

**TRƯỜNG ĐẠI HỌC HÀ NỘI
KHOA CÔNG NGHỆ THÔNG TIN**

KÝ YẾU

HỘI NGHỊ KHOA HỌC SINH VIÊN KHOA CÔNG NGHỆ THÔNG TIN

2020

MỤC LỤC

ONLINE EXAMINATION SYSTEM APPLYING EMBEDDED SYSTEM AND ENCRYPTION.....	5
<i>SVTH: Nguyễn Thị Bích Loan</i>	
<i>GVHD: TS Tạ Quang Hùng</i>	
RESEARCH NETWORK SECURITY MONITORING SOLUTION IN CLOUD COMPUTING	15
<i>SVTH: Lương Quỳnh Hương, Đào Văn Quang</i>	
<i>GVHD: ThS Vương Thị Nhung</i>	
HUMAN RESOURCE MANAGEMENT SYSTEM	25
<i>SVTH: Đào Thị Hiền, Đặng Thành Lam Vũ Mai Loan, Trịnh Ngọc Sơn</i>	
<i>GVHD: ThS Nguyễn Đình Trần Long</i>	
DEVELOPING WEB APPLICATIONS USING MERN STACK	35
<i>SVTH: Nguyễn Hải Dang, Nguyễn Hoàng, Trần Văn Khang, Ngô Quốc Trung, Vũ Thành Trung</i>	
<i>GVHD: ThS Nguyễn Đình Trần Long</i>	
ARDUINO AND IOT APPLICATION USING ESP8266 NODEMCU.....	44
<i>SVTH: Nguyễn Thị Ngà</i>	
<i>GVHC: TS Tạ Quang Hùng</i>	
SENTIMENT ANALYSIS USING LONG SHORT-TERM MEMORY RECURRENT NEURAL NETWORK.....	48
<i>SVTH: Nhâm Gia Hoàng Anh, Nguyễn Tiến Huy Nguyễn Khắc Phúc, Bùi Đình Quân</i>	
<i>GVHD: ThS Bùi Quốc Khanh</i>	
MASK DETECTING PROGRAM	56
<i>SVTH: Phạm Quang Chiến, Nguyễn Thu Hà Vương Khánh Linh, Trần Thu Hiền, Vũ Văn Tùng</i>	
<i>GVHD: ThS Bùi Quốc Khanh</i>	

ANALYZE INTERFACES AND SUGGEST IMPROVEMENT FOR ENTERTAINMENT WEBSITES IN VIETNAM.....	63
<i>SVTH: Phạm Thanh Hà, Ngô Tuấn Anh Nguyễn Sỹ Nhật, Đỗ Hồng Hà</i>	
<i>GVHD: ThS Nguyễn Thị Ý Linh</i>	
HCI APPLICATION IN NUCLEAR REACTOR SYSTEM INTERFACE DESIGN.....	68
<i>SVTH: Lương Minh Hoàng, Đỗ Hồng Hà</i>	
<i>GVHD: ThS Nguyễn Thị Ý Linh</i>	
E-MARKETING TECHNIQUES APPLIED IN MOVIE LAUNCHING EVENT	75
<i>SVTH: Đặng Cẩm Tú, Đào Văn Quang, Nguyễn Như Quỳnh</i>	
<i>GVHD: ThS Nguyễn Thị Ý Linh</i>	
DIABETES PREDICTION ON WOMEN USING DATA MINING TECHNIQUE	83
<i>SVTH: Bùi Thị Anh Nguyệt, Vũ Thị Liệu, Lê Thị Bích Ngọc</i>	
<i>GVHD: ThS Bùi Quốc Khánh</i>	
ACCEPTANCE TESTING TYPES AND PROCESSES	92
<i>SVTH: Trần Thị Mai Hương, Trương Ngọc Quỳnh</i>	
<i>GVHD: ThS Đặng Định Quân</i>	
PHP LOOSE COMPARISON & TYPE JUGGLING	98
<i>SVTH: Phạm Đức Tùng, Đỗ Nguyễn Hoàng Ân</i>	
<i>GVHD: ThS Đặng Định Quân</i>	
APPLYING THREE-TIER ARCHITECTURE TO DEVELOP JAVA WEB APPLICATION FOR TRACKING CORONAVIRUS.....	106
<i>SVTH: Do Hai Bình, Trương Ngọc Quynh Nguyễn Công Quang, Quan Trọng Tu</i>	
<i>GVHD: ThS Nguyễn Định Trần Long</i>	
RESEARCH ABOUT BLOCKCHAIN TECHNOLOGY, TYPE OF ATTACKS AND SOLUTIONS.....	117
<i>SVTH: Lương Quỳnh Hương, Cao Thu Huyền, Lê Đức Anh</i>	
<i>GVHD: ThS Đỗ Thị Phương Thảo</i>	

HANU STUDENT CORE – AN ANDROID APPLICATION BASED ON EXISTING WEB SERVICE	130
<i>SVTH: Đỗ Hải Bình</i>	
<i>GVHD: ThS Luyện Thu Trang</i>	
EFFECTIVE VIRTUAL LEARNING ENVIRONMENT THROUGH MOODLE.....	139
<i>SVTH: Nguyễn Khắc Hiếu Nguyễn Thủ An,</i>	
<i>Lâm Thị Thuong Huyền</i>	
<i>GVHD: ThS Luyện Thu Trang</i>	
SOCIAL NETWORK PLATFORM	152
<i>SVTH: Nguyễn Thu Hà, Phạm Quang Chiến</i>	
<i>Vương Khánh Linh, Trần Thu Hiền</i>	
<i>GVHD: ThS Luyện Thu Trang</i>	
APPLYING MULTIMEDIA TO SUPPORT ACTIVE LEARNING AT HANOI UNIVERSITY	160
<i>SVTH: Nguyễn Thị Diễm, Lê Đức Lâm</i>	
<i>Nguyễn Dương Bích Quỳnh, Phạm Thị Quỳnh</i>	
<i>GVHD: ThS Đinh Thị Minh Nguyệt</i>	

ONLINE EXAMINATION SYSTEM APPLYING EMBEDDED SYSTEM AND ENCRYPTION

SVTH: Nguyễn Thị Bích Loan

GVHD: TS Tạ Quang Hùng

Abstract—Test is always considered a fair method to evaluate learner's ability, however, some academic dishonesty issues discovered these recent years have proved that there are still weaknesses in organizing tests. Therefore, this paper will demonstrate the development of an online examination system using embedded system, data encryption and website development technology that expects to solve the current drawbacks in conventional examination..

Keywords—Embedded systems, microcontroller, Arduino, RFID, AES encryption, PHP, Laravel, NodeJS, Classical Theory.

I. OBTAIN INFORMATION ON CARD USING RFID AND ARDUINO

A. Arduino and applications

Arduino is an open-source electronics platform containing both the board and software that allow users to optimize and build embedded system. Arduino was born at Ivrea Interaction Design Institute as the result of a studying about a simple and cheap enough tool for students to learn with technology. Indeed, Arduino has become popular over the world these years being such an easy-to-use and flexible system. Up to now, many versions of Arduino board have been released such as Arduino Uno, Arduino Duemilanove, Arduino Mega, etc. However, Arduino Uno is considered the most common and suitable model when it refers to embedded system, robotics or IoT.

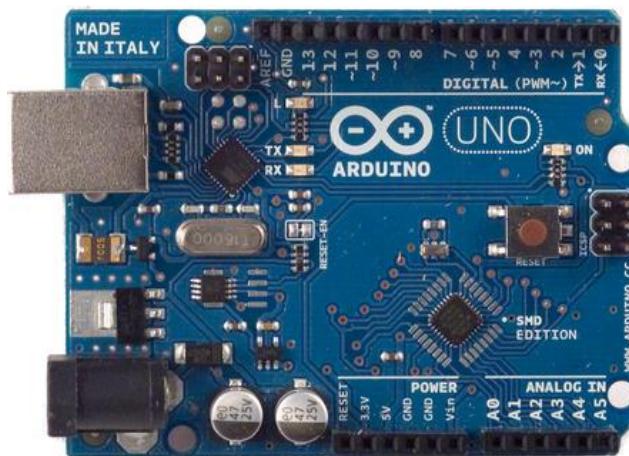


Figure 1. Arduino Uno R3 board

Regarding the hardware, though various of generations have been developed, all of Arduino boards were structured basing on an 8-bit Atmel AVR reduced instruction set computer (RISC) microprocessor. In particular, Arduino Uno contains an ATmega8U2 microcontroller programmed as a USB-to-serial converter, 6 analog input pins, 14 I/O digital ports which can receive electrical signal or create one as required. Besides, there is a

microcontroller that allows Arduino to connect to a computer via the USB.

One of the features making Arduino the flexible and convenient tool is the ability to give the physical part instructions through Arduino software IDE. This is an open-source cross-platform application whose programs (called “sketches”) written by C or C++, which can communicate with hardware like LEDs, LCDs or sensors. Moreover, if users have a stable Internet connection, Arduino also provides online editor which always updates the latest version of IDE.

In this paper, before Arduino is implemented into the system, there needs some preparations about hardware device. Apart from Arduino board connected with RFID module (RDM6300), some sample objects which contain RFID tags would be necessary. In reality, two plastic cards (each comprises a code of 7 numbers) have been used to communicate with Arduino sketch.

In particular, Arduino was applied to handle the information got from card via a RFID reader and return the pure ID. When the chain of numbers on card was obtained and sent to Arduino from RFID, it contains both the pure ID and a secure code to protect the completeness of the ID. The sketch written in Arduino IDE will process the calculation to check whether the ID is original. In the valid case, Arduino will return the card ID as output. On the next phrase, this output from Arduino will become input for authentication process on Laravel project.

```

readcard-full
1 #include <SoftwareSerial.h>
2 const int BUFFER_SIZE = 14; // RFID DATA FRAME FORMAT: 1byte head (value: 2), 10byte data (2byte version + 8byte tag)
3 const int DATA_SIZE = 10; // 10byte data (2byte version + 8byte tag)
4 const int DATA_VERSION_SIZE = 2; // 2byte version (actual meaning of these two bytes may vary)
5 const int DATA_TAG_SIZE = 8; // 8byte tag
6 const int CHECKSUM_SIZE = 2; // 2byte checksum
7
8 SoftwareSerial ssrfid = SoftwareSerial(6,8);
9
10 uint8_t buffer[BUFFER_SIZE]; // used to store an incoming data frame
11 int buffer_index = 0;
12
13 void setup() {
14   Serial.begin(9600);
15
16   ssrfid.begin(9600);
17   ssrfid.listen();
18
19   Serial.println("INIT DONE");
20 }
21 void loop() {
22   if (ssrfid.available() > 0){
23     bool call_extract_tag = false;
24
25     int ssvalue = ssrfid.read(); // read
26     if (ssvalue == -1) { // no data was read
27       return;
28     }
29     if (ssvalue == 2) { // RDM6300 found a tag => tag incoming
30       buffer_index = 0;
31     } else if (ssvalue == 3) { // tag has been fully transmitted

```

The serial monitor window shows the following output:

```

INIT DONE
TAG=9442876
-----
TAG=9442876
-----
TAG=9442876
-----
TAG=9442876
-----
TAG=9873276
-----
TAG=9873276
-----
TAG=9873276
-----
TAG=9873276
-----
TAG=9873276

```

Figure 2. Source for Arduino UNO to read, extract, and verify ID and result

B. RFID and applications

RFID (Radio-Frequency Identification) is a technology that allows users to get digital data encoded in RFID tags through a radio waves reader. Nowadays, a simple RFID set includes RFID tag encoded data, a RFID reader and an antenna. RFID tags are often

embedded in durable plastic as the passive ones because of the convenience and economy, while the active RFID allow the constant transmission of data through a power supply. A RFID reader system could be divided into different types according to the RFID tag types. There are Passive Reader Active Tag that only receives active tag signal, Active Reader Passive Tag which not only can transmit signals but also receive them from passive tags, and Active Reader Active Tag. The tag received from RFID readers can be transferred to a computer system and processed in the next phrases based on different intention of each project. Because of many advantages, which may be over the bar code technology when the covered tag is still obtainable, RFID has been applied in many areas like good tracking, people or animals tracking, access management, etc.



Figure 3. Access the building by card ID

In this study, RFID was applied to read ID from the plastic card that have been prepared. Once plastic card is placed on the antenna part (independently inserted on one side of device), antenna will receive the radio signals and send them to the RFID reader. Reader will become an intermediary acquiring raw information from card then transferring it to Arduino board for future analysis.

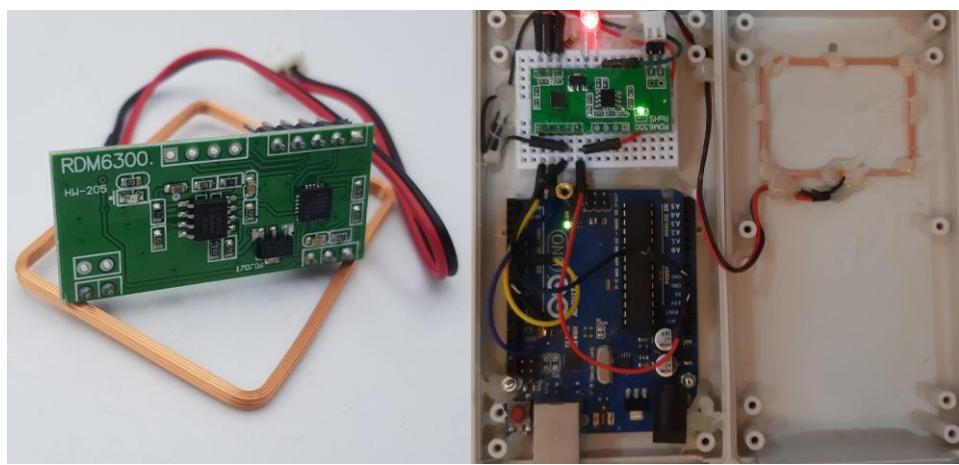


Figure 4. RFID module (RDM6300) is associated with Arduino Uno

II. ONLINE EXAM WEBSITE ALLOWS CONTENT ENCRYPTION USING AES ALGORITHM

A. AES encryption algorithm and application

AES is a block cipher, symmetric-key algorithm that uses only one key for both encryption and decryption process. AES was initially designed and named “Rijndael” by two Belgian cryptographers known as Joan Daemen and Vincent Rijmen, and was believed to replace its predecessor DES (Data Encryption Standard). The quantity of rounds in AES operation might vary and rely on the key’s length, for example, 128-bit keys will need 10 rounds, 192-bit keys demand 12 round. Each round in the encryption procedure contains four child-processes which are KeyExpansion and AddRoundKey, SubBytes, ShiftRows and MixColumns.

In the first step, the original key will be involved to generate a chain of keys used for the rest rounds of the encryption procedure. On the next AddRoundKey step, the block would be appended by original key with an XOR cipher (an additive encryption algorithm) because this is the first round. Following that, the SubBytes step allows every byte to be replaced by other characters based on a prearranged set. AES will pick up the necessary output from that collection. Next in ShiftRows phrase, rows would be moved forward by some spaces. For example, second row is shifted one space to the left, third row is shifted two spaces to the left, the last row is shifted three spaces to the left. Finally, linear transformation would be operated by matrix multiplication and bitwise XOR addition applied to each column. After that, AddRoundKey repeats, however, between the mixed column operation result and the new key cipher that was generated from the real key. Here is happening another round.

Refers to decryption, the process is just a piece of cake as long as long the key is provided. We only need to do every steps in the inverse way to get the origin secret message.

Moreover, user can obtain a higher level of security when using initialization vector (IV) which can be expressed in the form of arbitrary numbers or characters. In details, output from previous round of AES would be mixed with a part of data from the next round. For this action to be started, the first block needs to be mixed with a separate input value, which is initialization vector. Because of IV’s randomization, each encryption process with a certain key will use only one IV. As a result, AES could avoid the repetition when similar plaintext is encrypted under the same key, which leads to the same ciphertext.

AES has been implemented by various programming languages and shared in form of a library or package. In particular, this project will apply a library allowing to encrypt data using AES algorithm. This is a small part of OpenSSL, a well-known software library that secures communication over the Internet against eavesdropping. Initialization vector is also implemented on the library, user can simply invoke the method and pass enough required information the, library will give back a base-64 encoded output.

