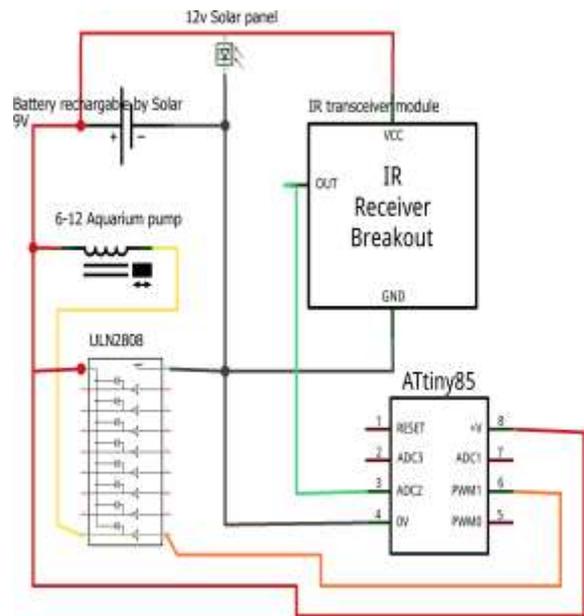


Figure 9 Circuit diagram of the system.

code by using Arduino Uno. Among conditions that code will control are duration of water flow and change of input signals, which would be resulted by hand motion in front of the sensor. The ATtiny will only send signal to ULN2803 amplifier after receiving signal from sensor and these conditions meet. The aquarium water pump is DC water pump, which works whenever supplied with minimal operating ratings. The ULN2803 will allow the rating to pump only after it receives signal from microcontroller.

From the diagram

Red line = Vcc power connection

Black line = Ground power connection

Yellow line= Input to pump from ULN2803/8 transistor

Orange line = Input to transistor from microcontroller

Green line = Input to microcontroller from sensor

The whole circuit excluding water pump will have about 6 cm in length and 4 cm in width. With PID package the components can interconnect even with double sided PCB board to reduce overall cost of production.

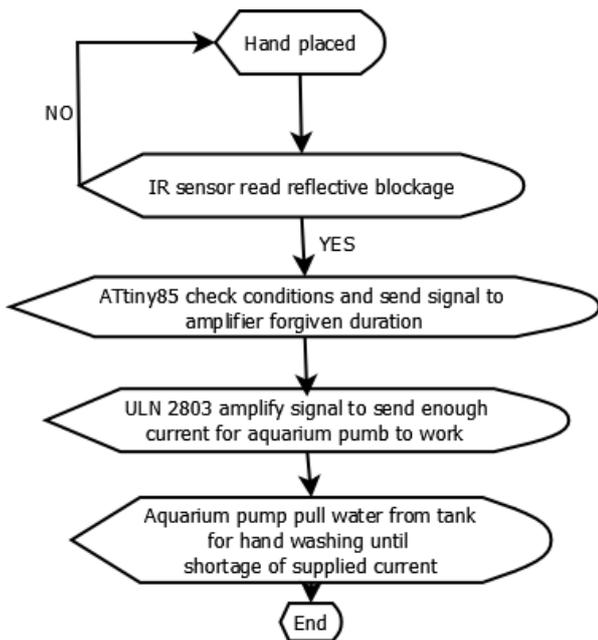


Figure 10 System operation flow diagram

#### 4. CONCLUSION AND FUTURE WORK

In this paper, a low cost instrumentation system has been designed for automatic water tap which will be used to fight against Coronavirus (Covid-19). The designed system uses simple, easy to get and low cost instruments. The system does not require a user to touch the water tap. It uses microcontroller to implement intelligence, like to allow flow of water for certain duration of time to ensure proper hand washing. The solar panel will be used as a source of power. The system will use green solution based automatic water tap which will be used in gatherings to avoid Coronavirus spreading. Such gathering places include hospitals, markets, bus stops, churches, mosques, and restaurants.

Further research work is required in this direction to implement the system towards lower price for production and durability, with high sensitivity of the sensor. An implementation of the low cost automatic water tap can be extended to other liquids such as Sanitizer, milk, and juice.

#### 5. ACKNOWLEDGMENTS

Authors hereby express the gratitude to the University of Iringa Staff and Students for the support.

#### 6. REFERENCES

[1] H. A. Rothan and S. N. Byrareddy, "The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak," *J. Autoimmun.*, no. February, p. 102433, 2020, doi: 10.1016/j.jaut.2020.102433.

[2] H. Luo *et al.*, "Can Chinese Medicine Be Used for Prevention of Corona Virus Disease 2019 (COVID-19)? A Review of Historical Classics, Research Evidence and Current Prevention Programs," *Chin. J. Integr. Med.*, vol. 11655, no. 100029, pp. 1–8, 2020, doi: 10.1007/s11655-020-3192-6.

[3] L. Zhang and Y. Liu, "Potential interventions for novel coronavirus in China: A systematic review," *J. Med.*

*Virol.*, no. January, pp. 479–490, 2020, doi: 10.1002/jmv.25707.

[4] C. C. Lai, T. P. Shih, W. C. Ko, H. J. Tang, and P. R. Hsueh, "Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges," *Int. J. Antimicrob. Agents*, vol. 2, p. 105924, 2020, doi: 10.1016/j.ijantimicag.2020.105924.

[5] M. Chinazzi *et al.*, "The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak.," *Science*, vol. 9757, no. March, pp. 1–12, 2020, doi: 10.1126/science.aba9757.

[6] C. Sohrabi *et al.*, "World Health Organization declares Global Emergency: A review of the 2019 Novel Coronavirus (COVID-19)," *Int. J. Surg.*, 2020, doi: 10.1016/j.ijsu.2020.02.034.

[7] URT, "The United republic of Tanzania. National Bureau of Statistics: 2012 Population and Housing Census Population Distribution by Administrative areas," 2013.

[8] S. Kwak, S. Yoo, and C. Kim, "Measuring the Willingness to Pay for Tap Water Quality Improvements: Results of a Contingent Valuation Survey in Pusan," *Water*, vol. 124, pp. 1638–1652, 2013, doi: 10.3390/w5041638.

[9] T. P. Lambrou, C. G. Panayiotou, and C. C. Anastasiou, "A low-cost system for real time monitoring and assessment of potable water quality at consumer sites," *Proc. IEEE Sensors*, pp. 3–6, 2012, doi: 10.1109/ICSENS.2012.6411190.

[10] H. Peter Wanga, "Designing a Machine Learning – Based Framework for Enhancing Performance of Livestock Mobile Application System," *Am. J. Softw. Eng. Appl.*, vol. 4, no. 3, p. 56, 2015, doi: 10.11648/j.ajsea.20150403.13.

[11] A. Harmon, "Do automatic water faucets actually save water?," 2016.

[12] K. Vani and P. Shrinidhi, "Automatic Tap Control System in the Smart Home using Android and Arduino," *Int. J. Comput. Appl.*, vol. 127, no. 8, pp. 19–23, 2015, doi: 10.5120/ijca2015906407.

[13] Agustawan and A. Hadi, "Efisiensi Rancang Bangun Keran Wudhu Otomatis Hemat Air," *Semin. Nas. Teknol. Informasi, Komunikasi, dan Ind.* 9, pp. 18–19, 2017.

[14] S. Adriansyah Hulukati, T. Pratiwi Handayani, R. Jaya, and S. Abdussamad, "A prototype of solar-powered automatic ablution tap," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 486, no. 1, 2019, doi: 10.1088/1757-899X/486/1/012078.

[15] H. Fitriyah, E. Rosana Widasari, E. Setiawan, and B. Angga Kusuma, "Interaction design of automatic faucet for standard hand-wash," *MATEC Web Conf.*, vol. 154, pp. 1–7, 2018, doi: 10.1051/mateconf/201815403003.

[16] R. Sood, M. Kaur, and H. Lenka, "Design and Development of Automatic Water Flowmeter," *Int. J. Comput. Sci. Eng. Appl.*, vol. 3, no. 3, pp. 49–59, 2013, doi: 10.5121/ijcsea.2013.3306.

# Examination of Corporate Social Reporting and Financial Performance

Cesar Castellanos  
Walden University  
Minneapolis MN, USA

Ify S. Diala-Nettles  
Contributing Faculty  
Walden University  
Minneapolis MN, USA

---

**Abstract:** Organizational leaders around the world spend millions of dollars on ineffective corporate social responsibility (CSR) programs and CSR reporting strategies. Understanding the relationship between CSR reporting, CSR indices (CSRi), and financial performance is necessary to minimize unnecessary expenditures among organizational leaders. The purpose of this quantitative correlational study, grounded in Frederick's CSR theory and Freeman's stakeholder theory, was to examine the relationship between CSR reporting, CSRi, and financial performance of hardware and software organizations. Data were collected from the Security Exchange Commission and the official websites of 25 hardware and software organizations that were part of Fortune 500 between the years 2010-2015. The results of the multiple linear regression indicated that there was no statistically significant relationship between CSR reporting, CSRi, and net income. Similarly, no significant relationship existed between CSR reporting, CSRi, and return on assets. The implications for social change include the development of socially responsible strategies that take into consideration the ethical variables of dignity and respect and the uncertainties faced by individuals within the community.