AES will be used to encode two factors on the website: file name of student’s test and the content of student’s test. File name of test contains every details of each student who has done that test, which includes full name, id of the card student used to log in the system and name of the shift when student do the exam. Because of its sensitiveness, this file name would be encoded immediately by AES algorithm after student submit their answer. The final result

stored in database later will be the ciphertext. Besides, system also encodes content of student's submission at the same time before storing. As a result, no one can have the possibility to recognize both the owner of each submission and file's content unless they can log in admin page.

```

62     $iv = openssl_random_pseudo_bytes(openssl_cipher_iv_length($this->cipher));
63
64     $input = $student->name . '+' . $student->card_id . '+' . $shift->name;
65     $test->file_name = openssl_encrypt($input, $this->cipher, $shift->key, 0, $iv);
66     $test->shift_id = $shift->id;
67     $test->iv = base64_encode($iv);
68
69     $test->save();
70
71     $content = openssl_encrypt(json_encode(array_slice($items, 1)), $this->cipher, $shift->key, 0, $iv);
72
73     Storage::disk('local')->put($test->id . '.txt', $content);
74
75     return view('student.ending', ['user' => $student]);
76
77 
```

Figure 5. Implementation of AES library in PHP

Encoding process will be successful when submission's secret details are hidden from irrelevant people. However, how can administrators collect the results and public those with the real information of students? Actually, administrators need to generate a key for every shift they create. This key later will be used for encoding process of file name and submission's content. After that, when the exam finishes, administrator can easily decode the file name that comprises details about student. Administrator become the only one who know exactly result of each student, however, is unable to change whatever relates to student's submission or result.

No	File name	Mark	Submitted at
1	V76AnxVEcvUDmA/I1t /hn2wwwun5kHFO10GeaGWj1Qevybqh31tyQQ+SGBaQkylM	2	2020-06-06 11:34:38
2	B5U4pqX9xoW89Czs2pkvK8FDX8gFsFTMW8HalRWO4arFE/PvCt6cNfo5dDgCrNap	6	2020-06-08 16:06:40

Figure 6. Table of encoded file name

No	File name	Mark	Submitted at
1	Nguyễn Thị Bích Loan+9873276+Math1	2	2020-06-06 11:34:38
2	Nguyễn Thị Bích Loan+9873276+Math1	6	2020-06-08 16:06:40

Figure 7. Table of decoded file name

B. Technology applied to build the website

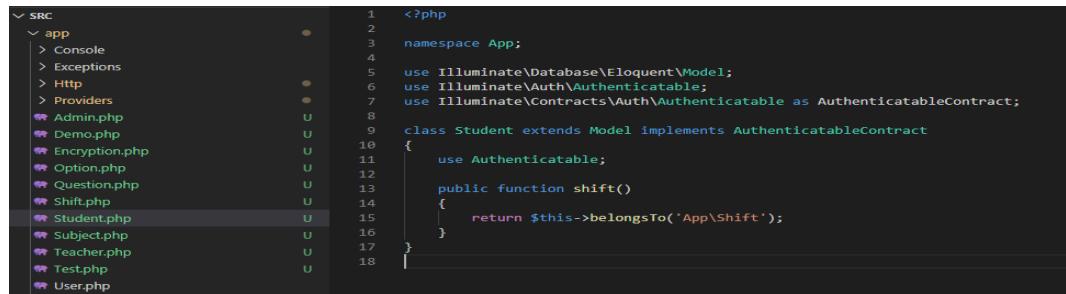
1) PHP programming language and PHP framework application

PHP was firstly introduced in 1995 by Rasmus Lerdorf, a Danish-Canadian programmer. After 25 years constantly being improved, PHP has become one of the most popular language for web applications. In the age of lightning development and mass appearance of websites, PHP is believed to continue to be favored over other languages in website development community.

Indeed, its advantages and effectiveness have been verified by a large community of programmers for the whole time. Firstly, the most impressive and comfortable about PHP is the simplicity which allows many people can easily learn this language and understand its concepts. Moreover, PHP owns a huge community. People could easily find many information as well questions about PHP on the Internet. These days, most of e-commerce websites are written in PHP using Laravel framework, a famous platform providing environment for bloggers like Wordpress uses PHP, even the most popular social network Facebook is written by PHP.

Because of these advantages, Laravel, a framework of PHP has been used for building this online exam model. Laravel is used to build web application based on model-view-controller (MVC) architectural pattern. It provides a series of feature that strongly support programmer such as Eloquent ORM (interact with database entities and their relationships), Blade templating engine (create final views by connecting some templates with data model), Migrations (serve a version control tool for database schemas), etc. Moreover, Laravel also provides powerful packages via Composer and Packagist as Socialite (authentication using social network account like Facebook or Google), Cashier (management of subscription billing services by Stripe) and many other ones.

Laravel is used to develop website based on MVC model, therefore, that pattern has been presented on this project through four main features. Refers to M (model), each database entity will be expressed via a model in “app” folder for future use. Normally, programmer need to use a relational database management system (for example MySQL) to interact with data in database. However, one of the most impressive feature of Laravel – Eloquent ORM – will allow programmer to easily work with database and relationships via relational mapping objects.



The screenshot shows a file browser interface with a tree view on the left and a code editor on the right. The tree view shows a directory structure under 'SRC': 'app' contains 'Console', 'Exceptions', 'Http', 'Providers', 'Admin.php', 'Demo.php', 'Encryption.php', 'Option.php', 'Question.php', 'Shift.php', 'Student.php', 'Subject.php', 'Teacher.php', 'Test.php', and 'User.php'. The 'Student.php' file is selected and shown in the code editor. The code is as follows:

```
1 <?php
2
3 namespace App;
4
5 use Illuminate\Database\Eloquent\Model;
6 use Illuminate\Auth\Authenticatable;
7 use Illuminate\Contracts\Auth\Authenticatable as AuthenticatableContract;
8
9 class Student extends Model implements AuthenticatableContract
10 {
11     use Authenticatable;
12
13     public function shift()
14     {
15         return $this->belongsTo('App\Shift');
16     }
17 }
18
```

Figure 8. Example of model Student in the online exam project

Now that system has the method to communicate with database, where could it have essential adjustment to the received data according to user's request? Though not be officially mentioned in the MVC design pattern, redirecting system "Route" in Laravel plays a crucial role in receiving user's orders and assign at which they will be processed. Every available request that system can serve in form of HTTP request via URL will be defined in web.php file (locates in "routes" folder), and controller which is responsible for handling request will be assigned simultaneously.



```

routes > 🌐 web.php > ⚙ #Function#a314c2c9
15
16 | Route::get('/', 'AuthenticateController@index')->name('index');
17 | Route::post('/check', 'AuthenticateController@check')->name('check');
18 | Route::post('/auth', 'AuthenticateController@auth')->name('auth');
19 | Route::post('/login', 'AuthenticateController@login')->name('login');
20 |
21 | Route::get('/admin', 'AuthenticateController@indexAdmin')->name('admin');
22 | Route::post('/admin/auth', 'AuthenticateController@authenticate')->name('authenticate');
23 | Route::post('/admin/logout', 'AuthenticateController@logoutAdmin')->name('logoutAdmin');
24 |
25 | Route::middleware('auth')->group(function () {
26 |     Route::get('/homepage', 'HomeController@index')->name('homepage');
27 |     Route::get('/test/{slug}', 'TestController@index')->name(['test']);
28 |     Route::post('/create', 'TestController@store')->name('create');
29 |     Route::post('/logout', 'AuthenticateController@logout')->name('logout');
30 |
31 |     Route::get('/demo', 'DemoController@encryptRSA');
32 });
33

```

Figure 9. A part of file web.php in online exam project

Depending on different objects or activities, there will be distinct controller classes created in folder app/Http/Controller. Here is where every logic process of system happens such as querying data from database, receiving inputs from user and handling them to return proper output, etc. Every controller files are placed inside app/Http/Controllers directory and extend a base Controller class. Inside each function defined in a controller class, after logic processes have been successfully handled, there is always a return statement to some views for showing result. Controller's task will conclude at that point.



```

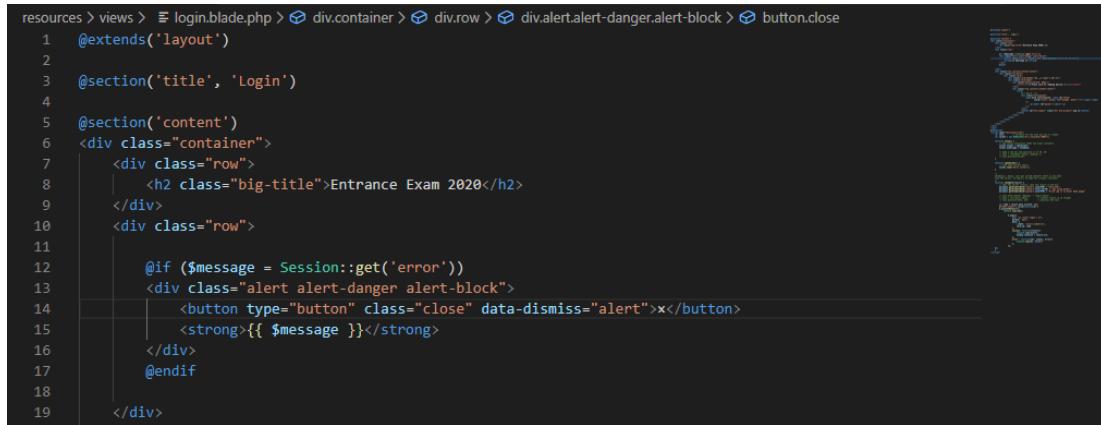
11  class ShiftController extends Controller
12  {
13      public function index()
14      {
15          $user = session()->get('user');
16          $items = Shift::orderBy('id', 'desc')->paginate(10);
17
18          return view('admin.shift.index', compact('items', 'user'));
19      }
20
21      public function search(Request $request)
22      {
23          $user = session()->get('user');
24
25          $items = Shift::query()
26              ->where('name', 'LIKE', "%{$request->keyword}%")
27              ->paginate(10);
28
29          return view('admin.shift.index', compact('items', 'user'));
30      }

```

Figure 10. Example of a Controller class in online exam system

The ending at each controller's function leads to the duty of View, the last feature in MVC pattern design. As can be seen from the previous section, the UI and UX that system releases for user will be totally included in view. Here is the place where Blade templating engine is responsible for. These file could be found in form of .blade.php file extension and

located in resources/views directory. View files allow programmers to associate not only HTML, CSS, Javascript (relating to front end development) but also PHP.



```

resources > views > login.blade.php > div.container > div.row > div.alert.alert-danger.alert-block > button.close
1  @extends('layout')
2
3  @section('title', 'Login')
4
5  @section('content')
6  <div class="container">
7      <div class="row">
8          <h2 class="big-title">Entrance Exam 2020</h2>
9      </div>
10     <div class="row">
11
12         @if ($message = Session::get('error'))
13             <div class="alert alert-danger alert-block">
14                 <button type="button" class="close" data-dismiss="alert">×</button>
15                 <strong>{{ $message }}</strong>
16             </div>
17         @endif
18
19     </div>

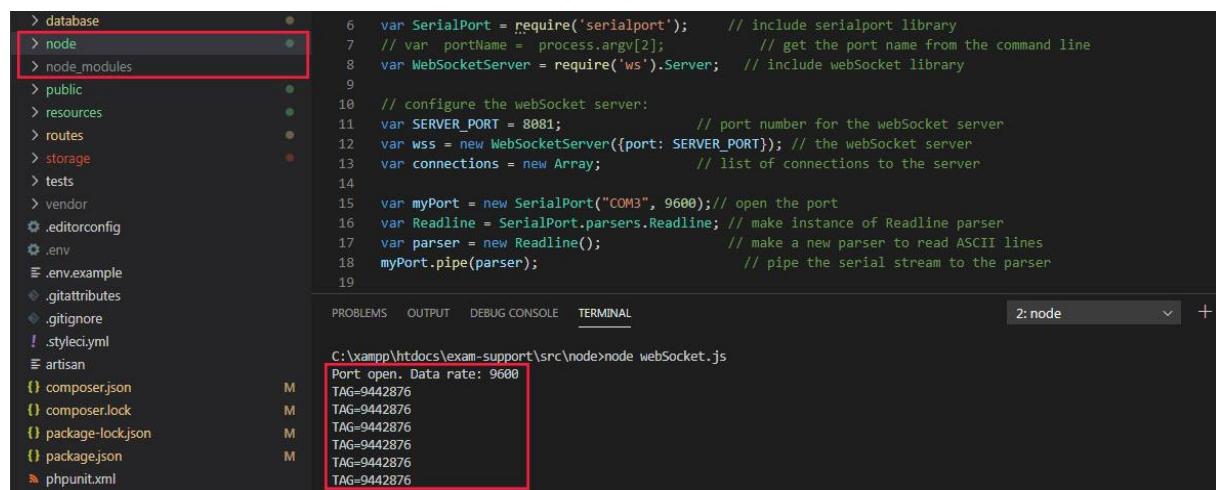
```

Figure 11. A snippet of View file login.blade.php

2) NodeJS in receiving information from Arduino

NodeJS is an open-source and cross-platform run-time environment written in Javascript that allows programmer to develop server side scripts. This platform was built by Ryan Dahl in 2009 on Google Chrome's JavaScript Engine. As a new technology but NodeJS has proved its efficiency via a plentiful library of Javascript functions which support considerably to the development of web applications.

NodeJS can be easily installed via command line interface of Visual Studio, with “node” is the location where programmers place classes for their project and “node_modules” for NodeJS library. In order to listen to the data sent from Arduino, NodeJS file need to include “serialport” and “ws” module. Particularly, the former is responsible for the reading the input from Arduino, in this situation, it is a string containing the card ID, then data will be stored for future transfer. On the other hand, the later creates a port to listen to event when user place card on RFID device. After card ID has been obtained, they will be sent to a controller class handling authentication in PHP via an Ajax request.



```

> database
> node
> node_modules
> public
> resources
> routes
> storage
> tests
> vendor
> .editorconfig
> .env
> .env.example
> .gitignore
> .gitattributes
> artisan
> composer.json
> composer.lock
> package-lock.json
> package.json
> phpunit.xml
6  var SerialPort = require('serialport'); // include serialport library
7  // var portName = process.argv[2]; // get the port name from the command line
8  var WebSocketServer = require('ws').Server; // include webSocket library
9
10 // configure the websocket server:
11 var SERVER_PORT = 8081; // port number for the webSocket server
12 var wss = new WebSocketServer({port: SERVER_PORT}); // the webSocket server
13 var connections = new Array(); // list of connections to the server
14
15 var myPort = new SerialPort("COM3", 9600); // open the port
16 var Readline = SerialPort.parsers.Readline; // make instance of Readline parser
17 var parser = new Readline(); // make a new parser to read ASCII lines
18 myPort.pipe(parser); // pipe the serial stream to the parser
19

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL

C:\xampp\htdocs\exam-support\src\node>node webSocket.js
Port open. Data rate: 9600
TAG=9442876
TAG=9442876
TAG=9442876
TAG=9442876
TAG=9442876
TAG=9442876

Figure 12. NodeJS structure and output from a request

III. APPLYING CLASSICAL TEST THEORY TO EVALUATE DIFFICULTY LEVEL OF MULTIPLE-CHOICE QUESTIONS

A. Overview of Classical Test Theory

Classical Test Theory (CTT) is a theory relating to educational measurement and psychometrics, which is designed to evaluate the difficulty of questions or the capacity of examinees. This hypothesis was beginning to be developed in the early of 20th century and mentioned approximately similar with true score theory. In particular, there are two typical features which could be obtained from CTT, the first one is tests and scores evaluation (based on reliability) and the other is items evaluation (based on a p value and item-total correlations). This paper will focus on items evalution.

B. Items evaluation and application

Items evaluation contains calculations to score the difficulty level of each question in a multiple-choice question test. This quantity is often referred as a p value, which is the proportion of candidates giving the correct option of a question over the total of candidates take that question. Ideally, the accepted proportion of difficulty for each question is in the range 0.25 – 0.75.

In this thesis, CTT theory has been applied to the group of multiple-choice questions. System wholly filters the question, separates multiple-choice question from short answer questions and gather their corresponding options. Most of the questions include four options A, B, C and D. On next step, system collects all the submissions and decode their content to figure out how many times each option was chosen. Then it does some simple calculation to obtain the proportion previously mentioned in the above paragraph. Finally, website returns result in form of table. Administrators have the authorization to access these statistics from their admin interface and could adjust and improve the question or its options for better in the future.

Statistics of multiple-choice questions

No	Question	Quantity	A	B	C	D	Total
1	She _____ French words for hours, but she still doesn't remember all of them	Sum	7	3	8	2	20
		Percentage	0.35	0.15	0.4	0.1	
2	The boy waved his hands to his mother, who was standing at the school gate, to _____ her attention	Sum	10	6	2	2	20
		Percentage	0.5	0.3	0.1	0.1	
3	He is one of the most _____ bosses I have ever worked with. He behaves rudely to not only me but also others in the staff	Sum	5	10	2	3	20
		Percentage	0.25	0.5	0.1	0.15	
4	If a boss wants to have a well-qualified staff, he should have to pay his employees _____	Sum	3	8	6	3	20
		Percentage	0.15	0.4	0.3	0.15	
5	I'm very happy _____ in India. I really miss being there.	Sum	10	6	0	4	20
		Percentage	0.5	0.3	0	0.2	
6	They didn't reach an agreement _____ their differences.	Sum	10	4	5	1	20
		Percentage	0.5	0.2	0.25	0.05	

Figure 13. Table of data gathered and analyzed using CTT

IV. CONCLUSION

These days, online test has been relatively popular, but there is rarely a secure enough model developed for an official examination. This paper has presented how a system of Arduino and RFID could be associated with AES encryption and website development tools to build a reliable online exam system. Besides, a simple illustration of Classical Test Theory has proved that educational measurement can be totally implemented effectively on this system for future improvement of question bank.

V. ACKNOWLEDGEMENTS

This research is a part of final-year project in the course Information Technology, Hanoi University. I would like to thank all the teachers in the Faculty of Information Technology Hanoi University who dedicated teaching, equipping me with valuable knowledge during the past 4 years. Especially, I would like to send my gratefulness to my instructor, Dr. Hung Ta, who has spent a lot of time and dedication to help me complete this research. Valuable comments and suggestions of the reviewers on this paper are highly appreciated.

REFERENCES

- [1] Arduino. [Online]. Available: <https://www.arduino.cc>. [Accessed May 25, 2020].
- [2] D. Hayes, “Why Use PHP in 2019?,” *Thoughtful Code*, June 12, 2018. [Online]. Available: <https://www.thoughtfulcode.com/why-use-php>. [Accessed June 3, 2020].
- [3] J. Lake, “What is AES encryption and how does it work?,” Comparitech, February 17. 2020. [Online]. Available: <https://www.comparitech.com/blog/information-security/what-is-aes-encryption/>. [Accessed Jun 1, 2020].
- [4] J. Parab, S. A. Shinde, V. G. Shelake, R. K. Kamat, and G. M. Naik, “*Practical Aspects of Embedded System Design using Microcontrollers*”, Springer, 2008.
- [5] NodeJS. [Online]. Available: <https://nodejs.org>. [Accessed May 25, 2020].
- [6] Project Hub. [Online]. Available: https://create.arduino.cc/projecthub/Arca_Ege/using-dht11-b0f365. [Accessed May 20, 2020]
- [7] R. Thomas, “Advanced Encryption Standard (AES): What It Is and How It Works”, *Hashedout*, April 23, 2020. [Online]. Available: <https://www.thesslstore.com/blog/advanced-encryption-standard-aes-what-it-is-and-how-it-works>. [Accessed Jun 2, 2020].
- [8] T. Q. Lâm, *Trắc Nghiệm và Ứng dụng*. Hanoi: Khoa Học và Kỹ Thuật, 2008.

RESEARCH NETWORK SECURITY MONITORING SOLUTION IN CLOUD COMPUTING

SVTH: Lương Quỳnh Hương, Đào Văn Quang

GVHD: ThS Vương Thị Nhung

Abstract – Nowadays, Cloud computing is becoming an importance part of human life. However, security monitoring in Cloud computing has received limited attention from the researchers. In this paper, we discuss some main concepts, advantages, solutions for security monitoring in Cloud computing.

Keywords - Cloud computing, security monitoring in cloud computing

I. INTRODUCTION

During the Covid-19 pandemic, along with the trend of working from home, cloud computing also changed the way businesses and governments operated. However, this development also created new security challenges. Therefore, it is necessary to have information security monitoring solutions for cloud computing services, thereby making recommendations to help businesses and users take preventive measures to minimize the data's loss and the risk of information leakage, when providing as well as participating in the use of cloud services. The aim of this paper is to discuss about information security issues in cloud computing and suggest some solutions for monitoring network security in cloud computing. Another key point featured in this research is that we have demo version to build and integrate monitoring tools into the cloud environment. As such, it is our belief that this report can serve as a prototype as well as a potential point of view for large and further research in the future.

II. BACKGROUND

A. Cloud computing

In mid-2007, the term "Cloud computing" was born with the aims to recap the new directions of information technology development thanks to the broadband Internet and the giant computing centers such as Google, Amazon, IBM, Microsoft ... According to NIST, "Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction." It includes three service models: Infrastructure as a service (IaaS), Platform as a service (PaaS), and Software as a service (SaaS).

1. Infrastructure as a service (IaaS)

IaaS providers provide the infrastructure such as virtual server, master repository, and APIs. There are some famous IaaS providers like Amazon AWS, Microsoft Azure, Google Compute Engine, and IBM Smart Cloud Enterprise.

2. Platform as a service (PaaS)

The PaaS model allows users to use development tools on the infrastructure of cloud service providers via the internet using APIs, software portal or website portal. The users can use this platform to create, develop and run own application. Some of famous PaaS are Amazon Web Services (AWS) Elastic Beanstalk, Oracle Cloud Platform (OCP), Google App Engine and Microsoft Azure.

3. Software as a service (SaaS)

SaaS is a software distribution model provides software applications, known as web services, over the internet. By using internet connection, users can access SaaS applications and services from anywhere. For example, Google Docs is one of the popular SaaS application of Google.

B. Network security monitoring

Network security monitoring is term which describes the collection of information on the components of the system, analyzing information and signs to evaluate and give warnings to system administrators. The object of network security monitoring is all components and equipment in the network. It includes workstations, database, applications, servers and network devices.

Monitoring network security plays an important role in securing cloud services. It is the process of assessing, monitoring and managing safety for services, applications and infrastructure on the cloud.

Automatic and manual management and monitoring techniques ensure the availability and performance of websites, servers, applications and cloud infrastructure. Constantly assess resource levels, server response time, speed, availability, and anticipate potential future issues.

Monitoring in the cloud environment mainly focuses on website monitoring, virtual machine monitoring, database monitoring, virtual network monitoring, storage monitoring. Cloud monitoring is an easy way to identify potential security vulnerabilities in cloud infrastructure. From there, help organizations more reliable in using the cloud to transfer and store data.

III. THE URGENCY OF NETWORK SECURITY MONITORING IN CLOUD COMPUTING AND SOLUTIONS

A. Problem

Microsoft recently released its 22-day cybersecurity report, as organizations move to cloud computing, increasing the frequency and complexity of cyberattacks targeting "the cloud" also increased.

In 2018, the Chinese hacker group nicknamed "Red Apollo" launched one of the largest

global cyber espionage campaigns ever maintained. Instead of attacking companies directly, it targets cloud service providers, trying to use their networks to spread spyware to many companies.

Thereby we realize the cloud computing environment is threatened by hackers. They take advantage of security breaches to penetrate into the system to perform behaviors that cause losses to businesses and users on cloud computing environment. Therefore, monitoring security in cloud computing environment is an urgent issue. That helps the cloud infrastructure work stably as well as prevent unauthorized access by hackers.

B. *Solution*

After studying and researching a number of solutions, including special cloud monitoring solutions designed specifically for a cloud with different purposes including commercial purposes such as:

1. Amazon CloudWatch

Amazon CloudWatch is a monitoring solution for Amazon Web Services (AWS). Amazon CloudWatch allows easy tracking of basic processing and storage data. It provides the ability to automatically create statistics tables or users can customize arbitrary configurations. Amazon CloudWatch is a good solution for Amazon cloud users and managers. However, it is limited to AWS products.

2. Private Cloud Monitoring Systems

This is a private cloud monitoring solution; it is open source that uses a layer called "Integration" to provide uniform access for cloud resources and users. It provides a unified infrastructure solution and is independent of resources stored in a cloud. However, this solution has to configure and install manually, which affects the flexibility when deploying.

3. Runtime Model for Cloud Monitoring

The runtime model for cloud monitoring with purpose to monitor resources through abstract models, which helps to consistently handle heterogeneous resources. However, it requires constant updating of monitoring resources to maintain a consistent model. The main drawback of this solution involves manual installation and configuration of specific agents. For this reason, cloud monitoring capabilities such as scalability and migration are limited.

4. Demo

In addition, we have researched, researched and demoed a common solution for monitoring network security in cloud computing. A common solution was created to monitor systems without worrying about the specificities associated with each type of system. These solutions are widely used in information systems to retrieve information about resources. Although common solutions are often used before the cloud computing environment appears, they can be completely used in the cloud. In cloud environments, these solutions can be used to monitor basic data (memory, network, etc.) of virtual machines or servers that provide

cloud computing services, they have It can collect information to build charts from aggregated data and show parameters such as bandwidth consumption, network links or processes and status of services in the system. Besides, IDS / IPS systems such as snort, Suricata can also be deployed to detect network intrusion in cloud environment. It can detect and issue alerts when detecting attacks inside and outside the network on cloud computing environment.

In demo for solution of network security monitoring in cloud computing. We were built an Openstack cloud computing environment to provide infrastructure services as a service (IaaS). Openstack is an open source software used to deploy cloud computing.

Use snort network security monitoring tools to detect network attacks. Snort is a network intrusion detection system, or NIDS (Network intrusion detection system). Snort is a free and open source with many great features in protecting the internal system, detecting and preventing outside attacks on the system.

Snort, developed by Martin Roesch in 1998, is not only free but they have many powerful features that not all products have. Snort is currently operated by Sourcefire and is part of the firewall development platform Checkpoint. Snort has been deployed throughout the world. With millions of downloads and more than 400,000 registered users, Snort has become the standard for intrusion prevention and detection systems. The main function of Snort is packet sniffing, packet logging and network-based intrusion detection. Designed on the module to check incoming and outgoing packets, detect abnormal packets. Snort can be run on many platforms such as Window, Linux, MacOS, OpenBSD, AIX ... Snort can be configured to run as an A network intrusion detection system (NIDS). They support the interoperability of protocols such as Ethernet, Token Ring, 802.11, Cisco HDLC etc. Although there are intrusion detection methods, Snort is considered the best system.

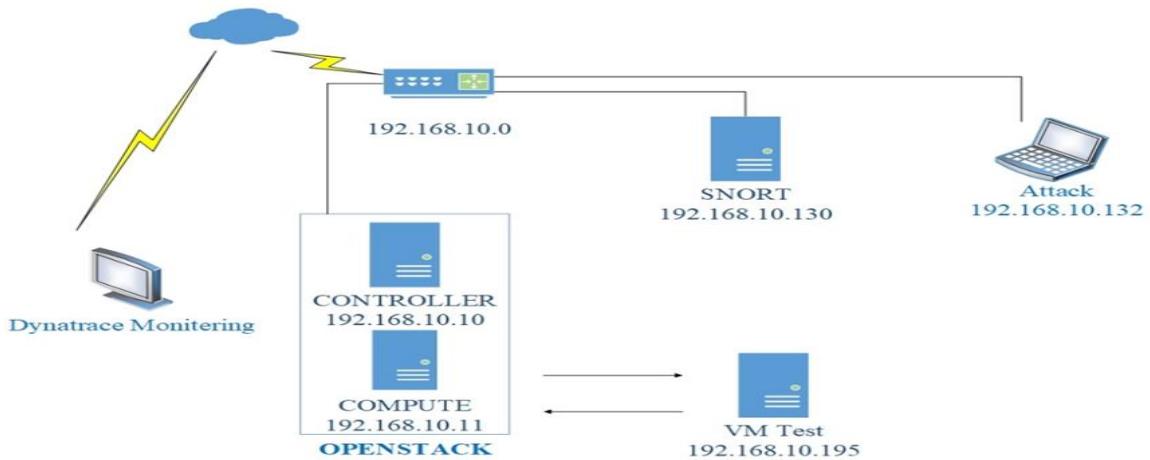


Figure 1: Model

When Snort works, it will listen for all the packets moving through it. The captured packet will be put into the module for analysis. Snort monitors all events in the network and reconciles them with the set of rules and then issues alerts when it detects network attacks.

The advantage of snort is Managing network segments. Simple installation and maintenance, without affecting the network. Avoid being attacked by denial of Dos service to a specific host. Able to determine the entrance at the network layer, independent of the operating system. Attackers cannot detect if the system is using Snort. Besides, the rules of snort are frequently added and updated with new forms, easily downloaded from www.snort.org. Able to detect large numbers of different intrusions. Snort operates 24/7 in real time, helping to continuously detect attacks. Has a large Snort user community. Snort does not need to replace any existing security infrastructure.

Use the Dynatrace Monitor installed on the system called OneAgent. OneAgent is basically a set of specialized processes that run on each monitored server. OneAgent collects metrics from the operating system it runs on in real time and provides server performance metrics. Users through the UI are provided on the website and login account to access performance monitoring of the server installed OneAgent.

Proceed 1:

Scenario of port scanning attack. Using an attacker machine to perform a Zenmap port scanning attack on a virtual VM test server created from cloud computing to provide IaaS services to customers. An attacker could perform a port scan to probe and search for information that targets the organization's virtual server.

When performing an attack, the snort will detect an attack on the virtual server and will issue a warning.

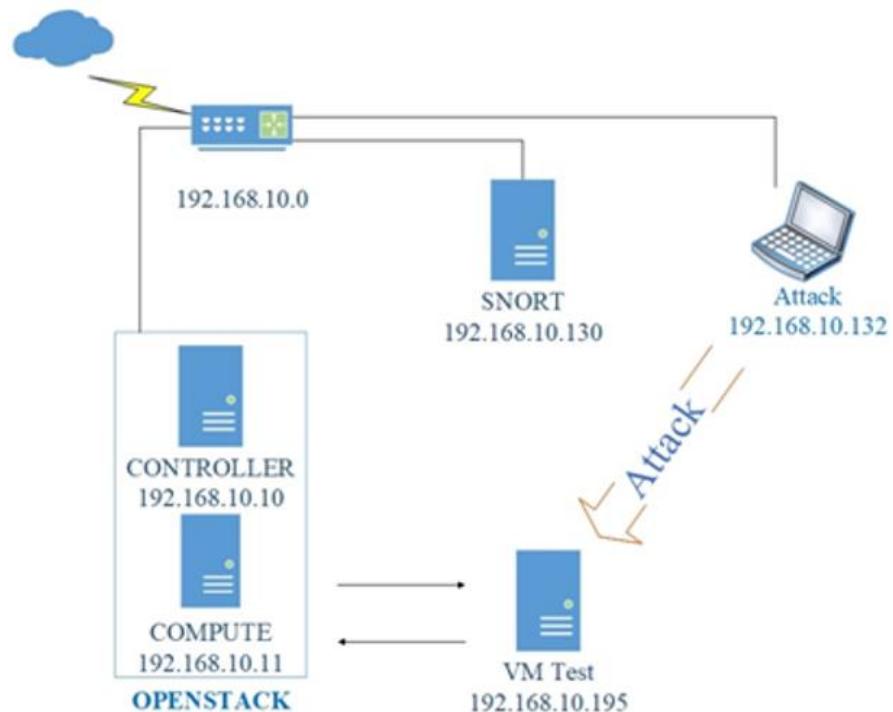


Figure 2: Zenmap attack model

On the attacker's machine, perform a port scan with Zenmap with the IP address of the virtual server VM test address.

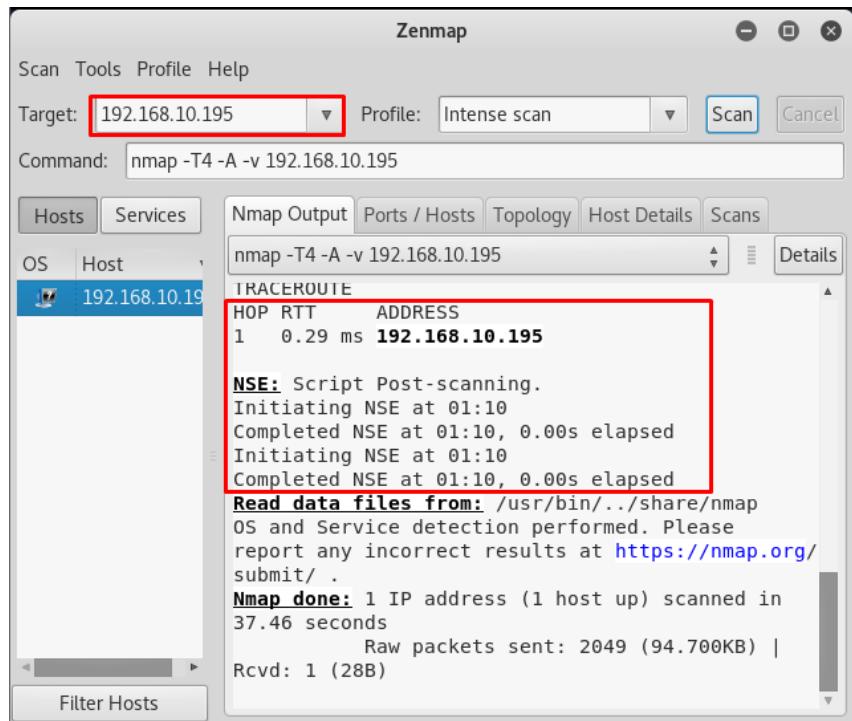


Figure 3: Perform a port scan with Zenmap

The results show that on the snort server, the snort listens for packets moving through the network and collates with the rule set to issue an alert.