**Keywords:** Corporate social responsibility (CSR); CSR reporting; CSR indices (CSRi); Financial performance; Sustainability; Stakeholders.

---

## 1. INTRODUCTION

Corporate social responsibility (CSR) is a topic of constant debate among scholars and business practitioners. Numerous researchers and practitioners advocate the view that CSR strategies encompass the efforts conducted by organizational leaders to achieve social stability. Frynas and Yamahaki (2016) pointed out that throughout history, CSR strategies were perceived as crisis handling mechanisms. Marples (2017) advocated the view that during a crisis management process, organizational leaders try to report their crisis-mitigating approaches, hoping to safeguard the already questioned ethical standpoint. CSR reporting strategies are to be used as planning mechanisms and should be included in today's complex business models. To account for the holistic view that governs our value-maximization business culture, it was necessary to employ a quantitative correlational model to confirm or deny the level of relationship between CSR reporting, CSRi, and financial performance. Understanding that CSR involvement may decrease information irregularities and potential debts associated with non-socially accepted activities (Svantesson, Silén, & James, 2017), it was essential to indulge the income-driven business population by examining accounting-based measures such as net income (NI) and return on assets (ROA) as the dependent variables. The sample's annual reports filed to the Security Exchange Commission (SEC) between the years 2010-2015 and the sample's CSR indices (as provided by the Reputation Institute), served as critical components of the data gathering process. To facilitate the future reproduction of this study, a relatively easy data analysis model was employed to assess the level of relationship between the variables. If duplicated in the future, the applied statistical model can facilitate the scrutiny process of organizational leaders interested in understanding the implications of adequately reporting CSR actions. As mentioned by CSR researchers such as Agudo, Garcés, and Salvador (2015), engaging in socially responsible activities can translate to positive economic performances.

## 2. PROBLEM & PURPOSE OF THE STUDY

Adverse profitability in the U.S. technology industry is associated with organizational leaders' inability to accurately report their corporate social responsibility strategies (Michelon, Pilonato, & Ricceri, 2015). During the years 2010-2015, 42% of hardware and software organizations listed in the United States experienced a change in their financial performance after employing or modifying their CSR reporting strategies. The general business problem was that various organizational leaders in the U.S technology industry do not use the proper CSR reporting tools to address stakeholders, resulting in a negative variation in their financial performance. The specific business problem was that limited reliable information is available for U.S. executives and business owners of small- and medium-sized hardware and software organizations in the United States to understand the relationship between CSR reporting, CSRi, and financial performance. The purpose of this quantitative correlational study was to examine the relationship between CSR reporting, CSRi, and financial performance. The independent variables were CSR reporting, obtained by coding the sample's annual reports to the SEC, and CSRi, as provided by the Reputation Institute. The dependent variables were NI and ROA. The targeted population was comprised of executives and business owners of small- and medium-sized software and hardware manufacturing organizations in the metropolitan area of Austin, Texas. This population was suitable for this study because 99% of software and hardware organizations in the United States are led by small- and medium-sized organizational leaders that do not understand the relationship between CSR reporting and financial performance. The implications for notable social change include the development of socially responsible strategies that take into consideration the ethical variables of dignity and respect, and the uncertainties faced by individuals in their communities.

### 3. DISCUSSIONS

#### 3.1 Background of CSR

##### CSR-1: 1930s to the 1970s

Scholars such as Babiak and Kihl (2018) argued that the notion of business responsibility emerged in the 20th century with the rise of an idea referred to as socially oriented business leaders. The conglomeration of historical events such as (a) World War I, (b) the rise of notorious dictatorships, (c) human rights movements, and (d) the United States' depression, marked the inauguration of the CSR concept. CSR supporters such as Fontana (2018) concurred that through history, business leaders had maintained an ambiguous understanding of business responsibility, but the concept of business social responsibility can be traced to 1937. To clarify the concept of business responsibility. To ease the understanding of CSR's complex evolutionary process, Frynas and Yamahaki (2016) cited an article published in the 1940s. Frynas and Yamahaki referenced this article to explain how the population of the 1940s viewed the concept of CSR. This article was published in a business magazine of the era, and it is constantly cited as the first scholarly attempt to confirm or deny a relationship between business responsibility and customer acceptance. Frynas and Yamahaki pointed out that the results showed in this article demonstrated that costumers of the epoch were willing to consume more from organizations that made an effort to improve their communities. Though Brown and William (2013) tried to determine the roots of the CSR theory by referencing topics such as organizational ethics and social responsibility, without a doubt, Frederick's (1978, 1986) work has been identified as the theoretical framework of the CSR field. Frederick (1978) argued that the history of CSR could be segmented into three notable evolutionary stages. Frederick determined that researchers belonging to the CSR-1 evolutionary stage started to address the notion that business leaders needed to consider the necessities of their communities and that there was a need for business professionals to develop business models that included strategic approaches tailored to the needs of their societies. Frederick observed that the CSR-1 period was filled with immense economic turmoil, not only because of the economic effects brought by the great depression of 1930 but also because of the involvement of the United States in World War II. Other scholars such as Abernathy, Stefaniak, Wilkins, and Olson (2017) and Frynas and Yamahaki (2016) comprehended that CSR's evolutionary process stopped as a reaction to the great depression of 1930; however, Frederick realized that the numerous social movements introduced by the post-World War II period offered a new way of associating businesses and societies. Frynas and Yamahaki explained that the first attempt to theorize CSR emerged from a new group of researchers schooled during and after World War II. The new wave of business leaders and scholars promoted the need for enhanced business management schooling and the establishment of employee-oriented human resource departments (see Brown & Zmora, 2015). Brown and Zmora indicated that scholars of the CSR-1 era began to discuss the benefits of becoming socially accepted organizational leaders. Frederick (1986) pointed out that business leaders of the postwar era understood the concept of social responsibility; however, a significant percentage of business leaders needed to understand factors such as:

- Businesses and their impact on the well-being of the citizens comprising their communities.

- The importance of not exploiting natural and human resources.
- The rise of a more educated and conscious society.
- The introduction of women to the workforce.
- The relationship between being a socially responsible organization and customer perception.

Researchers and practitioners of the CSR-1 period tended to define CSR as:

- Socially oriented organizational actions that range outside of the scope of standard business transactions.
- An organization's moral obligation towards society.
- The ethical and moral obligation of attending environmental issues, enhance the workforce, and meet social needs.
- The ability to allocate funds towards social and individual betterment.

In his second article, Frederick (1986) addressed the crucial points not discussed by scholars of the CSR-1 era and provided enough evidence to support the view that scholars of the CSR-2 and CSR-3 era were responsible for introducing CSR as a strategic managerial tool.

##### CSR-2: Early 1970s to Mid-1980s

Frederick (1978) coined CSR's second evolutionary stage as CSR-2. Abernathy et al. (2017) maintained that scholars of the CSR-1 era focused on the development of the necessary tools to facilitate formal research and the expansion of concepts that could fill literature gaps. Freeman and Dmytriyeu (2017) affirmed that researchers of the CSR-1 era tried to convince business leaders that engaging in CSR activities was the ethical and moral thing to do; however, the new wave of researchers and business professionals attempted to locate the necessary elements to engage in rewarding CSR activities. The need for a clear and concise definition of CSR was evident. Business leaders understood that there was a need for organizations to become socially responsible; however, no practical managerial tools had been addressed or introduced. Business leaders of the CSR-2 era struggled to find the answer to several fundamental questions:

- How can organizations develop suitable CSR strategies?
- What amount of resources should organizations allocate to the development and preservation of CSR programs?
- Is there a positive or negative relationship between CSR strategies and financial performance?

Frederick (1986) mentioned that the CSR-2 period was viewed as CSR's most notable revolutionary phase. Frederick pointed out that researcher of the CSR-2 era evaluated concepts such as (a) financial performance, (b) social involvement, and (c) the newly introduced stakeholder theory. Kang and Liu (2015) advocated the view that researchers of this era began to recognize the numerous stakeholders,

without obviating its primary goal, serving their communities through CSR programs and activities. Porter and Kramer (2007) concluded that the moment business professionals witnessed an attempt to measure CSR against financial performance was the moment business leaders began to accept the CSR theory as a strategical tool. Frederick (1986) clarified that researchers of the CSR-2 era rejected the notion that CSR was a trade-off between organizational leaders, and societies, and that there was a clear relationship between CSR and financial performance. Frederick also mentioned that within the CSR-2 period, an era of partnerships among business professionals of the same industries emerged. This macro approach materialized from the collaboration of leaders of multinational organizations who experienced ethical dilemmas on foreign soil and did not possess the ability or the financial capacity to resolve the ethical turmoil. According to Frynas and Yamahaki (2016), though researchers of the CSR-1 era were perceived as CSR's introductory living organisms, business professionals and scholars of the CSR-2 epoch embraced the revenue-driven mentality and decided to research CSR as an executing mechanism. As mentioned by Ranängen (2016), researchers of the CSR-2 epoch decided to further develop an analytical approach to CSR. Harrison, Freeman, and Sá de Abreu (2015) argued that this development allowed researchers to demonstrate that a clear relationship between CSR and stakeholder theories (see Harrison et al., 2015). The proposed relationship between the CSR and the stakeholder theories served as the fueling apparatus to increase the momentum and the interest of scholars who once showed fascination in examining CSR strategies and its relationship with financial performance (see Freeman & Dmytriiev, 2017). As explained by Yu and Choi (2016), scholars and professionals of the CSR-1 and CSR-2 era, introduced numerous revolutionary movements, but still experienced abundant shortcomings. Hategan, Sirghi, Curea-Pitorac, and Hategan (2016) argued that researchers of the CSR-2 period failed to (a) segregate CSR's ethical approach (b) to provide a model capable of measuring CSR (c) address the importance of adequately reporting CSR activities to stakeholders, and (d) clarify the vague concept of social betterment. Freeman and Dmytriiev (2017) commented that business professionals and scholars of the CSR-2 era failed to anticipate and mitigate the 1960's and late 1970's insertion of public awareness movements against businesses that profited from the exploitation of natural resources and human labor. Freeman and Dmytriiev emphasized that because practitioners of the upcoming CSR field did not develop strategies capable of combating the emerging social movements, lawmakers of the era were forced to increase the minimum wage and introduced multiple costly governmental directives that opened the doors of unionized employment movements. Brown and William (2013) argued that legislation was ultimately passed to calm the demands of a population that understood that organizations needed to be recognized as legitimate social citizens.