```
Preprocessor Object: SF_SIP Version 1.1 <Build 1>
Commencing packet processing (pid=48365)
05/23-01:09:57.389697 [**] [1:10000001:1] ZenMap test [**] [Priority: 0] {ICMP} 192.168.10.132 -> 1
92.168.10.195
05/23-01:09:57.415309 [**] [1:10000001:1] ZenMap test [**] [Priority: 0] {ICMP} 192.168.10.132 -> 1
92.168.10.195
05/23-01:09:57.544528 [**] [1:10000001:1] ZenMap test [**] [Priority: 0] {ICMP} 192.168.10.132 -> 1
92.168.10.195
05/23-01:09:57.570325 [**] [1:10000001:1] ZenMap test [**] [Priority: 0] {ICMP} 192.168.10.132 -> 1
92.168.10.195
05/23-01:09:57.697840 [**] [1:10000001:1] ZenMap test [**] [Priority: 0] {ICMP} 192.168.10.132 -> 1
92.168.10.195
05/23-01:09:57.723359 [**] [1:10000001:1] ZenMap test [**] [Priority: 0] {ICMP} 192.168.10.132 -> 1
92.168.10.195
05/23-01:09:57.852250 [**] [1:10000001:1] ZenMap test [**] [Priority: 0] {ICMP} 192.168.10.132 -> 1
```

Figure 4: Snort issues a warning

Warning includes time of attack, method of attack, attack via ICMP protocol address source of attack 192.168.10.132 to destination 192.168.10.195.

Proceed 2:

The slowloris attack script on the controller server running the openstack service. In fact, an attacker could make an attack on the cloud provider's cloud infrastructure, making it impossible for a customer to access the cloud infrastructure provider's website.

When an attacker performs an attack, the snort will detect and issue a warning about the attack. At the same time, the Dynatrace monitoring system also displays parameters, data related to network traffic, the ability to connect to the service and alerts to users.

Based on the information above, the administrator may have solutions to troubleshoot and maintain a stable state for the system.

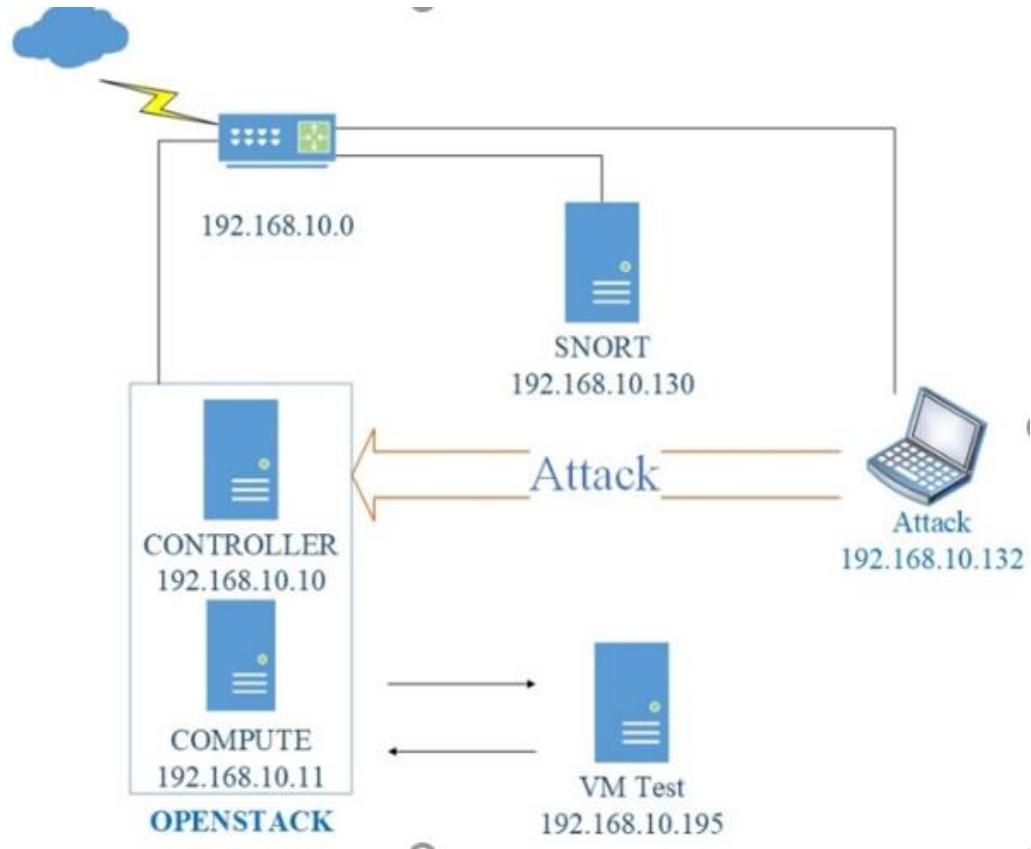


Figure 5: Slowloris attack model

The attacker performs a Slowloris attack on the controller server running the Openstack service with the IP address: 192.168.10.10.

```
File Edit View Search Terminal Help
root@castr0l:~/Desktop/slowloris.pl-master$ perl slowloris.pl -dns 192.168.10.10
Welcome to Slowloris - the low bandwidth, yet greedy and poisonous HTTP client by Laera
Loris
Defaulting to port 80.
Defaulting to a 5 second tcp connection timeout.
Defaulting to a 100 second re-try timeout.
Defaulting to 1000 connections.
Multithreading enabled.
Connecting to 192.168.10.10:80 every 100 seconds with 1000 sockets:
    Building sockets.
    Building sockets.
    Building sockets.
    Sending data.
Slowloris.pl
Current stats: Slowloris has now sent 358 packets successfully.
This thread now sleeping for 100 seconds...
    Building sockets.
    Building sockets.
    Sending data.
Slowloris.pl
Current stats: Slowloris has now sent 800 packets successfully.
This thread now sleeping for 100 seconds...
    Building sockets.
    Building sockets.
    Building sockets.
    Building sockets.
    Building sockets.
```

Figure 6: Performing a Slowloris attack

Using a web browser to access the controller's website, the server cannot serve the visitor at the moment, affecting the availability of the service.

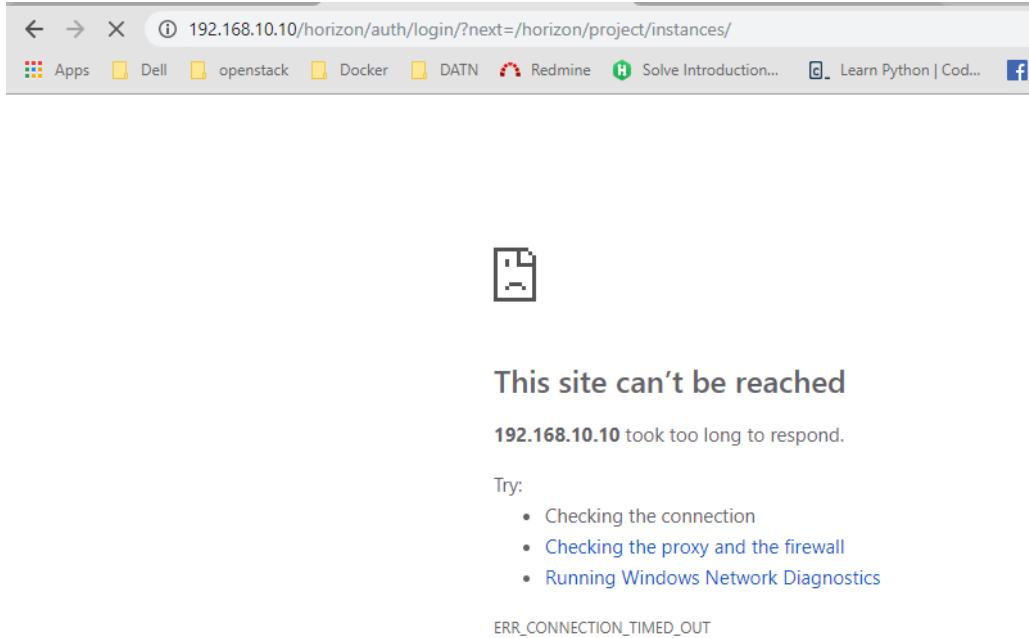


Figure 7: Unable to connect to the website

The results show that on snort, the system has detected and warned about the attack. The attack information includes the time affected, the method of attack, the source address attack protocol and the attack port, as well as the target address and the port under attack.