### **CSR-3: 1980 to the Present**

Frederick (1978) understood that researchers of the CSR field became stagnant after its second evolutionary stage. Frederick (1986) mentioned that during the CSR-3 period, scholars defined the concept of CSR as the economic, legal, and ethical actions taken by organizational leaders in pursuit of social betterment, without compromising the organization's current financial state. The newly adopted definition covered the gaps initially neglected by researchers of the CSR-1 era, and subsequently semi-covered by scholars of the CSR-2 epoch.

Kang, Germann, and Grewal (2016) pointed out that the definition presented by intellectuals of the CSR-3 period addressed the ethical, financial, and operational factors needed to gain credibility among business scholars and professionals. Yu and Choi (2016) argued that scholars of the CSR-3 era introduced a strategical operational approach designed to include and safeguard stakeholders such as (a) employees, (b) customers, (c) suppliers, (d) local communities, (e) government, and (f) international organizations. Recognizing that during the CSR-1 era business leaders were not capable of understanding the concept of CSR and that scholars of the CSR-2 period offered a vague understanding of CSR's core existence, academics of the CSR-3 era presented a revolutionary approach that postulated a clear understanding of the CSR concept from a strategical standpoint (see Frederick, 1986). Kang and Liu (2015) understood that researchers of the CSR-3 era focused on the development of ethical decision-making business leaders that could one day become capable of formulating socially-oriented strategies. Hsu and Cheng (2016) pointed out that the professional and academic literature presented by contemporary scholars offered a reactive solution to business leaders who tended to work towards the identification and the development of responsive measures that could alleviate any ethical or environmental dilemma. Smith and Colvin (2016) noted that business leaders and scholars tend to regularly address the benefits of openly reporting current and future CSR strategies to stakeholders. Leister and Maclachlan (2015) argued that business leaders of the CSR-3 period marked a pivotal point in CSR history, multinational executives and business leaders demanded the issuing of sustainability reports and decided to include their CSR actions in their letters to shareholders. Shareholders commenced to request supplementary CSR approaches, and organizational leaders understood that accomplishing or fulfilling those demands had a significant effect on market share value and future investments. Yang and Yao (2017) mentioned that an influential group of shareholders questioned current CSR reporting strategies and required the use of external CSR auditing firms. As a result, prominent CSR auditing firms and independent reporting tools such as (a) MSC KLD 400, (b) MSCI KLD, (c) Fortune Magazine Reputation Index, (d) Dow Jones Sustainability Index, (e) the Reputation Institute and (f) the Vigeo Index were introduced (see Kang & Liu, 2015).

### **3.2 The CSR Theory as a Planning Tool**

The concept of the CSR theory shifted the archaic value-maximization model and introduced a notion that focused on the development of stakeholders such as individuals and communities. Practitioners of the CSR theory not only presented a different approach on how to conduct businesses but also streamlined the gap among stakeholders. Business leaders tend to portray shareholders as the sole stakeholders to be satisfied, obviating that revenue emerges from selling products or rendering services to a group of individuals that ultimately form a community. When building business models or setting objectives and projections, business leaders need to evaluate strategies capable of inducing social change or at least provide benefits to their employees. Though business practitioners accepted the CSR theory, several scholars and practitioners, such as Kim and Woo (2019), maintained that CSR strategies are implemented as a reactive mechanism to alleviate social pressure, which then translates to spontaneous and costly reactive measures. Given today's high levels of (a) cost management activities, (b) business intelligence evaluations, and (c) advanced analytical tools, business leaders can weigh the financial burden of engaging in CSR

activities against the social cost of not doing so. While the CSR concept dates to the 1930s, its practice is viewed as a modern approach, an approach that business leaders can take to safeguard a role within their societies. It is evident that organizational culture and managerial expertise plays a vital role in the implementation of CSR strategies; however, Cantrell, Kyriazis, and Noble (2015) maintained that executives and business owners are required to lead adaptive organizations, adept enough to tackle the needs of their communities. Fontana (2018) sustained that practitioners of the CSR theory tend to understand that to maintain a good relationship with stakeholders, business leaders must create an organizational culture where stakeholders become familiar with the organization's mission, vision, and values. Though Frederick (1978) understood that an organization's financial situation limits its CSR engagement levels, Frederick (1986) maintained that the need for CSR strategies to be incorporated as an operational reality was needed. Organizational leaders experience or execute CSR in numerous ways (see Johnson, Ashoori, & Lee, 2018). Cheng, Ioannou, and Serafeim (2014) established that executives and business leaders strive to attain different degrees of CSR involvement; however, Frederick (1978) argued that business leaders should avoid the hasty reaction to society's wishes at any cost. Frynas and Yamahaki (2016) mentioned that organizational leaders are required to formulate CSR scanning mechanisms capable of detecting social adversities at an early stage. Fanti and Buccella (2017) and Kang et al. (2016) mentioned that Frederick offered a model that could be considered a static strategical tool and was not designed to evaluate social changes after the implementation of suitable CSR strategies. While understanding that Frederick set the conditions for today's CSR engagements, Babiak and Kihl (2018) emphasized that Frederick failed to provide an operationally-oriented model capable of measuring and delivering quantifiable CSR information to leaders at all levels. Bridoux and Stoelhorst (2016) concluded that Frederick introduced the CSR theory as a moral principle that would justify the attempts and efforts made by business leaders that decided to invest in social betterment; the reason, in this quantitative correlational study a clear and concise strategical model that offer the necessary tools to allow organizational leaders to measure CSR reporting levels at any given time was provided. The importance of treating CSR reporting as a planning mechanism needs to be a topic of constant debate among business leaders.

### 3.2.1 The Stakeholder Theory

Pérez and Rodríguez del Bosque (2016) presented arguments to emphasize that by introducing the stakeholder theory, Freeman (1984) (a) shaped the business world, (b) transformed the scholarly discussion of the stakeholder approach, and (c) established the foundation of today's business practice. Freeman suggested that the stakeholder terminology was first introduced during the 1960s; however, the stakeholder theory as a strategical approach emerged during the mid-1980s. Freeman (2010) indicated that the antiquated stakeholder approach limited the adaptation of organizational leaders, an action that prevented business leaders from reacting to the changes brought by the accelerated technological revolution. Brulhart, Gherra, and Quelin (2019) concluded that researchers in support of the newly introduced stakeholder approach embraced concepts such as ethics and business social responsibility. Miles (2017) argued that the inspiration behind the stakeholder approach was to build a framework capable of answering the concerns of organizational leaders that became stagnant due to the