```
nial of Service Attack] [Priority: 2] {TCP} 192.168.10.132:53186 -> 192.168.10.10:80
05/22-11:34:34.999124 [**] [1:1:1] SlowLoris.py DoS attempt [**] [Classification: Detection of a De
nial of Service Attack] [Priority: 2] {TCP} 192.168.10.132:53196 -> 192.168.10.10:80
05/22-11:34:34.999173 [**] [1:1:1] SlowLoris.py DoS attempt [**] [Classification: Detection of a De
nial of Service Attack] [Priority: 2] {TCP} 192.168.10.132:53204 -> 192.168.10.10:80
05/22-11:34:34.999253 [**] [1:1:1] SlowLoris.py DoS attempt [**] [Classification: Detection of a De
nial of Service Attack] [Priority: 2] {TCP} 192.168.10.132:53212 -> 192.168.10.10:80
05/22-11:34:34.999255 [**] [1:1:1] SlowLoris.py DoS attempt [**] [Classification: Detection of a De
nial of Service Attack] [Priority: 2] {TCP} 192.168.10.132:53252 -> 192.168.10.10:80
05/22-11:34:34.999348 [**] [1:1:1] SlowLoris.py DoS attempt [**] [Classification: Detection of a De
nial of Service Attack] [Priority: 2] {TCP} 192.168.10.132:53270 -> 192.168.10.10:80
05/22-11:34:34.999349 [**] [1:1:1] SlowLoris.py DoS attempt [**] [Classification: Detection of a De
nial of Service Attack] [Priority: 2] {TCP} 192.168.10.132:53282 -> 192.168.10.10:80
05/22-11:34:35.203211 [**] [1:1:1] SlowLoris.py DoS attempt [**] [Classification: Detection of a De
nial of Service Attack] [Priority: 2] {TCP} 192.168.10.132:54874 -> 192.168.10.10:80
05/22-11:34:35.203220 [**] [1:1:1] SlowLoris.py DoS attempt [**] [Classification: Detection of a De
nial of Service Attack] [Priority: 2] {TCP} 192.168.10.132:54872 -> 192.168.10.10:80
05/22-11:34:48.158942 [**] [1:1:1] SlowLoris.py DoS attempt [**] [Classification: Detection of a De
nial of Service Attack] [Priority: 2] {TCP} 192.168.10.132:54904 -> 192.168.10.10:80
05/22-11:34:48.163045 [**] [1:1:1] SlowLoris.py DoS attempt [**] [Classification: Detection of a De
nial of Service Attack] [Priority: 2] {TCP} 192.168.10.132:54908 -> 192.168.10.10:80
05/22-11:34:57.923085 [**] [1:1:1] SlowLoris.py DoS attempt [**] [Classification: Detection of a De
nial of Service Attack] [Priority: 2] {TCP} 192.168.10.132:54944 -> 192.168.10.10:80
```

Figure 8: Snort issued a warning about the Slowloris attack

IV. EVALUATION

Through the above two proceedings, we realize that attacks on cloud service providers or customers are detected and alerted by snort security monitoring tools. Despite the many solutions outlined above, we find snort to be an easy-to-use, free tool suitable for initial

research in cloud security and attack analysis basic.

In addition, building Snort intrusion prevention and detection system is the solution to improve the security of the system. This system to supplement and collect information for the process of preventing attacks, not mean that replacing the firewall. Because of the nature and characteristics of Snort, it is not only suitable for companies providing cloud computing services, customers using the service, but also suitable for those who are researching and learning background solutions. for network security monitoring in cloud computing. Through the demo, we also learn more about Snort's working structure to optimize its use for the right purpose, for the specific situations of each system. The results returned after using Snort are notified in writing, each line is easy to manage and detect attack errors. Seldom give false results, especially effective for current forms of aggression. Quick, reliable identification of attack tools and techniques. From there, the system administrator can quickly take measures to handle promptly.

V. CONCLUSION

In this research, we realize that the attacker can make an attack on the cloud service provider itself and also perform an attack on the service. The cloud is provided to customers so cloud security monitoring is extremely necessary and important. It helps us to monitor and detect cyber-attacks in order to have appropriate solutions to overcome problems and enhance the security of the system. *In addition*, the solution for network security on cloud computing we aim to be a common solution that uses Snort to detect attacks and intrusions. Because Snort is a highly regarded solution, many of the features and advantages we have mentioned.

REFERENCES

- [1] B. Balakrishnan, Book: “Cloud Security Monitoring”, SANS, 2017.
- [2] B. Wheeler, B. Griffin, “Cloud computing monitoring and management system”, 2014.
- [3] D. Linthicum, “Learning Cloud Computing: Monitoring and Operations”, 2019.
- [4] D. T. Truong, “How Cloud Computing Enhances Competitive Advantages: A Research Model for Small Businesses”, 2010.
- [5] G. D. C. Rodrigues, “Monitoring of Cloud Computing Environments: Concepts, Solutions, Trends and Future Directions”, 2016.
- [6] J. Koziol, “Intrusion Detection with Snort”, 2003.
- [7] L. Dupré, T. Haeberlen, “Cloud Computing Benefits, risks and recommendations for information security”, Rev.B, December 2012.
- [8] M. J. Kavis, “Architecting the Cloud: Design Decisions for Cloud Computing Service Models (SaaS, PaaS AND IaaS)”, 2014.

- [9] M. Zareapoor, P. Shamsolmoali, M. A. Alam, “Establishing Safe Cloud: Ensuring Data Security and Performance Evaluation”, *International Journal of Electronics and Information Engineering*, vol.1, no.2, pp.88-99, 2014.
- [10] N. T. Khanh, “Nghiên cứu về bảo mật trong điện toán đám mây”, Học viện công nghệ bưu chính viễn thông, 2012.
- [11] R. L. K. R. D. Vines, R. L. Krutz, “Cloud Security: A Comprehensive Guide To Secure Cloud Computing”, 2010.
- [12] T. C. Đệ, “Tổng quan về an ninh trên điện toán đám mây”, *Tạp chí Khoa học Trường Đại học Cần Thơ, Số chuyên đề: Công nghệ Thông tin*, pp.39-46, 2013.
- [13] T. Kuldeep, S.S Tyagi, A. Richa, “Snort Intrusion Detection System in Cloud Environment”, ISSN 0974-2239, Vol. 4, No, 3, pp. 329-334, 2014.
- [14] TS. Chou, “Security threats on cloud computing vulnerabilities”, vol. 5, no. 3, June 2013.
- [15] V. Casola, A. D. Benedictis, “Critical Infrastructure Protection Research”, pp. 123-139, 2015.

HUMAN RESOURCE MANAGEMENT SYSTEM

SVTH: Đào Thị Hiền, Đặng Thành Lam

Vũ Mai Loan, Trịnh Ngọc Sơn

GVHD: ThS Nguyễn Đình Trần Long

Abstract: Studies on HRM field at contrasting levels have gotten much curiosity for researchers recently because of the enlargement of the society, industry, and contemporary proceeding that enable a grouping to maintain and appropriate the Human resource productively. Because the 21st century has considered a incredible development in information technology (IT) especially the application of IT to the aspect of human life where the services provider were expanding. It is essential for the institution to retain and manage their employees and to improve their aggressive improvement. Hence, this paper that was undertaken to identify the impact and benefit when manage human resource through website. The information related to human resource was collected from Q-cash Company. Finally, the website was done using PHP language to code function. The experimental results demonstrate that the use of website to manage personnel is absolutely necessary, easy, proactive and economical

Keywords. Customer, Human Resource Management (HRM), Framework, Functions Problem, QCASH, Timesheets.

I. LITERATURE REVIEW

To all the organizations, their employees value more than anything does, because people working in a company or for a business decide its fate. Their productivity, passion, and determination are essential for the success of any organization and you cannot manage your workforce without a Human Resource Management Department (HRM). And that's reason why my group has decision on building a website to manage better.

Obviously, HRM brings to enterprise various benefits. In terms of economic benefits, it easily manage factors, components from general things to details, manager can manage and find information staff easily, cut down staffs who are unnecessary and save time and cost and find down new methods and ideas for working thanks to salary tables. With regards to social ones, HRM grasps the general situation of staff to have decision for awarding and punishment, makes everything clear, easily restore information and find them if necessary. To build it, we use PHP Language for this project. As we know, PHP is common one which applied popularly in website project because of simple algorithms and syntax. Different with other languages, PHP is highly recommended in website application like HRM Project.

1. Problems Definition

1.1 Introduction about the company

Qcash Vietnam Co., Ltd is a company operating in the field of financial information services, financial technology, specializing in the development of financial science and

technology and user risk control; create and develop APPs about financial technology; provide intelligent financial services and the most effective solutions to users. Qcash currently operates businesses in Beijing, Indonesia, Philippines, Vietnam, India and other countries.

The location is Tower A, Sky City, 88 Lang Ha, Ba Dinh District, Ha Noi City. This company is set up about 5 years, with the root company is at Singapore and a branch in Vietnam. The manager is Chinese, sub-manager is Vietnamese, and the legal representative is ZHAI BING TAO. The total of staff has about 30- 40 staffs, include two main departments: telesales and collect taxes.

Number of tax is 0108625977 and date begin working 1st March 2019

1.2 Language

In terms of language, we use PHP Language to create HRM website, and at the same time, the language used in website will be English. Because most of whom are foreigners and they all use English or Chinese to exchange information.

1.3 Sources for coding

Actually, to create a perfect website, The Internet will be useful and abundant tool to find sources, so what is source here? Obviously, they are documents related to the instructions to consult and implement code lines. Besides the knowledge we learned from university, specialty courses, online documents helps us not only how to implement command, code lines but also how to fix and debug. In other words, there are advanced functions in HRM Project which requires many supported tools. To be specific, to create an alert when users log in successfully, we used some template for this alerts [1]. Another examples, when upload images or send email functions, the sources used in websites are [2], [3]. The most important thing is perhaps User Interface. If you are IT-er, make sure that no one is unknown Bootstrap, a valuable and salutary library to boost the beauty of interface [4], [5]. So, perfect website brings the satisfactions to customer!

II. PROBLEMS OF CUSTOMERS

Problems of managing staff are popular for many companies and business. The companies that can efficiently utilize their human resources will become superior to their rivals. The employees, who have been considered as a cost component in previous practices, have become the most valuable asset of the organizations, and every employee started to be considered as a talent creating value for the organization.

Specifically, Q-cash company has still manage staff by hand. With a large number of staffs, controlling them manually make a difficulty for the company. Excel can help them to manage staffs but only with a small number of staffs. Therefore, they need a solution for this problem. Moreover, the company wants an app that has more convenient functions. Following the requirement of Q-cash, we have a plan to build a website that have many functions and specially, it can solve almost the difficulties.

Furthermore, the web has strengths that it works online, it's more flexible for reaction of users. Manager can manage staff from distance or restore data in an easy way,

1. Inovation and expected results

Through those present problems of company, our group made decision to erect Human Resource Web Application with basic and advanced functions. Specifically, instead of using traditional timekeeping in Excel, staff can access to website and then click on Start / Stop working status, the system will immediately calculate the total time in / out in your day-working. Moreover, you can send a request for approval such as off-working, personal reasons, family matter, and manager will response them after your request transferred.

Specifically, the system allows staff and manager not only using online signature to sign related contracts but also role-based control permission for each account. For example, this staff can use his/her account to view/edit/delete some news or documents, but others may not. I think it is plus point for system.

Perhaps, the biggest expected results from this web is satisfaction from customers side, which will make the success for our whole project. I think if the proposed system run softly, it will help us to gain certain experience and motivation to carry out other related project. Especially, it will be new chance for our group to upgrade advanced functions, create creative and new ways to deal with new challenges in the future.

III. OVERVIEW OF THE SYSTEM

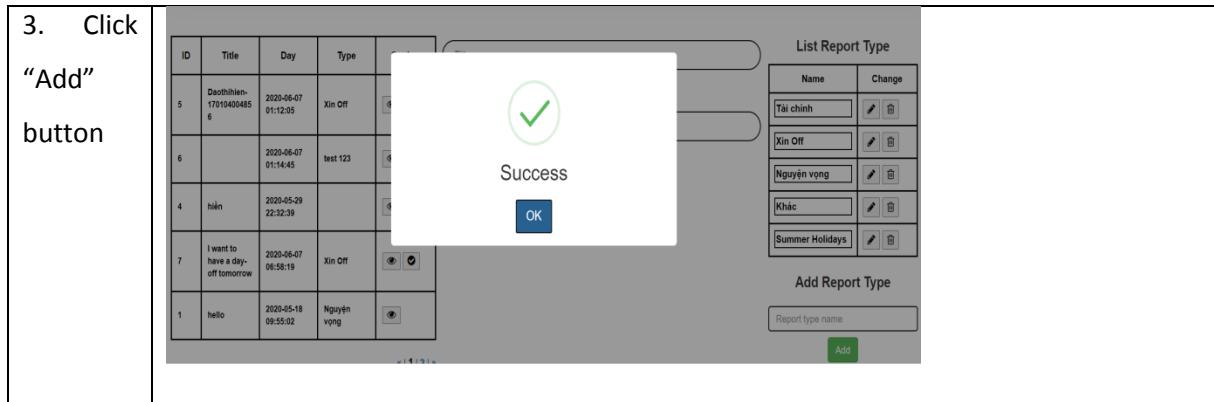
In this part, we focus on each step in each function that will make clearly how to use the website. We divided into 2 main parts: Staff and Admin, because each has different works, therefore, the following detail is regarded as depiction.

Staff	Admin
1. Log in/out 2. Registration 3. View Report 4. View News	5. Check Registration 6. Manage staff 7. Manage Attendance 8. Manage Department 9. Manage Profile of Staff 10. Manage Report 11. Manage Contract 12. Manage Contract Type 13. Manage Gross Salary 14. Manage News 15. Signature 16. Sign Contract

- | | |
|--|---|
| | <p>17. Manage Training Program
 18. Manage Staff in Training Program
 19. Manage Permission
 20. Evaluate and Classify Staff
 21. List statistics</p> |
|--|---|

1.4 Manage Report (Admin)

Steps	The display of website														
1. Admin add/edit/delete report type (if wants)	<p>List Report Type</p> <table border="1"> <thead> <tr> <th>Name</th> <th>Change</th> </tr> </thead> <tbody> <tr> <td></td> <td> </td> </tr> <tr> <td>Tài chính</td> <td> </td> </tr> <tr> <td>Xin Off</td> <td> </td> </tr> <tr> <td>Nguyễn vọng</td> <td> </td> </tr> <tr> <td>Khác</td> <td> </td> </tr> <tr> <td></td> <td> </td> </tr> </tbody> </table> <p>Add Report Type</p> <p>Report type name</p> <p>Add</p>	Name	Change			Tài chính		Xin Off		Nguyễn vọng		Khác			
Name	Change														
Tài chính															
Xin Off															
Nguyễn vọng															
Khác															
2. Input information in Report type name field	<p>List Report Type</p> <table border="1"> <thead> <tr> <th>Name</th> <th>Change</th> </tr> </thead> <tbody> <tr> <td>Tài chính</td> <td> </td> </tr> <tr> <td>Xin Off</td> <td> </td> </tr> <tr> <td>Nguyễn vọng</td> <td> </td> </tr> <tr> <td>Khác</td> <td> </td> </tr> </tbody> </table> <p>Add Report Type</p> <p>Summer Holidays</p> <p>Add</p>	Name	Change	Tài chính		Xin Off		Nguyễn vọng		Khác					
Name	Change														
Tài chính															
Xin Off															
Nguyễn vọng															
Khác															



1.5 View Report

3. After add, admin will receive your report	<p>Xin Off</p> <table border="1"> <thead> <tr> <th>ID</th><th>Title</th><th>Day</th><th>Type</th><th>Setting</th></tr> </thead> <tbody> <tr> <td>5</td><td>Daothihien-170104004856</td><td>2020-06-07 01:12:05</td><td>Xin Off</td><td><input type="checkbox"/> <input checked="" type="checkbox"/></td></tr> <tr> <td>6</td><td></td><td>2020-06-07 01:14:45</td><td>test 123</td><td><input type="checkbox"/> <input checked="" type="checkbox"/></td></tr> <tr> <td>7</td><td>I want to have a day-off tomorrow</td><td>2020-06-07 06:58:19</td><td>Xin Off</td><td><input type="checkbox"/> <input checked="" type="checkbox"/></td></tr> <tr> <td>4</td><td>hiền</td><td>2020-05-29 22:32:39</td><td></td><td><input type="checkbox"/> <input checked="" type="checkbox"/></td></tr> <tr> <td>1</td><td>hello</td><td>2020-05-18 09:55:02</td><td>Nguyễn vọng</td><td><input type="checkbox"/></td></tr> </tbody> </table> <p>I want to have a day-off tomorrow Because today I am sick File Attachment</p> <p>ADD MESSEAGE</p>	ID	Title	Day	Type	Setting	5	Daothihien-170104004856	2020-06-07 01:12:05	Xin Off	<input type="checkbox"/> <input checked="" type="checkbox"/>	6		2020-06-07 01:14:45	test 123	<input type="checkbox"/> <input checked="" type="checkbox"/>	7	I want to have a day-off tomorrow	2020-06-07 06:58:19	Xin Off	<input type="checkbox"/> <input checked="" type="checkbox"/>	4	hiền	2020-05-29 22:32:39		<input type="checkbox"/> <input checked="" type="checkbox"/>	1	hello	2020-05-18 09:55:02	Nguyễn vọng	<input type="checkbox"/>
ID	Title	Day	Type	Setting																											
5	Daothihien-170104004856	2020-06-07 01:12:05	Xin Off	<input type="checkbox"/> <input checked="" type="checkbox"/>																											
6		2020-06-07 01:14:45	test 123	<input type="checkbox"/> <input checked="" type="checkbox"/>																											
7	I want to have a day-off tomorrow	2020-06-07 06:58:19	Xin Off	<input type="checkbox"/> <input checked="" type="checkbox"/>																											
4	hiền	2020-05-29 22:32:39		<input type="checkbox"/> <input checked="" type="checkbox"/>																											
1	hello	2020-05-18 09:55:02	Nguyễn vọng	<input type="checkbox"/>																											
4. After then, admin can click on "Confirm" icon or add message to reply	<p>Xin Off</p> <table border="1"> <thead> <tr> <th>ID</th><th>Title</th><th>Day</th><th>Type</th><th>Setting</th></tr> </thead> <tbody> <tr> <td>5</td><td>Daothihien-170104004856</td><td>2020-06-07 01:12:05</td><td>Xin Off</td><td><input type="checkbox"/> <input checked="" type="checkbox"/></td></tr> <tr> <td>6</td><td></td><td>2020-06-07 01:14:45</td><td>test 123</td><td><input type="checkbox"/> <input checked="" type="checkbox"/></td></tr> <tr> <td>7</td><td>I want to have a day-off tomorrow</td><td>2020-06-07 06:58:19</td><td>Xin Off</td><td><input type="checkbox"/> <input checked="" type="checkbox"/></td></tr> <tr> <td>4</td><td>hiền</td><td>2020-05-29 22:32:39</td><td></td><td><input type="checkbox"/> <input checked="" type="checkbox"/></td></tr> <tr> <td>1</td><td>hello</td><td>2020-05-18 09:55:02</td><td>Nguyễn vọng</td><td><input type="checkbox"/></td></tr> </tbody> </table> <p>I want to have a day-off tomorrow Because today I am sick File Attachment</p> <p>Yes, I totally agree. Get well soon!</p> <p>ADD MESSEAGE</p> <p style="text-align: center;">« 1 2 »</p>	ID	Title	Day	Type	Setting	5	Daothihien-170104004856	2020-06-07 01:12:05	Xin Off	<input type="checkbox"/> <input checked="" type="checkbox"/>	6		2020-06-07 01:14:45	test 123	<input type="checkbox"/> <input checked="" type="checkbox"/>	7	I want to have a day-off tomorrow	2020-06-07 06:58:19	Xin Off	<input type="checkbox"/> <input checked="" type="checkbox"/>	4	hiền	2020-05-29 22:32:39		<input type="checkbox"/> <input checked="" type="checkbox"/>	1	hello	2020-05-18 09:55:02	Nguyễn vọng	<input type="checkbox"/>