unprecedented changes in their business environments. Bridoux and Stoelhorst (2016) established that business leaders that followed traditional strategical approaches were not capable of developing deliberate business models, nor could they comprehend the opportunities that emerged from the uninterrupted business environments and social transformations. As indicated by Freeman (1984, 2010), stakeholder approaches introduced before the 1980s were inconsistent due to the number of disconnected theoretical frameworks presented by previous scholars, and the reactive nature that characterized them. Freeman (1984) identified the need for a new conceptual framework, one capable of addressing the challenges faced by business leaders of the era. Freeman introduced the refined stakeholder approach, and with it, the intention of expanding the concept of strategic management. Freeman presented the term by defining stakeholders as any group or individuals that can be affected or can affect an organization's financial objective. Yang and Bentley (2017) suggested that Freeman tried to devise a strategic methodology that would allow business leaders to assess the needs of a myriad group of stakeholders and its relationship with organizational performance. Kim (2017) confirmed that although the technological revolution of the 1980s facilitated the introduction of the stakeholder approach as a strategical tool, the idea was not utterly innovative. However, Freeman (1984) was the first scholar to indicate that business leaders needed to comprehend the interests of not only shareholders but also (a) employees, (b) customers, (c) suppliers, (d) lenders, and (e) societies. Jones, Harrison, and Felps (2018) mentioned that by analyzing the needs of stakeholders such as (a) employees, (b) customers, (c) suppliers, (d) lenders, and (e) societies, business leaders could develop business objectives and strategies capable of earning the support of key stakeholders. Weitzner and Deutsch (2019) pointed out that the stakeholder approach earned the respect of business professionals and scholars after Freeman decided to include strategical frameworks such as (a) corporate planning, (b) the system theory, (c) the organizational theory, and (d) corporate social responsibility. Freeman's newly introduced list of stakeholders gave the power to internal and external groups that were not traditionally addressed by business leaders (see Freeman & Dmytriiev, 2017). Harrison et al. (2015) explained that the comprehensive view presented by Freeman (1984, 2010) seamlessly segregated stakeholders and assigned levels of importance to each group. Babiak and Kihl (2018) suggested that business strategies introduced before Freeman's broad definition of stakeholders offered business leaders a limited sense of corporate social responsibility. Guibert and Roloff (2017) pointed out that organizational leaders who understand the influence that groups of interest and entities have on their organization's overall performance are more likely to engage in productive, sustainable CSR practices. Brulhart et al. (2019) presented arguments to emphasize the view that organizational leaders need to demonstrate high levels of stakeholder awareness and include strategical approaches tailored to their organizational needs. Freeman and his stakeholder theory are targeted by constant criticism (see Miles, 2017). Freeman's attempt to balance the feedback received after the introduction of his stakeholder theory had created a sense of ambiguity. Weitzner and Deutsch (2019) mentioned that dividing the stakeholders into internal and external groups created a sense of difficulty. Yang and Bentley (2017) argued that stakeholders should be segmented by financial impact, not by an ethical or moral lens. Jones et al. (2018) maintained that not balancing the diverse stakeholders' interest, demonstrate a sense of ambiguity that can deviate organizational leaders from their

original organizational mission and vision. Weitzner and Deutsch (2019) maintained that Freeman covered the importance of balancing the stakeholder's interest but failed to deliver a comprehensive answer on how business leaders could accomplish so. Ranängen (2016) argued that failing to determine the legitimacy of groups of interest, restricted the usefulness of the stakeholder theory. Guibert and Roloff (2017) observed that while providing a framework to segregate stakeholders, Freeman did not provide the tools to balance the needs of stockholders and stakeholders. Authors such as Bridoux and Stoelhorst (2016) concluded that Freeman's stakeholder theory continues to be criticized for a simple reason, it does not cover the issues that emerge when organizational leaders deal with situations where conflict of interest among stakeholders exists. Scholars such as Ranängen and Miles suggested that Freeman and supporters of the stakeholder theory needed to refocus the theory and converted it into a theory of shared responsibility among business leaders and stakeholders. Guibert and Roloff explained that Freeman's stakeholder approach does not cover the innate liability shared between stakeholders and business leaders, representing a critical limitation to the stakeholder theory. Brulhart et al. argued that scholars tend to address the responsibilities held by business leaders and stakeholders; however, minimal academic and professional literature can be found to address the opposite. Babiak and Kihl defended the notion that a sense of (a) reciprocity, (b) interdependence, and (c) accountability is essential to satisfy the relationship between businesses and stakeholders. Babiak and Kihl defined (a) reciprocity as the shared responsibility between stakeholders and businesses, (b) interdependence as the collective consequences emerged from organizational and stakeholder actions, and (c) accountability as the ability to hold each other accountable for their actions. Leister and Maclachlan (2015) believed that adding stakeholder's responsibility to Freeman's stakeholder theory delivers a distinguishable structure when discussing CSR from a managerial standpoint. Hsu and Cheng (2016) explained how the dual responsibility methodology could motivate business leaders and stakeholders to collaborate and support CSR initiatives. Following a similar rationale, Cohen, Holder, and Khalil (2017) attempted to convince business leaders and stakeholders, that an increase in stakeholder contribution can translate to enhanced CSR performance. CSR performance in relation to the stakeholder theory, equates to social betterment and that the obligation of improving our societies is a combined effort. Business leaders and stakeholders are fundamental parts of society; the reason, enough energy, and capital should be allocated to the betterment of their communities when financially permissible. By introducing the stakeholder theory, Freeman exemplified the notion of stakeholder management (see Freeman and Dmytriiev, 2017). By embracing the stakeholder management approach, Freeman and Dmytriiev established that (a) organizational leaders need to enhance their ability to categorize and distinguishing stakeholders, (b) business leaders need to determine the organizational impact of each stakeholder and (c) that organizational leaders need to create, nurture, maintain and cherish the professional relationship with stakeholders. Pérez and Rodríguez del Bosque (2016) explained that when Freeman introduced the stakeholder theory, he intended to demonstrate that organizational leaders needed to address the interest of stakeholders at the time of making crucial business decisions. Freeman's strategical approach segregated the stakeholder theory from other managerial strategies, rendering a sense of uniqueness (see Agudo et al., 2015) while allowing business leaders to assign

a direction to their CSR practices (Marples, 2017). Freeman's stakeholder theory is often associated with Frederick's CSR theory and approach.

### 3.2.2 *The Stakeholder Theory and its Relationship with CSR*

The traditional understanding of stakeholders did not offer the tools needed to address the rapidly changing business world. Freeman (2010) offered his gratitude and commented that scholars of his epoch expanded the definition of stakeholders, and introduced a larger group of interest that organizational leaders needed to attend to survive in today's volatile business environment. As an acceptance gesture, Freeman extended the definition of stakeholder and pointed out that stakeholders are individuals and entities that can be affected by any organizational actions (see Freeman, 2010). Freeman argued that individuals and organizations that can be affected by any organizational actions could be grouped among the following stakeholders (a) employees, (b) customers, (c) suppliers, (d) creditors, (e) communities, (f) governmental institutions, (g) political groups, and (h) competitors. After additional scrutiny, and seeking to address the received initial scholarly criticism, Freeman (2010) introduced the term key stakeholders and mentioned that organizations should focus on (a) customers, (b) employees, (c) investors, (d) suppliers, (e) communities, and (f) governments. Freeman continued to improve his stakeholder approach and later recognized the second wave of stakeholder, the external stakeholders (see Freeman & Dmytriiev, 2017). Brown and William (2013) mentioned that after the introduction of external stakeholders such as, (a) consumer advocates, (b) preservationists, (c) groups of interests, and (d) nonprofit organizations, the same became relevant to business leaders. Freeman (2010) mentioned that the external stakeholders provided a new approach, and with it, the need for a new wave of organizational leaders capable of addressing their necessities. Though external stakeholders represent a crucial and pivotal element for today's business models, scholars such as Hategan et al. (2016) maintained that organizational leaders should solely focus on critical stakeholders because external stakeholders do not pose a threat to an organization's survival.

Scholars such as Jones et al. (2018) argued that supporters of the stakeholder theory tend to focus on value maximization, obviating any external factors or external stakeholders. Jones et al. also mentioned that the reality is that business leaders exist to provide shareholders with their expected return on investment. Harrison et al. (2015) argued that Freeman offered a framework that could be employed to analyze internal and external factors capable of inducing organizational change. Researchers and business professionals tend to post arguments to establish a clear connection between financial performance and proper stakeholder management. Business leaders must possess the aptitude to (a) segregate key stakeholders, (b) assess their needs, and (c) maintain an ethical relationship. Understanding that supporters of the stakeholder theory understood that Freeman's framework is an essential element of today's business environments, CSR scholars emphasized the need for attending not only the community but also the rest of stakeholders. Frynas and Yamahaki (2016) explained that the alignment of both theories reinforces the need for practical CSR strategies. Hetze (2016) argued that users of these strategies need to become capable of addressing not only stakeholders, such as employees but also their communities. Frynas and Yamahaki understood that satisfying a community while attending the social demands imposed by investors, and governmental

agencies can become a harsh task. Jones et al. (2018) mentioned that business leaders who combine the stakeholder and the CSR theories become capable of evaluating key stakeholders and developing CSR strategies tailored to their necessities. Michelon et al. (2015) maintained that organizational leaders tend to face difficulties at the time of quantifying CSR implementation, or at the time of developing operationally compatible CSR strategies. Frynas and Yamahaki argued that by utilizing Freeman's stakeholder theory as a framework, business leaders could become capable of (a) assessing CSR performance, (b) gain stakeholder acceptance, and (c) increase the stakeholder's willingness to maintain the proposed CSR initiatives. Agudo et al. (2015) mentioned that contemporary theorists tend to relate to Frederick's work with Freeman's stakeholder theory. Freeman (2010) argued that business leaders could become capable of determining and addressing the necessities presented by stakeholders but struggle to segment their communities and discuss their current necessities. Freeman conveyed that communities play a vital role in an organization's lifespan and should not be precluded. Harrison et al. (2015) maintained that researchers could easily argue that the stakeholder theory is an integral aspect of the CSR theory. Jones et al. (2018) emphasized that supporters of the stakeholder theory tend to establish the need for developing business models that encompass the betterment of their communities, and the strong relationship between business leaders and stakeholders. Researchers such as Zhang and Zhang (2016) endorsed a more modern approach and advocated the view that CSR performance is associated with an organizational leaders' ability to identify groups of interest, as well as their proficiency at the time of developing appropriate CSR programs. Although trying to pinpoint the etymology of the word CSR can be a difficult task, Frynas and Yamahaki stated that aligning Freeman's stakeholder theory with Frederick's CSR approach facilitated their research process. When analyzing a significant amount of peer-reviewed articles and books, attempting to assign a chronological line to CSR became a tedious assignment. Frederick's CSR theory and Freeman's stakeholder approach intertwine and can one day become a single methodology; however, this ideology cannot become a reality until scholars develop a CSR-quantifiable model.

### 3.3 The Development of a Sustainable CSR Program

CSR actions are the organizational efforts employed by business leaders capable of inducing social change. Business professionals that tend to embrace CSR are recognized as individuals seeking social betterment. However, the normative that CSR initiatives can positively benefit the financial performance of a company is a topic of constant debate among scholars and business professionals. Understanding that the employment of CSR is driven by significant economic developments such as globalization and sustainability, Kang et al. presented arguments to emphasize that organizational leaders must evaluate CSR at the time of developing or refining today's complex business models. Cheng et al. (2014) pointed out that CSR initiatives are an underestimated planning tool, and that it requires the same attention organizational leaders give to crucial strategic elements such as (a) price, (b) quality, and (c) demand. It is evident that transforming CSR from a reactive tool to a planning mechanism is essential to develop the CSR field even further (see Zhang & Zhang, 2016). Brown and William (2013) presented arguments to advocate the view that scholars and CRS-practitioners need to present empiric data capable of

demonstrating that CSR strategies can transcend and become more than an instrument applied to cushion scandals involving (a) labor abuse, (b) international labor rights, and (c) environmental exploitation. CSR consultants can provide an outside perspective of the problems experienced by the communities under scrutiny, allowing business leaders to determine a suitable organizational approach. Hsu and Cheng (2016) pointed out that business leaders attempting to develop valuable CSR programs need to assess the available CSR literature and seek for pertinent external validation. Smith and Colvin (2016) proposed that today's business practices call for not only environmentally sustainable approaches but also a business environment free of unethical behaviors. Organizational leaders belonging to the same industry tend to experience similar unfavorable circumstances. Harrison et al. (2015) understood that after analyzing the common ethical trends, and by embracing CSR strategies as a preventive mechanism, business leaders could mitigate the negative consequences associated with a previously assessed crisis. Smith and Colvin recognized that additional occurrences could pose a different set of difficulties, for which CSR preventing tools can serve as the cushion business leaders need at the time of tackling a crisis. Greater volume of academic literature would one day allow the CSR field to transcend and become a widely-renown and respected strategic instrument among business leaders. The managerial implementation of CSR strategies will one day become a requirement among business practitioners at all organizational levels.

## 4. METHODOLOGY

Bernard and Bernard (2017) mentioned that the formulation of a research question and hypothesis typically guides the selection of a research methodology. For the purpose of this study, we selected the quantitative method. The research question: "What is the relationship between CSR reporting, CSR<sub>i</sub>, and the financial performance of hardware and software organizations listed in the United States?" was designed to ignite curiosity among scholars and business practitioners that understand that a possible association between CSR reporting and financial performance exists. The independent variables were CSR reporting and CSR<sub>i</sub>. The dependent variables were NI and ROA. To confirm or deny the relationship between the independent and dependent variables, we (a) exported the sample's financial information, (b) CSR reporting values, and (c) CSR<sub>i</sub> information from Microsoft Excel to IBM SPSS version 26 and conducted a complete correlational analysis. Thanks to IBM SPSS user-friendly analytical platform, we were able to evaluate the following hypotheses:

H1<sub>0</sub>: There is no significant relationship between CSR reporting, CSR<sub>i</sub>, and the NI of hardware and software organizations listed in the United States.

H1<sub>1</sub>: There is a statistically significant relationship between CSR reporting, and the NI of hardware and software organizations listed in the United States.

H2<sub>0</sub>: There is no significant relationship between CSR reporting, CSR<sub>i</sub>, and the ROA of hardware and software organizations listed in the United States.

H2<sub>1</sub>: There is a statistically significant relationship between CSR reporting, CSR<sub>i</sub>, and the ROA of hardware and software organizations listed in the United States.

The correlation values are interpreted as such (a)  $r^2 < .34$  is weak (b)  $r^2$  between  $.35$  and  $.60$  is significant, and (c)  $r^2 > .61$

is strong. The information surrounding the dependent and independent variables were exported from Microsoft Excel to IBM SPSS. We organized the financial information surrounding the dependent variables NI, and ROA and the independent variables CSR reporting and CSRI into four different columns. The vertical placement of the variables facilitated the employment of IBM SPSS version 26 to test the assumptions of (a) multicollinearity, (b) outliers, (c) normality, (d) linearity, (e) homoscedasticity, and (f) independence of residuals. To quantify the data associated with the independent variable CSR reporting, we first used IBM SPSS Modeler Text Analytics version 18.0 to code the sample's annual reports as issued to the SEC between the years 2010-2015. After coding the annual reports, the results were augmented by a five-point Likert scale. Through the coding process, we attempted to locate criteria such as (a) environment (b) sustainability, (c) social responsibility, (d) employee development, (e) safe labor practices, and (f) charity. Besides the internet, Bloomfield et al. (2016) recommended the use of (a) physical and web-based libraries, (b) surveys, (c) observation logs, and (d) the employment of formal interviews. The sample's financial information was retrieved from (a) the sample's official web pages, (b) the SEC official web page, (c) macro trends.net, (d) NASDAQ.com, and (e) reuters.com. The dependent variable ROA was re-calculated by appraising the sample's total assets between the years 2010-2015. This information was retrieved from the same web pages. The independent variable CSRI was extracted from the sample's consolidated Global CSR RepTrak scores as provided by the Reputation Institute. The Reputation Institute evaluates seven general dimensions (Fombrun, 2007). The Reputation Institute evaluates dimensions such as (a) products and services, (b) innovation, (c) workplace, (d) governance, (e) citizenship, (f) leadership, and (g) financial performance. The dimensions constitute a score that ranges from 0-100. The sample's RepTrak scores for the years 2010-2015 were donated by the Reputation Institute to further enhance current CSR literature.

## 5. SUMMARY OF THE FINDINGS

The initial intention was to utilize the entirety of the software and hardware organizations that were part of Fortune 500 between the years 2010-2015 as the population sample. After applying industry filters such as (a) computer office equipment, (b) computer peripherals, (c) computer software, (d) information technology services, (e) electronic and office equipment, (f) network, and other communication equipment, (g) semiconductors, and (h) other electronic components to the Fortune 500 lists between the years 2010-2015, a total of 51 organizations emerged. Out of the 51 hardware and software organizations, (a) three were acquired by larger corporations between the 2010-2015 period, (b) 13 were discarded due to a lack of CSRI information, and (c) 10 were identified as critical outliers. The sample was reduced to 25. The critical outliers emerged from the organizations that showed significantly large NIs. After removing the NIs identified as outliers, the distribution of the data showed a relatively normal distribution, one that presented a skewness and kurtosis Z-value that was not less than -1.96 or greater than 1.96. Organizations with a ROA of 5% or higher are considered great investment opportunities. The statistical mean of the data surrounding ROA was 4.3%, proving that the gathered information was suitable for this study. CSR reporting displayed a statistical mean of 1.4, revealing low

CSR reporting values among the sample. CSRI, on the other hand, showed a moderate level of CSR by displaying a statistical mean of 67.4.

## Inferential Results

For the purpose of this quantitative correlational study, we used a multiple linear regression model to examine the relationship between CSR reporting, CSRI, and the financial performance of 25 hardware and software organizations. The multiple linear regression analysis is typically used to determine the relationship between two independent variables and one dependent variable. The independent variables were CSR reporting and CSRI. The dependent variables were NI and ROA. The null hypothesis was that CSR reporting and CSRI were not significantly correlated with the financial performance of 25 hardware and software organizations. Following the recommendations of Poole and O'Farrell (1971), the occurrence of Type I and Type II errors were mitigated by testing the assumptions of (a) multicollinearity, (b) outliers, (c) normality, (d) linearity, (e) homoscedasticity, and (f) independence of residuals; no significant violations were encountered. A multiple linear regression model was used to statistically assess the independent variables CSR reporting, and CSRI against the dependent variable NI. The model showed a positive but weak relationship between the independent variables CSR reporting, CSRI and the dependent variable NI,  $F(2, 22) = .389, p = .682, r^2 = .03$ . Since  $p > .05$ , the null hypothesis stating that there is no significant relationship between CSR reporting, CSRI, and NI was accepted. The  $r^2$  value of .034 demonstrates that when measured against CSR reporting and CSRI, NI has no significant relationship. Utilizing the same multiple linear regression model, we attempted to obtain a significance level of  $p \leq .05$  or an  $r^2 > .35$  to confirm or deny the relationship between the independent variables CSR reporting, CSRI, and the dependent variable ROA. The model showed a positive but weak relationship between CSR reporting, CSRI, and ROA,  $F(2, 22) = .552, p = .584, r^2 = .048$ . Since  $p > .05$ , the null hypothesis stating that there is no significant relationship between CSR reporting, CSRI, and ROA was accepted. The  $r^2$  value of .034 demonstrated that when measured against CSR reporting and CSRI, ROA has no significant relationship.