ID	Title	Day	Type	Setting																											
5	Daothihien-170104004856	2020-06-07 01:12:05	Xin Off	<input type="checkbox"/> <input checked="" type="checkbox"/>																											
6		2020-06-07 01:14:45	test 123	<input type="checkbox"/> <input checked="" type="checkbox"/>																											
7	I want to have a day-off tomorrow	2020-06-07 06:58:19	Xin Off	<input type="checkbox"/> <input checked="" type="checkbox"/>																											
4	hiền	2020-05-29 22:32:39		<input type="checkbox"/> <input checked="" type="checkbox"/>																											
1	hello	2020-05-18 09:55:02	Nguyễn vọng	<input type="checkbox"/>																											

1.6 News

Steps	The display of website
1. Admin click "Manage News" button to add/edit/delete news	<p>Off-week because COVID -19 (From 12/3 - 13/3)</p> <p>PHÒNG CHỐNG DỊCH COVID-19</p> <p>5 ĐIỂM CẦN LÀM TỐT</p> <ul style="list-style-type: none"> 01 KHÁM CHÉT BÌ LẠU Viết chữ số 01 là để xác định số thứ tự của tin tức 02 GIỮA NGHỈ HỌC Hãy luôn giữ khoảng trống, tránh chen nhau trong phòng thông gió, tối thiểu là 2m 03 ỦA TỰ THÔNG TIN Thường xuyên cài lên laptop và điện thoại ứng dụng trực tuyến để cập nhận tin tức 04 Ủ CƠM VỚI CÁ Thường xuyên ăn cơm và đồ uống nóng để tăng cường tăng cường sức khỏe 05 KHOA BẢO Y TẾ Thực hiện khai báo y tế hàng ngày qua ứng dụng Zalo, Viber hoặc điện thoại <p>SỨC KHỎE ĐỒNG Để được tư vấn chi tiết vui lòng liên t่อ 1900 9095 1900 3228</p> <p>News Details</p>

<p>2. Staff go to the website and click on “News”, where all news related to hiring, off-days, off-public holidays...</p>	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th colspan="4" style="text-align: center; color: #800080;">News</th></tr> </thead> <tbody> <tr> <td style="padding: 5px;">Off-week because COVID -19 (From 12/3 - 13/3)</td><td style="padding: 5px; text-align: right;">2020-06-07 00:38:39</td><td style="padding: 5px;">admin</td><td style="padding: 5px; text-align: center;"></td></tr> <tr> <td style="padding: 5px;">pilot test</td><td style="padding: 5px; text-align: right;">2020-06-07 00:34:10</td><td style="padding: 5px;">admin</td><td style="padding: 5px; text-align: center;"></td></tr> <tr> <td style="padding: 5px;">test 1</td><td style="padding: 5px; text-align: right;">2020-06-07 00:09:19</td><td style="padding: 5px;">admin</td><td style="padding: 5px; text-align: center;"></td></tr> <tr> <td style="padding: 5px;">test 1</td><td style="padding: 5px; text-align: right;">2020-05-29 22:27:52</td><td style="padding: 5px;">admin</td><td style="padding: 5px; text-align: center;"></td></tr> <tr> <td style="padding: 5px;">Test 2</td><td style="padding: 5px; text-align: right;">2020-05-14 09:21:17</td><td style="padding: 5px;">admin</td><td style="padding: 5px; text-align: center;"></td></tr> <tr> <td style="padding: 5px;">Công ty mới chi nhánh mới</td><td style="padding: 5px; text-align: right;">2020-05-10 15:22:40</td><td style="padding: 5px;">admin</td><td style="padding: 5px; text-align: center;"></td></tr> <tr> <td colspan="4" style="text-align: center; padding-top: 5px;">« 1 »</td></tr> </tbody> </table>	News				Off-week because COVID -19 (From 12/3 - 13/3)	2020-06-07 00:38:39	admin		pilot test	2020-06-07 00:34:10	admin		test 1	2020-06-07 00:09:19	admin		test 1	2020-05-29 22:27:52	admin		Test 2	2020-05-14 09:21:17	admin		Công ty mới chi nhánh mới	2020-05-10 15:22:40	admin		« 1 »			
News																																	
Off-week because COVID -19 (From 12/3 - 13/3)	2020-06-07 00:38:39	admin																															
pilot test	2020-06-07 00:34:10	admin																															
test 1	2020-06-07 00:09:19	admin																															
test 1	2020-05-29 22:27:52	admin																															
Test 2	2020-05-14 09:21:17	admin																															
Công ty mới chi nhánh mới	2020-05-10 15:22:40	admin																															
« 1 »																																	
<p>2. Staff can click on “View” button to view detail</p>																																	

1.7 Manage Department

Steps	The display of website
<p>1. Admin log in with admin account and add / edit/ delete department</p>	<p>The screenshot shows the Admin interface with the following sections:</p> <ul style="list-style-type: none"> User Management: A table titled "List User with Department" showing users like admin, admin1, admin2, admin3, and admin5, each assigned to a specific department (Staff, Manager, None, Director). Department Management: A table titled "List Department" showing departments: None, Staff, Manager, and Director, each with edit and delete icons. Modified Department: A form to add a new department with fields for "Department name" and a "Add" button. Set Department to user: A form to assign a department to a user, with dropdowns for "Select user" and "Select Department" and an "Add" button.

2. Admin input all fields and click on “Add” button	<p style="text-align: center;">List Department</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">Name</th><th style="text-align: center;">Change</th></tr> </thead> <tbody> <tr><td>None</td><td style="text-align: center;"> </td></tr> <tr><td>Staff</td><td style="text-align: center;"> </td></tr> <tr><td>Manager</td><td style="text-align: center;"> </td></tr> <tr><td>Director</td><td style="text-align: center;"> </td></tr> </tbody> </table> <p style="margin-top: 10px;">Modified Department</p> <p style="margin-top: 10px;">Set Department to user</p>	Name	Change	None		Staff		Manager		Director	
Name	Change										
None											
Staff											
Manager											
Director											
3. Admin set Department for staff	<p style="text-align: center;">List Department</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">Name</th><th style="text-align: center;">Change</th></tr> </thead> <tbody> <tr><td>None</td><td style="text-align: center;"> </td></tr> <tr><td>Staff</td><td style="text-align: center;"> </td></tr> <tr><td>Manager</td><td style="text-align: center;"> </td></tr> <tr><td>Director</td><td style="text-align: center;"> </td></tr> </tbody> </table> <p style="margin-top: 10px;">Modified Department</p> <p style="margin-top: 10px;">Set Department to user</p>	Name	Change	None		Staff		Manager		Director	
Name	Change										
None											
Staff											
Manager											
Director											

IV. CONCLUSION

Building advanced functions is one of the most important aspects of Human Resources Management Project. In this paper, we proposed an efficient approach for solving the website by Agile model. We interacted and worked face-to-face to customers to get requirement about the project. After many meetings and feedbacks from customer during the process of implementation, we consistently update and edit requirements that get from them. Then, we conducted building this website to show that the necessary of management online human resources based on Internet. Both director and Human Resources Manager can use it to select appropriate strategies for company and boost benefits in the future.

REFERENCES

- [1] Jorge A. Ruiz-Vanoye, et al, *Editorial for Volume 9 Number 3: Models, Resources and Activities of Project Scheduling Problems*, International Journal of Combinatorial Optimization Problems and Informatics, Vol. 9, No. 3, Sep-Dec 2018, pp. 1-17. ISSN: 2007-1558.
- [2] Nestor Raul Ortiz-Pimiento, Francisco Javier Diaz-Serna, *The Project Scheduling Problem with Non-Deterministic Activities Duration: A Literature Review*, JIEM, 2018 - 11(1): 116-134 - Online ISSN: 2013-0953 - Print ISSN: 2013-8423
<https://doi.org/10.3926/jiem.2492>

- [3] Icmeli, Oya., *Scheduling problems in project management*, University of Florida, George A. Smathers Libraries, 1992, pp.7-22.
- [4] Dayanand, N. and Padman, R. *Payments in projects: a contractor's model*, Technical Report, The Heinz School, Carnegie Mellon University, Pittsburgh 1993
- [5] Richard C. Grinold, *The payment scheduling problem*, Volume 19, Issue 1, <https://doi.org/10.1002/nav.3800190110>
- [6] Kazaz, B. and Sepil, C.B. *Project scheduling with discounted cash flows and progress the payments*, Journal of Operational Research Society, 47 (1996), 1262-1272
- [7] Bao Ngoc Trinh, Quyet Thang Huynh, Xuan Thang Nguyen, *Nash Equilibrium model for conflicts in project management*, Journal of Computer Science and Cybernetics, V.34, N.3 (2019), 1– DOI 10.15625/1813-9663/34/3/13095.
- [8] Deng Ze-min, Gao Chun-ping, Li Zhong-xue, *Optimization of project payment schedules with Nash equilibrium model and genetic algorithm* [J]. J Chongqing Univ: Eng Ed (ISSN 1671-8224), 2007, 6 (2):107-112.
- [9] José Ramón San Cristóbal, *The use of Game Theory to solve conflicts in the project management and construction industry*, International Journal of Information Systems and Project Management
- [10] Piotr Skowron, Krzysztof Rzadca, *Non-monetary fair scheduling — cooperative game theory approach*, in Annual ACM symposium on parallelism in algorithms and architectures, pp 288–297, 2013
- [11] George Christodoulou et al., *A lower bound for scheduling mechanisms*, Article in Algorithmica 55, December 2009
- [12] Cyril Briand, Jean-Charles Billaut, *Cooperative project scheduling with controllable processing times: a game theory framework*, IEEE 5-9 Sept. 2011
- [13] Pradeep Varakantham, Na Fu, *Mechanism Design for Strategic Project Scheduling*, in Twenty-Sixth International Joint Conference on Artificial Intelligence, August 2017
- [14] Andres Abeliuk et al., *Interdependent Scheduling Games*, Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence
- [15] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, *A fast and elitist multiobjective genetic algorithm: NSGA-II*, IEEE Transactions on Evolutionary Computation, vol. 6, no. 2, pp. 182–197, 2002.
- [16] Miqing Li, Shengxiang Yang, Xiaohui Liu, Kang Wang, *IPESA-II: Improved Pareto Envelope-Based Selection Algorithm II*, EMO 2013: Evolutionary Multi-Criterion Optimization, pp143-155

[17] L.C. Jiao, Handing Wang, R.H. Shang, F. Liu, *A Co-evolutionary Multi-Objective Optimization Algorithm Based on Direction Vectors*, Information Sciences 228:90 – 112, April 2013, DOI: 10.1016/j.ins.2012.12.013

[18] MOEA framework, A Java library for multi-objective evolutionary algorithms, <http://www.moeaframework.org/>, accessed: March 10, 2-15.

[19] DTIC ADA615240: Tactical AI in Real Time Strategy Games, Defense Technical Information Center, 2015-03-26.

[20] Deb, K. et al., *A Fast Multi-Objective Evolutionary Algorithm for Finding Well-Spread Pareto-Optimal Solutions*, KanGAL Report No 2003002, Feb 2003.

[21] Chih-Hao Lin and Pei-Ling Lin (2010), *A New Non-dominated Sorting Genetic Algorithm for Multi-Objective Optimization*, Modeling Simulation and Optimization Focus on Applications, Shkelzen Cakaj (Ed.), ISBN: 978- 953-307-055-1.

[22] Lin, C.H. & Chuang, C.C. (2007), *A rough set penalty function for marriage selection in multiple-evaluation genetic algorithms*, Lecture Notes in Computer Science, Vol. 4481, pp. 500-507.

[23] Kukkonen and Lampinen (2005), *GDE3: The Third Evolution Step of Generalized Differential Evolution*, KanGAL Report Number 2005013.

[24] David W. Corne, Nick R. Jerram, Joshua D. Knowles, Martin J. Oates (2001), *PESA-II: Region-based Selection in Evolutionary Multi-objective Optimization*, In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'2001), Pages 283-290, San Francisco, California — July 07 - 11, 2001, ISBN:1-55860-774-9

[25] A.J. Nebro, J.J. Durillo, J. García-Nieto, C.A. Coello Coello, F. Luna, E. Alba (2009), *SMPSO: A New PSO Metaheuristic for Multi-objective Optimization*. Conference: Computational intelligence in multi-criteria decision-making, DOI: 10.1109/MCDM.2009.4938830

[26] A Guide to the Project Management Body of Knowledge (PMBOK), *Project Management Institute*, 1996.

[27] SweetAlert for Bootstrap [online] Available at: https://lipis.github.io/bootstrap-sweetalert/?fbclid=IwAR3ukh4hqGDrcb5TOYjbNTPXV6kz8YOHeUxo6_vQVyKG0bsMAWDfwSuBn8w. [Accessed 16 May 2020]

DEVELOPING WEB APPLICATIONS USING MERN STACK

SVTH: Nguyen Hai Dang, Nguyen Hoang, Tran Van Khang,

Ngo Quoc Trung, Vu Thanh Trung

GVHD: ThS Nguyen Dinh Tran Long

Abstract: Thanks to the evolution of the Internet and World Wide Web, webpages at first basically are hypertext media that have been formatted in Hypertext Markup Language and displayed in text, audio, video and other components . Nowadays, new technologies help web developers to build, manage, maintain web applications a lot easier, more efficient with better performances. In this paper, we will discuss about the technique, reasons to use the stack, the way it works and our result using the technique.

Keywords – ExpressJS, MERN stack, MongoDB, NodeJS, ReactJs, Web applications.

I. INTRODUCTION

In the early day of the Internet, website are developed using a combination of basic HTML, CSS elements along with some Javascipt to handle the logic. However, as our society progress, the plain old websites of the early days were no longer suffice to satisfy the increasing needs of mankind. Websites not only have to have a user friendly interface but also need to timely handle user requests while also maintain a database to store and retrieve all the necessary data for the site. As the effort the for the web development became more and more absurb to accompany the scope, many techniques and approaches were introduced and widely addopted to not only handle most of the tedious process of development but also allow the application to be highly efficient and can be scaled appropriately on demand. One of the most popular approach to web development is the stack, which is collection of software and tools that allow the developers to deal with the front-end (i.e User Interface) and the back-end (i.e Server, Database,...) to their website.

II. MAIN CONTENT

A.What is MERN stack?

A web application is built from multiple different technologies and the combination of these technologies is called a “stack” such as LAMP stack, which stands for Linux, Apache, MySQL, and PHP. Single page application (SPA) is getting more and more popularr due to the avoidance of refreshing a web page to display new screen with new data. Therefore, front-end frameworks and NoSQL databases also getting more attentions from the developers.

MERN stack is a new open-source stack that “M” is MongoDB, a NoSQL Database, also known as a cross-platform document oriented database, where each record is a document including key-value pairs that are related to JSON (JavaScript Object Notation) objects. “E”

stands for ExpressJS, a web server built on Node.js. “R” is ReactJS, a JavaScript library developed by Facebook that is used for building user interfaces of SPAs and mobile applications. It allows users to code in JavaScript programming language and create UI (User interface) components. And the last one, “N”, is Node.js which provides a JavaScript Runtime Environment and allows user to run code on server. Node.js with Node package manager lets user to download thousands of free packages (node modules) for their project.

B. Why use MERN stack ?

MERN stack gains popularity from developers and architects by using four powerful technologies which provide ability to build a full stack application with front-end, back-end, database, mobile application and testing. Project manager can save budget by hiring a full stack developer, and the full stack developer may have ability to take control of design and implementation and deliver a complete solution to problems and many more of benefits.

The stack can bring the ability of easier and faster deployment of full-stack web applications thanks to the advantage from each technology of MERN.

Initially, MongoDB is a fast database due to its easy documents index, has ability to handle large data and divide into machines, stores data in form of JSON. Because of schema-less, any data type in a separate document, no dependency. Besides, MongoDB is also easy to set up environment and supports document-model which is faster and easier.

Secondly, ExpressJS makes the way to write code for back-end server easier by supporting middlewares. It also has a big community, biggest for Node.js, and reusability with the built-in router, delivers robust API.

Thirdly, ReactJS supports JSX – JavaScript XML which helps developers easier to get familiar with write React components code due to HTML/XML JavaScript Extension. React has the reusability by supporting Components where component code has logic and contributes to the overall UI. Front-end framework from Facebook also has high performance when Virtual DOM, JSX and Components makes it faster than the rest of frameworks. ReactJS with React Native helps developers to code mobile applications easier with the base knowledge of JavaScript and React.

And finally, Node.js, the environment of application, is an open source JavaScript Runtime Environment with high scalability, fast code execution due to being built on Google Chrome’s JavaScript Engine and data streaming, follows single threaded model.

C. The way MERN stack works:

MERN stack is a full stack modern JavaScript solution to dynamic web applications. Building a web application using one programming language is a great attraction to development teams and corporations. It's also included heavily vetted components with large open source communities. The diagram below shows how the stack basically works.

Figure 1 – Application Structure

```
const MongoClient = require('mongodb').MongoClient;
const assert = require('assert');

// Connection URL
const url = 'mongodb://localhost:27017';

// Database Name
const dbName = 'myproject';

// Use connect method to connect to the server
MongoClient.connect(url, function(err, client) {
  assert.equal(null, err);
  console.log("Connected successfully to server");

  const db = client.db(dbName);

  client.close();
});
```

MongoDB, the database server, connected to the back-end server through a driver.

Figure 2 – MongoDB Node.js driver

In the example, the URI connection string specifies connecting to a MongoDB instance in Node.js on *localhost* using port 27017 (Default MongoDB port). Then it connects to the database with the name *dbName*.

Next, web back-end framework for Node.js is Express which is used to start a server connecting with the database and running API controller so the front-end framework can get data to display on the client-side.

```
const express = require('express')
const app = express()
```

Figure 3 – Creates an Express Application

The API is including CRUD – Create, Read, Update, Delete, also referred to as a RESTful API that uses HTTP requests to GET, PUT, POST and DELETE data.

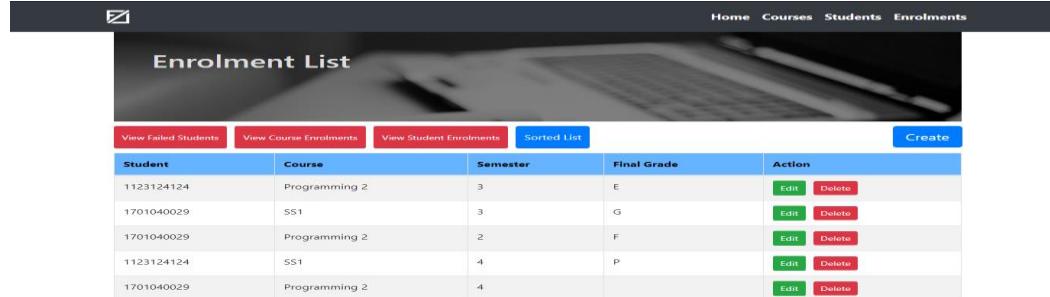
Resource	POST create	GET read	PUT update	DELETE delete
/dogs	Create a new dog	List dogs	Bulk update dogs	Delete all dogs
/dogs/1234	Error	Show Bo	If exists update Bo If not error	Delete Bo

Figure 4 – RESTful API

And when the API controller is ready to request, React needs to send request to the back-end server to get data for the server rendering.

Without having to deal with UI complexity, React uses Component as a simple state machine. Whenever the *state* changes, screen re-renders the component and the view automatically changes without reloading page. The following is an example for changing state by methods without reloading page.

Figure 5 – First State



The screenshot shows a web application interface titled "Enrolment List". At the top, there is a navigation bar with links: "Home", "Courses", "Students", "Enrolments", and a "Create" button. Below the navigation bar is a table with the following columns: "Student", "Course", "Semester", "Final Grade", and "Action". The table contains five rows of data:

Student	Course	Semester	Final Grade	Action
1123124124	Programming 2	3	E	<button>Edit</button> <button>Delete</button>
1701040029	SS1	3	G	<button>Edit</button> <button>Delete</button>
1701040029	Programming 2	2	F	<button>Edit</button> <button>Delete</button>
1123124124	SS1	4	P	<button>Edit</button> <button>Delete</button>
1701040029	Programming 2	4		<button>Edit</button> <button>Delete</button>

*Figure 5 –
First State*

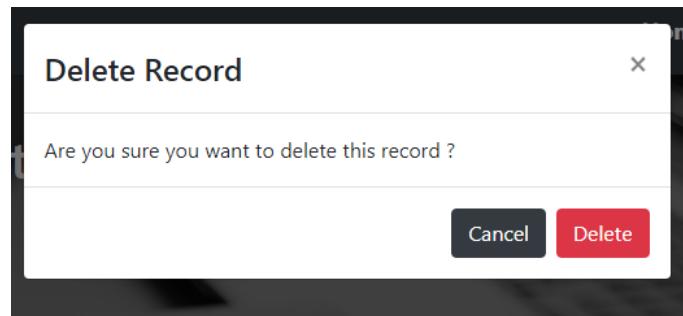
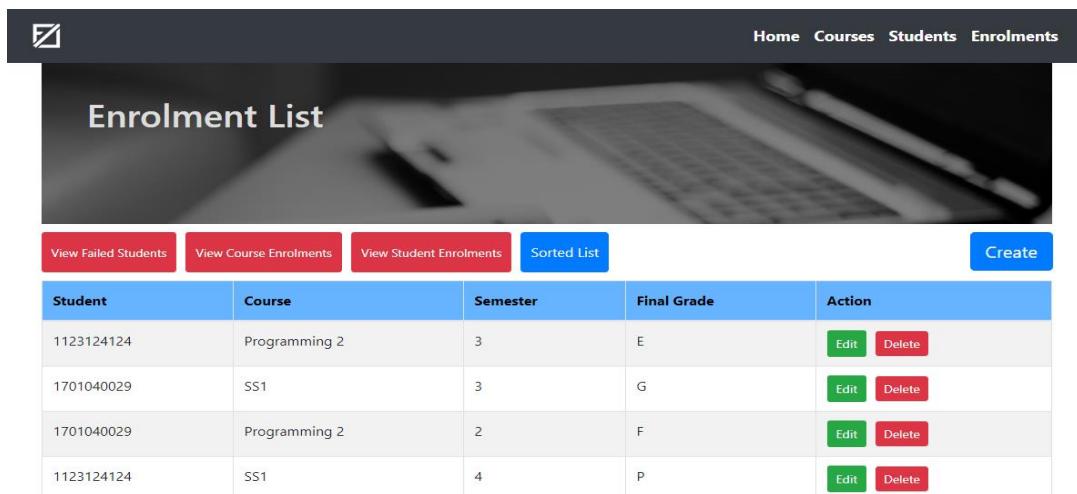


Figure 6 – Trigger Delete Method

At this example, first state of the component, the view displays a table with 5 records. Then triggering the Delete method of a row (a record), the view immediately deletes the record without reloading page if there is no error.



The screenshot shows the same "Enrolment List" page as Figure 5, but now it only displays four rows of data, indicating that one record has been deleted. The table structure is identical to Figure 5.

Student	Course	Semester	Final Grade	Action
1123124124	Programming 2	3	E	<button>Edit</button> <button>Delete</button>
1701040029	SS1	3	G	<button>Edit</button> <button>Delete</button>
1701040029	Programming 2	2	F	<button>Edit</button> <button>Delete</button>
1123124124	SS1	4	P	<button>Edit</button> <button>Delete</button>

Figure 7 – Result

D.Simple web application – Result of work:

In this section, we will briefly describe our simple web application using MERN stack to work. The application is for student-course management.

Firstly, for the database, we designed it with 3 collections (like 3 tables in SQL) courses, students and enrolments. Due to focusing on the main functions of the application is the CRUD for Course – Student – Enrolment and no authentication, there is only three collections.

Figure 8 – Database

Collection Name	Documents	Avg. Document Size	Total Document Size	Num. Indexes	Total Index Size	Properties
courses	3	98.0 B	294.0 B	2	73.7 KB	
enrolments	4	121.0 B	484.0 B	1	36.9 KB	
students	3	141.3 B	424.0 B	2	73.7 KB	

Students and Courses collection can work independently, however, for the Enrolment collection, it has a relationship with both two collections because an enrolment record is for a student to enrol in a course so only when student and course existed in the database, the enrolment is valid.

The image below is the very first output when running Course Management App. A homepage which display a slide show of picture related to university. The section above is a navigation bar which got three links Home, Course, Student, Enrolment corresponding to each management screen.

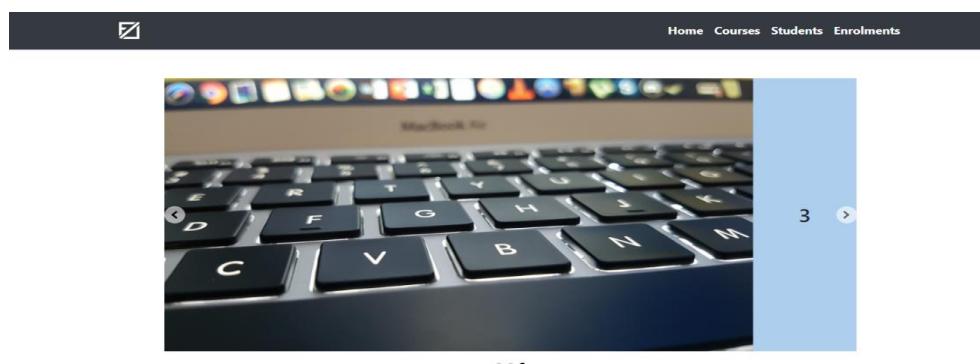


Figure 9 – Homepage

To access course management, we click on the item named **Courses**, the browser will display a course list screen which contain all exist courses with 3 attributes based on the

requirement in the database. Besides, there are also a button **Create** to make a new course record and 2 actions which are **Delete** and **Edit** for delete a exist record or edit it respectively.

Figure 10 – Course View

The screenshot shows a web application interface titled "Course List". At the top right, there are navigation links: Home, Courses, Students, and Enrolments. Below the title, there is a table with the following data:

Course ID	Course Name	Prerequisites	Actions
IPG	Nhập môn lập trình	-	Edit Delete
PPL	Lập trình II	IPG	Edit Delete
SEG	Kỹ nghệ phần mềm	PPL	Edit Delete
JAV	Công nghệ JAV	SEG	Edit Delete

A blue "Create" button is located at the top right of the table area. A small number "1" is visible in the bottom right corner of the table.

Click on button **Create**, a create new course screen will show up. In this case, we will input invalid input and it will display an error for each input for user to notify then inputting and correcting the form for another submit.

Figure 11 – Create Course View with validation

The screenshot shows a "Create Course" form. The fields are as follows:

- Course Id ***: Input field contains "S", with an error message: "Please provide correct course id (Length: 3-10)".
- Course Name ***: Input field contains "Course Name", with an error message: "Please provide correct course name (Length: 2-30)".
- Prerequisites**: Input field contains "Prerequisites".

A large red "Create Course" button is at the bottom.

And if we input the valid form, browser will rollback to course list screen and the new one record will appear at the bottom of the datatable.

Now move back to the course list, if we click on the button **Delete** there will be an modal confirmation to make sure that user truly want to delete that record as image below. If user click **Delete**, the record will be removed otherwise **Cancel**, the modal will be closed back to course list.

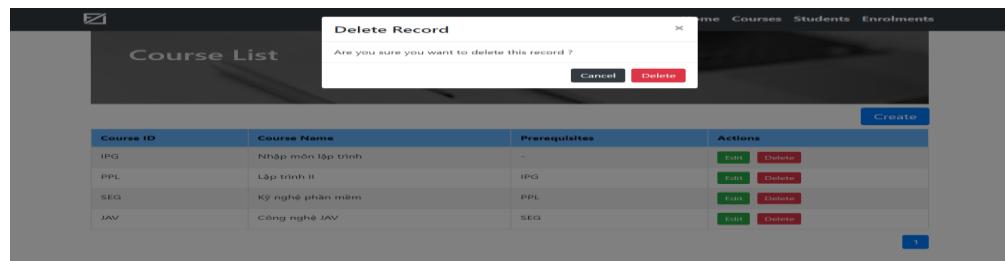


Figure 12 – Delete Method

To edit a course, user have to click on button **Edit** in the corresponding Actions row of the record. Follow the requirement, the *courseID* input is disabled.

Course Id *

PPL

Course Name *

Lập trình II

Prerequisites

IPG

Update Course

Figure 13 – Update Course View

The Student and Course have the similar methods and views so we will show the Enrolment part instead of Student. Enrolment View with List of Enrolments, also has Create, Delete and Edit methods like the others. However, it also has other functions for the reports, View Failed Students, Search for Enrolments of a specific student and Search for students enrolled in a specific course.

Figure 14 – Enrolment View

View Failed Students

View Course Enrolments

View Student Enrolments

Sorted List

Create

Student	Course	Semester	Final Grade	Action
1701040001	Nhập môn lập trình	2	F	<button>Edit</button> <button>Delete</button>
1701040002	Lập trình II	2	P	<button>Edit</button> <button>Delete</button>
1701040004	Lập trình II	1	F	<button>Edit</button> <button>Delete</button>
1701040001	Kỹ nghệ phần mềm	4	E	<button>Edit</button> <button>Delete</button>
1701040004	Lập trình II	5	G	<button>Edit</button> <button>Delete</button>

To view all failed students, user just have to click on the button View Failed Students, a page contains information about all failed students will display.

The screenshot shows a table titled "Enrolment List" with columns: Student ID, Student Name, Course, Semester, and Final Grade. Two rows are visible: one for student ID 1701040001 (Name: Trần Minh Lộc, Course: Nhập môn lập trình, Semester: 2, Grade: F) and one for student ID 1701040004 (Name: Nguyễn Mỹ Hoa, Course: Lập trình II, Semester: 1, Grade: F). A blue button labeled "1" is located in the bottom right corner of the table.

Student ID	Student Name	Course	Semester	Final Grade
1701040001	Trần Minh Lộc	Nhập môn lập trình	2	F
1701040004	Nguyễn Mỹ Hoa	Lập trình II	1	F

Figure 15 – Failed Students View

To view courses enrolled by a specific student, user have to click on button View Student Enrolments, there will be a modal that contain an input for user to enter the student ID they want to see. When user click on the input, a list of existed students will be suggested for user.

The screenshot shows a modal window titled "Student Enrolments" with a search input for "Student ID". A dropdown menu lists suggested student IDs: 1701040001, 1701040002, 1701040003, and 1701040004. Below the modal is a table of student enrolments with columns: Student, Course, Semester, Final Grade, and Action (Edit, Delete). The table contains five rows of data.

Student	Course	Semester	Final Grade	Action
1701040001	Nhập môn lập trình	2	F	<button>Edit</button> <button>Delete</button>
1701040002	Lập trình II	2	P	<button>Edit</button> <button>Delete</button>
1701040004	Lập trình II	1	F	<button>Edit</button> <button>Delete</button>
1701040001	Kỹ nghệ phần mềm	4	E	<button>Edit</button> <button>Delete</button>
1701040004	Lập trình II	5	G	<button>Edit</button> <button>Delete</button>

Figure 16 – View Student's Enrolments

The screenshot shows a modal window titled "Student Enrolments" with a search input for "Student ID". The dropdown menu now shows the selected student ID: 1701040001. Below the modal is a table of student enrolments with columns: Student, Course, Semester, Final Grade, and Action (Edit, Delete). The table contains five rows of data.

Student	Course	Semester	Final Grade	Action
1701040001	Nhập môn lập trình	2	F	<button>Edit</button> <button>Delete</button>
1701040002	Lập trình II	2	P	<button>Edit</button> <button>Delete</button>
1701040004	Lập trình II	1	F	<button>Edit</button> <button>Delete</button>
1701040001	Kỹ nghệ phần mềm	4	E	<button>Edit</button> <button>Delete</button>
1701040004	Lập trình II	5	G	<button>Edit</button> <button>Delete</button>

Figure 17 – View Student's Enrolments – success

View Course Enrolments is the same as View Students Enrolments.

III. CONCLUSION

Over the years, Internet has shown its outstanding development and users have the opportunity to approach the modern applications, web applications with many powerful features for functions and graphical UI. Therefore, developers can create websites for many different purposes without limitation or obstacles due to technology. In this paper, we delivered a brief discussion of the MERN stack and a simple demo application about Student – Course management which follow this modern web application technology stack.

REFERENCES

- [9] Ganguly, D. S. (2019, February 26). Most Popular Technology Stack To Choose From: Full Stack Vs. MEAN Stack Vs. MERN Stack In 2019 [Web blog post]. Retrieved from <https://hackernoon.com/most-popular-technology-stack-to-choose-from-full-stack-vs-mean-stack-vs-mern-stack-in-2019-d12c0a17439a>
- [10] Patel, D. (2019) . *Switching to the Modern Day MERN stack!*. Retrieve from <https://medium.com/nybles/switching-to-the-modern-day-mern-stack-574bb478fc64>
- [11] Subramanian, V. (2017). *Pro MERN Stack*. doi: 10.1007/978-1-4842-2653-7

ARDUINO AND IOT APPLICATION USING ESP8266 Nodemcu

SVTH: Nguyễn Thị Ngà

GVHC: TS Tạ Quang Hùng

Abstract— *ESP8266 NodeMCU Lua is an RF-based SoC ESP8266 chip-based development kit with an easy-to-use design and especially can be used directly by the Arduino compiler to program and load code, which makes using and programming the application on ESP8266 is very simple. The Wifi RF ESP8266 NodeMCU Lua RF transceiver kit is used for applications that need to connect, collect data and control via Wifi waves, especially applications related to IoT.*

Keywords—*Embedded systems, microcontroller, ESP8266, IoT.*

I. INTRODUCTION

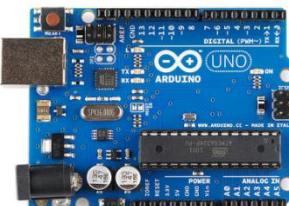
Arduino is an open-source platform used for building electronics projects. Arduino consists of both a physical programmable circuit board (often referred to as a microcontroller) and a piece of software, or IDE (Integrated Development Environment) that runs on your computer, used to write and upload computer code to the physical board.