## Analysis Summary

The purpose of this quantitative correlational study was to confirm or deny the relationship between CSR reporting and financial performance. A multiple linear regression model demonstrated that implementing higher levels of CSR reporting does not necessarily equate to an increase or decrease in financial performance. To verify the validity of the chosen statistical model, we decided to test the assumptions of (a) multicollinearity, (b) outliers, (c) normality, (d) linearity, (e) homoscedasticity, and (f) independence of residuals. No major violations were encountered. The model showed that there was no significant relationship between the dependent variable NI and the independent variables CSR reporting, and CSRI,  $F(2, 22) =$

.389,  $p = .682$ ,  $r^2 = .034$ . The model also showed that the dependent variable ROA is not significantly correlated with the independent variables CSR reporting, and CSRI,  $F(2, 22) = .552$ ,  $p = .584$ ,  $r^2 = .048$ .

### Theoretical Conversation on Findings

Freeman's (1984) stakeholder and Frederick's (1978) CSR theories served as the theoretical framework for this study. As applied to this study, the stakeholder theory suggested that the independent variables CSR reporting, and CSRI drive financial performance. As applied to this study, the CSR theory suggested that a strong correlation between CSR reporting and financial performance existed. The statistical models did not support the notion that higher levels of CSR reporting translated to enhanced financial performance. Kang and Liu (2015) conducted a similar study and encountered an  $r^2 = .157$ . Kang and Liu used the ratings provided by the Dow Jones Sustainability Group index and measured it against the dependent variables ROA and return on equity. Abernathy et al. (2017) recommended the use of CSR rating entities such as the KLD 400 as a key independent variable; however, Hetze (2016) utilized the KLD 400 ratings and proved that a negative relationship between CSR and financial performance existed ( $r^2 = -.001$ ). The constant apparitions of weak but positive relationships between CSR and financial performance further enhance Brown and William's (2013) argument that organizational leaders and investors need to visualize CSR as a responsible act, not as an income-driven strategy.

## 6. CONCLUSION

Though limited empiric data exist to determine the impact CSR initiatives have on small- and medium-sized organizations, it is evident that shareholders continue to demand more CSR participation. The professional and academic literature that we used to shape this study helped to contribute to current CSR literature. We used two dependent variables to ignite critical thinking among professionals and scholars. Based on the encountered positive but insignificant relationship, it is safe to assume that if CSR-enthusiasts decide to evaluate additional variables to create a different statistical model, the model could reveal a higher level of correlation. Understanding that organizational actions need to be justified, in this study, we aimed to confirm or deny the relationship between CSR reporting, CSRI, and financial performance by employing a multiple linear regression model. The statistical models were designed to convey quantifiable data and to satisfy the value-maximization needs of the targeted audience. Small- and medium-sized organizational leaders could replicate this study by measuring CSR variables such as customer and employee satisfaction against dependent variables such as NI, and market share. Regardless of organizational size, NI, or social reputation, social responsibility is a collective obligation that requires a shared contribution.

## 7. REFERENCES

- [1] Frynas, J. G., & Yamahaki, C. (2016). Corporate social responsibility: Review and roadmap of theoretical perspectives. *Business Ethics: A European Review*, 25(3), 258-285. doi:10.1111/beer.12115
- [2] Marples, R. (2017). Art, knowledge and moral understanding. *Ethics and Education*, 12(2), 243-258. doi:10.1080/17449642.2017.132342
- [3] Svantesson, M., Silén, M., & James, I. (2017). It's not all about moral reasoning: Understanding the content of moral case deliberation. *Nursing Ethics*, 25(2), 212-229. doi:10.1177/0969733017700235
- [4] Agudo, V. J. M., Garcés, A. C., & Salvador, F. M. (2015). Corporate social performance and stakeholder dialogue management. *Corporate Social Responsibility & Environmental Management*, 22(1), 13-31. doi:10.1002/csr.1324
- [5] Michelon, G., Pilonato, S., & Ricceri, F. (2015). CSR reporting practices and the quality of disclosure: An empirical analysis. *Critical Perspectives on Accounting*, 33(3), 59-78. doi:10.1016/j.cpa.2014.10.003
- [6] Babiak, K., & Kihl, L. A. (2018). A case study of stakeholder dialogue in professional sport: An example of CSR engagement. *Business & Society Review*, 123(1), 119-149. doi:10.1111/basr.12137
- [7] Fontana, E. (2018). Corporate social responsibility as stakeholder engagement: Firm-NGO collaboration in Sweden. *Corporate Social Responsibility & Environmental Management*, 25(4), 327-338. doi:10.1002/csr.1463
- [8] Brown, J. A., & William, F. R. (2013). CSR and stakeholder theory: A tale of Adam Smith. *Journal of Business Ethics*, 112(2), 301-312. doi:10.1007/s10551-012-1251-4
- [9] Frederick, W. C. (1978). From CSR1 to CSR2: The maturing of business-and-society thought. *Business & Society*, 33(1), 150-164. doi:10.1177/000765039403300202
- [10] Frederick, W. C. (1986). Toward CSR3: Why ethical analysis is indispensable and unavoidable in corporate affairs. *California Management Review*, 28(2), 126-142. doi:10.2307/41165190
- [11] Abernathy, J., Stefaniak, C., Wilkins, A., & Olson, J. (2017). Literature review and research opportunities on credibility of corporate social responsibility reporting. *American Journal of Business*, 32(1), 24-41. doi:10.1108/AJB-04-2016-0013
- [12] Brown, H., & Zmora, V. (2015). The role of corporate social responsibility (CSR) assurance in investors' judgments when managerial pay is explicitly tied to CSR performance. *Auditing: A Journal of Practice & Theory*, 34(1), 75-96. doi:10.2308/ajpt-50813
- [13] Freeman, R. E., & Dmytriiev, S. (2017). Corporate social responsibility and stakeholder theory: Learning from each other. *Emerging Issues in Management*, 12(4), 7-15. doi:10.4468/2017.1.02
- [14] Kang, H., & Liu, S. (2015). Corporate social responsibility and corporate performance: A quantile regression approach. *Quality and Quantity*, 48(6), 311-325. doi:10.1007/s11135-013.9958-6

- [15] Porter, M., & Kramer, M. (2007). Strategy and society: The link between competitive advantage and corporate social responsibility. *Strategic Direction*, 23(5), 78-93. doi:10.1108/sd.2007.05623ead.006
- [16] Ranängen, H. (2016). Stakeholder management theory meets CSR practice in Swedish mining. *Mineral Economics*, 30(1), 15-29. doi:10.1007/s13563-016-0098-z
- [17] Harrison, J. S., Freeman, R. E., & Sá de Abreu, M. C. (2015). Stakeholder theory as an ethical approach to effective management: Applying the theory to multiple contexts. *Revista Brasileira de Gestão de Negócios*, 17, 858-859. doi:10.7819/rbgn.v17i55.2647
- [18] Yu, Y., & Choi, Y. (2016). Stakeholder pressure and CSR adoption: The mediating role of organizational culture for Chinese companies. *Social Science Journal*, 53(2), 226-235. doi:10.1016/j.soscij.2014.07.006
- [19] Hategan, C. D., Sirghi, N., Curea-Pitorac, R. I., & Hategan, V. P. (2016). Doing well or doing good: The relationship between corporate social responsibility and profit in Romanian companies. *Sustainability*, 10(4), 1-27. doi:10.3390/su10041041
- [20] Kang, C., Germann, F., & Grewal, R. (2016). Washing away your sins: Corporate social responsibility, and firm performance. *Journal of Marketing*, 80(2), 59-79. doi:10.1509/jm.150324
- [21] Hsu, J. L., & Cheng, M. (2016). What prompts small and medium enterprises to engage in corporate social responsibility. *Environmental Management*, 19(5), 288-305. doi:10.3215/151321-4
- [22] Smith, K. J., & Colvin, L. (2016). Creating and implementing codes of ethical conduct. *Employment Relations Today*, 43(1), 95-101. doi:10.1002/ert.21555
- [23] Leister, D. V., & Maclachlan, D. L. (2015). Organizational self-perception and environmental image measurement. *Academy of Management Journal*, 18(2), 205-223. doi:10.2307/25552
- [24] Yang, C., & Yao, L. (2017). Testing ambiguity theories with a mean-preserving design. *Quantitative Economics*, 8(1), 219-238. doi:10.3982/qe460
- [25] Pérez, A., & Rodríguez del Bosque, I. (2016). The stakeholder management theory of CSR. *International Journal of Bank Marketing*, 34(5), 731-751. doi:10.1108/IJBM-04-2015-0052
- [26] Kim, Y., & Woo, C. W. (2019). The buffering effects of CSR reputation in times of product-harm crisis. *Corporate Communications: An International Journal*, 24(1), 21-43. doi:10.1108/ccij-02-2018-0024
- [27] Cantrell, J., Kyriazis, E., & Noble, G. (2015). Developing CSR giving as a dynamic capability for salient stakeholder management. *Journal of Business Ethics*, 130(2), 403-421. doi:10.1007/s10551-014-2229-1
- [28] Johnson, Z., Ashoori, M. T., & Lee, Y. J. (2018). Self-reporting CSR activities: When your company harms, do you self-disclose? *Corporate Reputation Review*, 21(4), 153-164. doi:10.1057/s41299-018-0051-x
- [29] Cheng, B., Ioannou, I., & Serafeim, G. (2014). Corporate social responsibility and access to finance. *Strategic Management Journal*, 35(1), 1-23. doi:10.1002/smj.2131
- [30] Fanti, L., & Buccella, D. (2017). Profit raising entry effects in network industries with corporate social responsibility. *Economics and Business Letters*, 6(3), 59-67. doi:10.17811/ebl.6.3.2017.59-68
- [31] Bridoux, F., & Stoelhorst, J. W. (2016). Stakeholder relationships and social welfare: A behavioral theory of contributions to joint value creation. *Academy of Management Review*, 41(2), 229-251. doi:10.5465/amr.2013.0475
- [32] Freeman, R. E. (1984). *Strategic management: A stakeholder approach*. Boston: Pitman.
- [33] Freeman, R. E. (2010). *Strategic management: A stakeholder approach*. New York, NY: Cambridge University Press.
- [34] Brulhart, F., Gherra, S., & Quelin, B. V. (2019). Do stakeholder orientation and environmental proactivity impact firm profitability? *Journal of Business Ethics*, 158(1), 25-46. doi:10.1007/s10551-017-3732-y
- [35] Miles, S. (2017). Stakeholder theory classification: A theoretical and empirical evaluation of definitions. *Journal of Business Ethics*, 142(3), 437-459. doi:10.1007/s10551-015-2741-y
- [36] Yang, A., & Bentley, J. (2017). A balance theory approach to stakeholder network and apology strategy. *Public Relations Review*, 43(2), 267-277. doi:10.1016/j.pubrev.2017.02.012
- [37] Kim, Y. (2017). A qualitative research on the organizing and institutionalizing processes of social venture entrepreneurship. *The Korea Entrepreneurship Society*, 12(5), 19-43. doi:10.24878/kes.2017.12.5.019
- [38] Jones, T. M., Harrison, J. S., & Felps, W. (2018). How applying the instrumental stakeholder theory can provide sustainable competitive advantage. *Academy of Management Review*, 43(3), 371-391. doi:10.5465/amr.2016.0111
- [39] Weitzner, D., & Deutsch, Y. (2019). Why the time has come to retire instrumental stakeholder theory. *Academy of Management Review*, 44(3), 694-698. doi:10.5465/amr.2018.0342
- [40] Guibert, L., & Roloff, J. (2017). Stakeholder dialogue: strategic tool or wasted words? *Journal of Business Strategy*, 38(5), 3-11. doi:10.1108/JBS-07-2016-0071
- [41] Cohen, J., Holder, L., & Khalil, S. (2017). A further examination of the impact of corporate social responsibility and governance on investment decisions. *Journal of Business Ethics*, 146(1), 203-218. doi:10.1007/s10551-015-2933-5
- [42] Hetze, K. (2016). Effects on the (CSR) reputation: CSR reporting discussed in the light of signaling and stakeholder perception theories. *Corporate Reputation Review*, 19(3), 281-296. doi:10.1057/s41299-016-0002-3
- [43] Zhang, H., & Zhang, M. (2016). The corporate social entrepreneur: From concept to practice. *Global Business & Organizational Excellence*, 35(2), 50-59. doi:10.1002/joe.21655
- [44] Fombrun, C. J. (2007). List of lists: A compilation of international corporate reputation ratings. *Corporate Reputation Review*, 10(2), 144-153. doi:10.1057/palgrave.crr.1550047
- [45] Poole, M. A., & O'Farrell, P. N. (1971). The assumptions of the linear regression model. *Transactions*

*of the Institute of British Geographers*, 145-158.  
doi:10.2307/621706

# Feature Extraction Phase for Offline Arabic Handwritten Character Recognition

Dr. Rawia I. O. Ahmed  
Department of Computer Science  
College of Community Female  
University of Ha'il  
Saudi Arabia

Dr. Mohamed E. M. Musa  
College of Computer Science and Information  
Technology Sudan  
University of Science and Technology Khartoum,  
Sudan

**Abstract:** —In this paper we reviewed the importance issues of the optical character recognition, gives more emphases for OCR and its phases. We discuss the main characteristics of Arabic language, furthermore it focused on Feature Extraction phase of the character recognition system. We described and implemented the algorithm of Feature Extraction based on Freeman Chain codes. The algorithms are tested using 47,988 isolated character sample taken from SUST/ ALT dataset and achieved better results. The Feature Extraction phase developed by using MATLAB software.

**Keywords:** optical character recognition; offline recognition; online recognition; handwritten, preprocessing; Feature Extraction: Freeman Chain codes.

## 1. INTRODUCTION

Over the past three decades, many studies have been concerned with the recognition of Arabic words. Offline handwritten Arabic characters' recognition have received more attention in these studies, because of the need to Arabic document digitalization.

In this paper, Feature Extraction system for an isolated Arabic handwritten are design and tested by using SUST/ ALT dataset. it's a new dataset developed and published by SUST/ALT (Sudan University of Science and Technology-Arabic Language Technology group) group. It contains numerals datasets, isolated Arabic character datasets and Arabic names datasets[1]. 40 common Arabic (especially in Sudan) males and females' name[2]. Each form written by one writer resulting 40,000 sample. it used for researching purpose.

The rest of the paper is organized as follows: Then the main characteristics of Arabic language are discussed in Section 2. Then the concepts and phases of OCR system are described in Section 3. The proposed Feature Extraction phase based on Freeman Chain codes discussed in Section 4. conclusion and future work are presented in Section 5.

## 2. The MAIN CHARACTERISTICS OF ARABIC LANGUAGE

Many studies have been conducted on recognition of Chinese, Japanese and Latin languages, but few were done on Arabic handwritten recognition[3]. One of the main reasons for this is that characteristics of Arabic language do not allow direct implementation of many algorithms used in other languages. The characteristics of Arabic language can be summarized as follows:

- Arabic language is represented in 28 characters and appears in different four shapes isolated, initial, medium or final.
- Arabic language is written from right to left, rather than from left to right this is useful for human

reader rather than for the computer. As seen in Figure.1

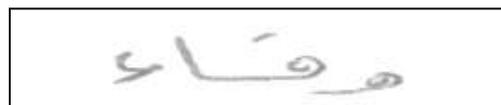


Figure.1 Arabic Word Written from Right to Left

- Arabic characters of a word are connected a long baseline, and character position above and below the baseline.
- Some Arabic character have the same shape and differ in the number of dots by which it will be identified, for example characters ث, ت, ب have the same shape but differ in number of dots, one dot in character Baa, two dots in character Taa, and three dots in character Thaa.
- Some Arabic character have the same shape and differ in the position of dots by which it will be identified, for example characters ن, ب the two characters have the same shape and identify with one dot, but they differ in position of dot one is above the baseline (character Noon), and other under the base line (character Baa), this differentiation can change the meaning of a word.
- The width and high of Arabic characters are differ from one character to another.
- The shape of Arabic character varies per writer.
- Arabic writing is cursive, most of Arabic characters are connected from two sides; right and left, only six characters are connected from right side only, as shown in Figure.2.



Figure.2 Arabic characters which can be connected from right to left

- Moreover, Arabic language has some diacritics called Tashkeel. The names of these Tashkeel: Fatha, Dhamma, Kasra, Sukun, Shadda, Fathatain, Kasratain, Dhammatain also combination of them are possible. These diacritics may change the meaning of specific word, for example: when we put Fatha diacritic on the word “حر” it became “حَر” which meaning “hot weather”, when we put dhamma diacritics on the same word, it became “حُر” which meaning “free”.
- Some Arabic words consists of more than one sub-words. A sub-word is the basic standalone pictorial block of the Arabic writing [4]. A brief details of Arabic handwritten characteristic were reviewed by Lorigo [5].

### 3. THE OPTICAL CHARACTER RECOGNITION SYSTEM

The Optical Character Recognition (OCR) is one of important tasks in computer area. It has many definitions, OCR defined as a process that attempts to turn a paper document into a fully editable form, which can be used in word processing and other applications as if it had been typed through the keyboard[3]. Also OCR was defined by Srihari et al. as the task of transforming text represented in the special form of graphical marks into its symbolic representation[6].

The recognition of handwritten can be applied in many areas such as names of persons, companies, organizations, newspapers, letters, archiving and retrieving texts, proteins and genes in the molecular biology context, journals, books, bank cheques, personal signatures and digital recognition, etc.[7]. A recognition system can be either online or offline[8]. It is online if the data being captured during the writing process. It always captured by special pen on an electronic interface. Online recognition has several interesting characteristics: firstly, recognition is performed on one dimensional rather than two dimensional images, secondly, the writing line is represented by a sequence of dots which its location is a function of time[3]. A recognition system is offline if its data scanned by scanner after writing process is over, such as any images scanned in by a scanner. In this case, only the image of the handwriting is available.

When we compared online handwriting recognition systems with offline systems, we found that offline systems are considered more difficult than online systems. This difficulty due to several reasons, out of which online handwriting recognition depends on temporal information, which facilitate the recognition system, but the temporal information is lacked in offline handwriting, it depends on passive images stored in files. This lemma makes offline systems less accurate than online systems. Furthermore, offline systems are more complex than online systems, because they depend on human writing which had more feature and characteristic specially for Arabic language.

There is no great variation exists between phases of the online and offline handwriting recognition systems.

the general phases of OCR systems are: data capture, preprocessing, segmentation, feature extraction, classification and, post processing as shown in Figure .3

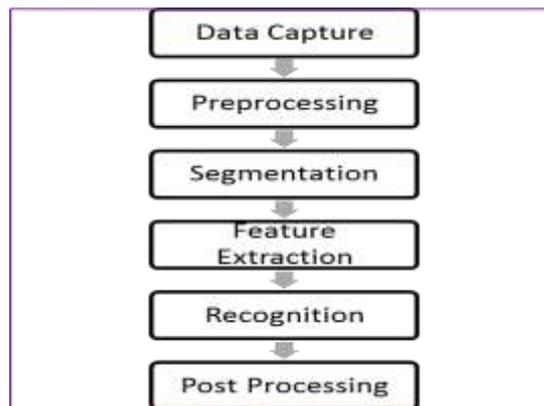


Figure.3 General recognition system phases

### 4. THE PROPOSED FEATURES EXTRACTION PHASE

Prior to the features extraction phase preprocessing phase must be done. Pre-processing of the handwritten character image is an important factor, to simplify the task of recognition. Usually several operations can be performed in this phase. Since in SUST isolated characters dataset, some preprocessing method are done during the development stage[1], minimal number of preprocessing processes are used in this work. An image file of isolated handwritten character will first be introduced to the system as gray scale bmp image. Then obtained images are binarized to be in digital form. When the study focus on characters' body only, dots is removed from some characters. Thinning is very important process in OCR, therefore we applied it the binary images. The next sub sections give a brief detail of these phases were discussed in our previous work [9].

The next step is to extract the useful features that will have used in classification phase. Many researchers agree that feature extraction phase play an important role in a handwriting recognition systems.

A human being can differentiate between various objects by observing their colors, shapes and attributes. To simulate this intelligence idea into a computer system, we need to implement geometrical and topological representation methods to help the system in recognizing the shapes of objects.

After studying several features methods, we found that geometrical and topological representation methods are reverent methods to recognize the isolated handwritten character.

To represent each image in a feature vector form, a mathematical model with a finite number of parameters is required. But unfortunately, they are no reasonable mathematical model currently exists.

In this section, we describe an efficient method for extracting features from handwritten Arabic character body using freeman chain code

#### 4.1 Freeman Chain Code Algorithm

FCC can be 4-connectivity or 8-connectivity. It traced the boundary of an image in a clockwise or anticlockwise directions. The main weakness of 4-connectivity is that we be unable to find the transverse points[10]. These points are very valuable in image recognition. So, to overcome the weakness of 4-connectivity we use 8-connectivity FCC. In

8-connectivity each code can be considered as the angular directions (8 directions).

In our experimentation, we labeled the direction from 0 to 7 and decide to trace the image in a clockwise direction, it is commonly in many research and suitable for Arabic characters, because it written from right to left [11-13].

The big challenge of FCC is how to find the starting pixel and the directions of image traverses. We have produced different chain code for the same character, if we start with different starting points or traverse on different directions. Therefore, consistency plays an important factor to overcome the variations of the chain code for the same character and preserve the success of the algorithm.

As mention before, we obtained the boundary of the character image by traverse the FCC in a clockwise direction. We assigned numbers 0 to 7 to each direction.

The character images are binarized to 1's and 0's pixels in preprocessing phase. Instead of storing the absolute location of each 1 pixel, we stored it is direction from its previously coded neighbor. A neighbor is any of the adjacent pixels in the 3x 3 pixels' neighborhood surround that current pixel as shown in Figure .4

The literature of Freeman chain code was introduced in [14]. To define the starting point, we move from the top of character body image to the bottom raw by raw to find first nonzero pixel. Furthermore, we assume that this first pixel has one neighbor. When we find the first nonzero pixel we defined it as starting point of the chain code and stored it in chain code list. In some cases, the first nonzero pixel has two neighbors. This case holds when the character body written as loop, for example, see the “Haa” character in Figure in Figure .4, If this case arises, we assume the starting chain code is zero. After finding the start point of the chain code we traverse to the next neighbor pixel in the image of character body.



Figure.4 “Haa” Character Written as Loop

In fact, there must be at least one neighbor boundary pixel at one of the eight locations surrounding the current starting boundary pixel (note each location has one value from 0 to 7 marked per the chain code direction). Again, the starting point can have more than one neighbor. In this case the chain code direction plays and important factor, to determine which neighbor will be chosen. In our experimentation, we chose a clockwise direction. When the neighbor found, we stored it in chain code list. Then finding the next neighbor.

The process of finding the next neighbor continuous until we reached to the starting point. This algorithm followed to find the FCC are summarized in Figure 5.

```

Input: Thinning binary character body image.
Output: 8-connective Freeman chain code
While there are still images to be traversed
Begin
    move from the top of character body image to the bottom, raw
    by raw to find first nonzero pixel
If one nonzero pixel then
Begin
    Assume that it is a starting pixel
        Stored it is direction in chain code list
    End
    Else
        Stored in chain code list 0 value as starting pixel
    End
    From starting pixel to end pixel do
        Assign 0-7 values to the eight directions
        Travels the neighbor pixel in clockwise direction
        Find and store the direction code of the neighbor pixel in
        chain code list
        Move to next position
    End
End
    
```

Figure.5 The Algorithm for Generating

8- Connective Freeman Chain Code

4.2 Normalized Freeman Chain Code

In this section, the feature vector for each character body image is obtained by applying 8-connective chain code. Then the two-dimensional (2D) matrix is converting to one-dimensional (1D) contain 10 digits by applying the proposed normalized chain code algorithm[14]. Samples of results obtained from implementation these algorithms on character "ف" samples are shown in Table 1.

Table .1 The Results Obtained from Implementation of Normalized Chain Code on Character "ف"

No	Original Image	Image After Preprocessing Operations	Normalized Chain Code
1			5 5 6 6 7 0 0 1 2 3
2			5 5 6 7 7 0 0 1 2 3
3			5 5 6 7 1 1 0 2 3 4
4			5 6 7 0 0 1 2 3 4 4

The proposed algorithm worked well and gives 99.23% result, this makes the algorithm more effective only about 193 of the samples their chain code be zero, this due to unconnected properties of these samples.

In making comparisons with the existing work, it is difficult to compare with work in[17] since those authors used other dataset and they have chosen 200 images so the proposed algorithm cannot be implemented on the same data.

## 5. CONCLUSION AND FUTURE WORK

In this paper, we present a review about optical character recognition system and its phases. Also, we list the characteristics of Arabic language, and focused in one of important phases in recognition systems which is feature extraction. Moreover, we described and implemented feature extraction algorithms based on free man chain code on isolated Arabian characters. In the future, we will have implemented this algorithm to Arabic word handwritten.

## 6. REFERENCES

- [1] Musa, M.E. Arabic handwritten datasets for pattern recognition and machine learning. in 2011 5th International Conference on Application of Information and Communication Technologies (AICT).
- [2] Wahby, T.M., I.M. Osman, and M.E. Musa, On Finding the Best Number of States for a HMM-Based Offline Arabic Word Recognition System. 2011.
- [3] Mori, S., H. Nishida, and H. Yamada, Optical character recognition. 1999: John Wiley & Sons, Inc.
- [4] Abed, M.A., Freeman chain code contour processing for handwritten isolated Arabic characters recognition. Alyrmook University Magazine,Baghdad, 2012.
- [5] Lorigo, L.M. and V. Govindaraju, Offline Arabic handwriting recognition: a survey. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2006.
- [6] Srihari, S.N., A. Shekhawat, and S.W. Lam, *Optical character recognition (OCR)*. 2003.
- [7] Amin, A., Off-line Arabic character recognition: the state of the art. Pattern recognition, 1998.
- [8] Khorsheed, M.S., Off-line Arabic character recognition–a review. Pattern analysis & applications, 2002.
- [9] Rawia, Ahmed; Mohammed, Musa, Preprocessing Phase for Offline Arabic Handwritten Character Recognition, International Journal of Science and Research (IJSR), ISSN (Online): 2319-7064 Volume 5 Issue 11 ,2016.
- [10] Yang, M., K. Kpalma, and J. Ronsin, A survey of shape feature extraction techniques. Pattern recognition, 2008.
- [11] Sampath, A., C. Tripti, and V. Govindaru, Freeman code based online handwritten character recognition for Malayalam using backpropagation neural networks. International journal on Advanced computing, 2012.
- [12] Omer, M.A.H. and S.L. Ma. Online Arabic handwriting character recognition using matching algorithm. in Computer and Automation Engineering (ICCAE), 2010 The 2nd International Conference on. 2010. IEEE.
- [13] Izakian, H., et al., Multi-font Farsi/Arabic isolated character recognition using chain codes. World Academy of Science, Engineering and Technology, 2008.
- [14] Ahmed, Rawia Ibrahim Omer. Offline Recognition System for Isolated Arabic Handwritten Characters using Hidden Markov Models. Diss. Sudan University of Science and Technology, 2017.
- [15] Abed, M.A., Freeman chain code contour processing for handwritten isolated Arabic characters recognition. Alyrmook University Magazine,Baghdad, 2012.