The Arduino platform has become quite popular with people just starting out with electronics, and for good reason. Unlike most previous programmable circuit boards, the Arduino does not need a separate piece of hardware (called a programmer) in order to load new code onto the board -- you can simply use a USB cable. Additionally, the Arduino IDE uses a simplified version of C++, making it easier to learn to program. Finally, Arduino provides a standard form factor that breaks out the functions of the micro-controller into a more accessible package.

Arduino makes several different boards, each with different capabilities. In addition, part of being open source hardware means that others can modify and produce derivatives of Arduino boards that provide even more form factors and functionality. Here are a few options that are well-suited to someone new to the world of Arduino:

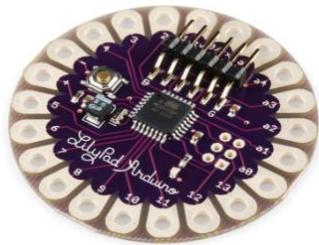
A. Arduino Uno (R3)

The Uno is a great choice for your first Arduino. It's got everything you need to get started, and nothing you don't. It has 14 digital input/output pins (of which 6 can be used as PWM outputs), 6 analog inputs, a USB connection, a power jack, a reset button and more. It contains everything needed to support the microcontroller; simply connect it to a computer with a USB cable or power it with a AC-to-DC adapter or battery to get started.



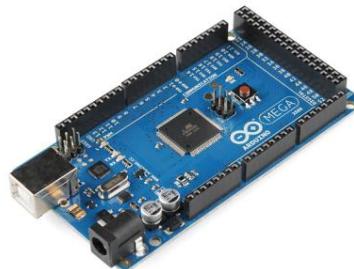
B. LilyPad Arduino

LilyPad is a wearable e-textile technology developed by Leah Buechley and cooperatively designed by Leah and SparkFun. Each LilyPad was creatively designed with large connecting pads and a flat back to allow them to be sewn into clothing with conductive thread. The LilyPad also has its own family of input, output, power, and sensor boards that are also built specifically for e-textiles. They're even washable!



C. Mega Arduino (R3)

The Arduino Mega is like the UNO's big brother. It has lots (54!) of digital input/output pins (14 can be used as PWM outputs), 16 analog inputs, a USB connection, a power jack, and a reset button. It contains everything needed to support the microcontroller; simply connect it to a computer with a USB cable or power it with a AC-to-DC adapter or battery to get started. The large number of pins make this board very handy for projects that require a bunch of digital inputs or outputs (like lots of LEDs or buttons).



D. Arduino Leonardo

The Leonardo is Arduino's first development board to use one microcontroller with built-in USB. This means that it can be cheaper and simpler. Also, because the board is handling USB directly, code libraries are available which allow the board to emulate a computer keyboard, mouse, and more!



II. ESP8266

A. What is ESP8266?

ESP8266 (technically ESP8266EX) is a WiFi Module based on Cadence Tensilica L106 32-bit MCU manufactured by Espressif Systems. The ESP8266 SoC contains a fully functional WiFi Stack and TCP/IP Stack that allows any Microcontroller to get connected to WiFi Network.

B. ESP8266 library

1) ESP8266WiFi library

ESP8266WiFi library has been developed basing on ESP8266 SDK, using naming convention and overall functionality philosophy of the Arduino WiFi Shield library. Over time the wealth Wi-Fi features ported from ESP8266 SDK to this library outgrew the APIs of WiFi Shield library and it became apparent that we need to provide separate documentation on what is new and extra.

The ESP8266WiFi library is broken up into several classes. In most of cases, when writing the code, the user is not concerned with this classification. We are using it to break up description of this library into more manageable pieces.

B BufferDataSource	E ESP8266WiFiAPClass	S SList	WiFiClient WiFiClientSecure WiFiEventHandlerOpaque WiFiEventModeChange WiFiEventSoftAPModeProbeRequestReceived WiFiEventSoftAPModeStationConnected WiFiEventSoftAPModeStationDisconnected WiFiEventStationModeAuthModeChanged WiFiEventStationModeConnected	WiFiEventStationModeDisconnected WiFiEventStationModeGotIP WiFiServer WiFiUDP
C ClientContext	ESP8266WiFiClass ESP8266WiFiGenericClass ESP8266WiFiMulti ESP8266WiFiScanClass ESP8266WiFiSTAClass	SSLContext		
D DataSource	P ProgmemStream	U UdpContext		
		W WifiAPlist_t		

2) Ticker

Library for calling functions repeatedly with a certain period: Ticker Basic, Ticker Functional and Ticker Parameter. It is currently not recommended to do blocking IO operations (network, serial, file) from Ticker callback functions. Instead, set a flag inside the ticker callback and check for that flag inside the loop function.

3) EEPROM

EEPROM library uses one sector of flash located just after the embedded file system. This is a bit different from standard EEPROM class. You need to call EEPROM.begin(size) before you start reading or writing, size being the number of bytes you want to use. Size can be anywhere between 4 and 4096 bytes. EEPROM.write does not write to flash immediately, instead you must call EEPROM.commit() whenever you wish to save changes to flash. EEPROM.end() will also commit, and will release the RAM copy of EEPROM contents.

III. IOT APPLICATIONS

The ESP8266 is often used for sensor devices especially in measure temperature and moisture. Therefore, people can control the environment statement from far distance because

all the collected data will be delivered by wifi. This paper will explain further about the way ESP8266 operates in a specific device which help people receive the temperature and moisture so we can see the wide applications in daily life.

Firstly, we are now going to use the sensor. Again, remember that we are using the Arduino IDE, so we can code just like we would do using an Arduino board.

If it has not been done yet, install the Adafruit DHT sensor library using the Arduino IDE library manager. Then, we read data from the DHT11 sensor, and print the value of the temperature and humidity on the web. ESP8266 Node MCU can be applied to any micro-controller design as a Wi-Fi adaptor through SPI/SDIO or I2C/UART interfaces. Node MCU operates as a webserver, once user logs in exactly IP address you can see the data output.



IV.CONCLUSION

The Internet of Things (IoT) has been a trending field in the world of technology. It has changed the way we work. Physical objects and the digital world are connected now more than ever. This paper presents an overview of Arduino family and introduces the IoT application using ESP8266 NodeMCU. The kit is at an unbelievable price and it is perfect for just about any IoT project.

V. ACKNOWLEDGEMENTS

This research is the first project in the course of embedded systems. I would like to thank all the teachers in the Faculty of Information Technology Hanoi University who dedicated teaching, equipping me with valuable knowledge during the past two years. Especially, I would like to send my gratefulness to my instructor, Dr. Hung Ta, who has spent a lot of time and dedication to help me complete this research. Valuable comments and suggestions of the reviewers on this paper are highly appreciated.

REFERENCES

- [12] J. Parab, S. A. Shinde, V. G. Shelake, R. K. Kamat, and G. M. Naik, “*Practical Aspects of Embedded System Design using Microcontrollers*”, Springer, 2008.
- [2]. Arduino, <https://www.arduino.cc/>
- [3]. Project Hub, https://create.arduino.cc/projecthub/Arca_Ege/using-dht11-b0f365

SENTIMENT ANALYSIS USING LONG SHORT-TERM MEMORY RECURRENT NEURAL NETWORK

SVTH: Nhâm Gia Hoàng Anh, Nguyễn Tiến Huy, Nguyễn Khắc Phúc, Bùi Đình Quân

GVHD: ThS Bùi Quốc Khánh

Tóm tắt

Bài báo nghiên cứu về đề tài phân tích cảm xúc trong câu nói, chủ yếu tập trung những bình luận, đánh giá của người dùng trên các mạng xã hội. Trong bài nghiên cứu khoa học này, chúng tôi sẽ sử dụng mạng thần kinh nhân tạo (mạng hồi quy LSTM) để giải quyết bài toán tìm được liệu trong câu bình luận mang ý nghĩa tích cực hay tiêu cực. Bài làm của chúng tôi đạt được kết quả tốt sau khi so sánh với các giải pháp hiện có trong vấn đề này. Cụ thể là chúng tôi đã giải quyết vấn đề overfitting của những bài làm trước đó. Tuy nhiên bài cũng chúng tôi vẫn cần nhiều thời gian cho nghiên cứu và thí nghiệm để nâng cao độ chính xác cũng như tìm giải pháp để áp dụng cho bài toán bằng Tiếng Việt.

Abstract: With the explosion of internet and social networks coming, there will be a lot of knowledge as well as useful information that can be derived from the emotions of people participating in social networks. The sentiment analysis, in an easy-to-understand way, is to listen and understand what is being said about brands, products on social media, how to say it, good or bad. Thus, to measure sentiment, discussion will be divided into Positive, Negative. This paper will explain how to sentiment analysis using machine learning, and specifically RNN and LSTM.

I. INTRODUCTION

RNN - Recurrent Neural Network is an algorithm that has gained a lot of attention recently because of the good results obtained in the field of natural language processing [1]. In this paper, the focus will be on RNN and one more special form is Long Short-term Memory (LSTM). And from the obtained models, we will have evidence to compare the effectiveness of this algorithm compared to another type of network, Convolution Neural Network (CNN) [2].

The main idea of RNN (Recurrent Neural Network) is to use sequences of information [3]. In traditional neural networks, all inputs and outputs are independent of each other. That means they are not chained. But these models do not fit in many problems. For example, if you want to guess the next word that might appear in a sentence, you also need to know how the previous words appear. RNNs are called recurrent because they perform the same task for all elements of a sequence with the output depending on previous calculations. In other words, RNN has the ability to remember previously calculated information. Traditional neural network models cannot do that, which can be considered as a major drawback of traditional neural networks. For example, if you want to categorize scenes that occur at all times in a movie, it is not clear how it is possible to understand a situation in the film but depend on previous situations. then if using traditional neural networks. The Recurrent Neural Network

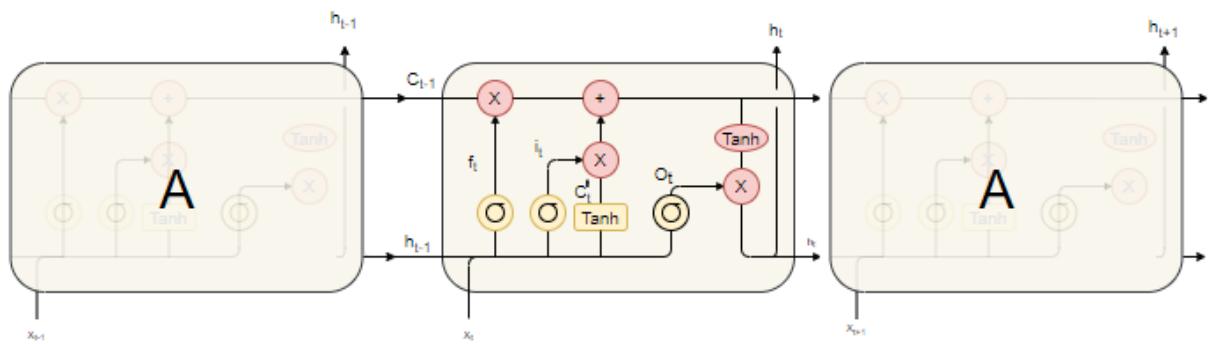
was born to solve that problem.

This network contains internal loops that allow information to be saved. In theory, RNN can use information of a very long document, but in reality, it can only remember a few steps before, it is called “long-term dependency”. This continues to lead to the improvement of this network, and from there, the LSTM network is created with the same basic structure as the RNN. LSTM is designed to avoid “long-term dependency”. Remembering information over a long period of time is their default feature, no need to train it to be able to remember it. That is, it was able to memorize without any intervention. In the next part, how the LSTM network work will be clarified.

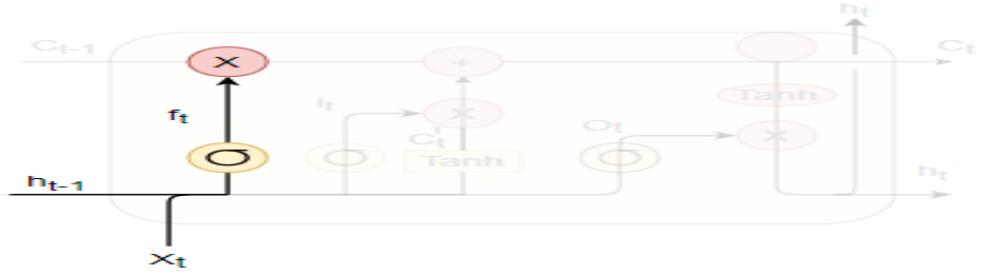
II. MODEL

RNN can handle difficult mathematical problems when the input is a sequence of data, however, when the data becomes too long and unrelated, there will be a large deviation. This phenomenon is called vanishing gradient (sometimes exploding gradient), which causes the network to forget the original words because the weights of that word will reach 0. This is the main reason for the arise of “Long-term dependency” challenge in RNN. To tackle these issues, LSTM was born [4].

LSTM is an improved version of RNN, introduced by Hochreiter & Schmidhuber in 1997. The great idea of LSTM is that it creates a cell state, a memory that can store and update information throughout the process of running RNN. The cell state is kind of like a conveyor belt. It runs straight down the entire chain, with only some minor linear interactions [5]. In order to do that, LSTM uses 3 “gates” of neural network instead of single neural network layer in normal RNN. Those gates are respectively forget, input and output, and for each LSTM there will be 3 main inputs: new data x_t , hidden state h_{t-1} (can be known as previous output) and previous cell state C_{t-1} .



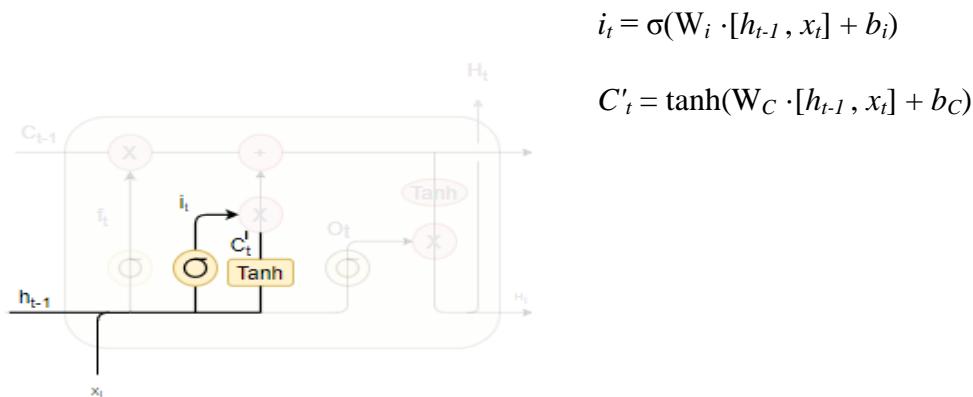
Based on the diagram above, there is a horizontal line (located in the upper corner of the model) that goes through all the LSTM networks, that horizontal line is the cell state, acting as the main brain of the process. Along with the support from 3 “gates” step by step as follows:



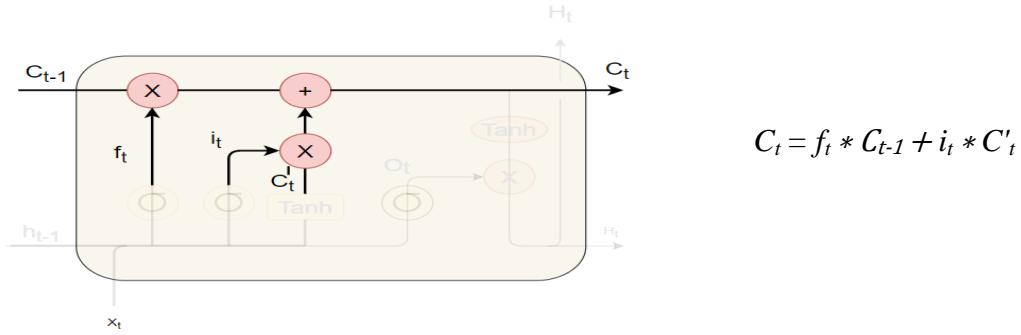
In the first step, Cell State from previous LSTM C_{t-1} removes irrelevant information by interacting with the forget gate f_t through matrix multiplication. The f_t will apply the sigmoid activation function to give values from 0 to 1 to decide whether to “forget” or “not to forget” (if a result of zero means forgetting and vice versa). The forget gate will take short-term memory from previous LSTM h_{t-1} and new data x_t , multiply them by the weights W_f and bias b_f of it. After performing the operations with the necessary inputs, include them in the sigmoid function to calculate f_t .

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

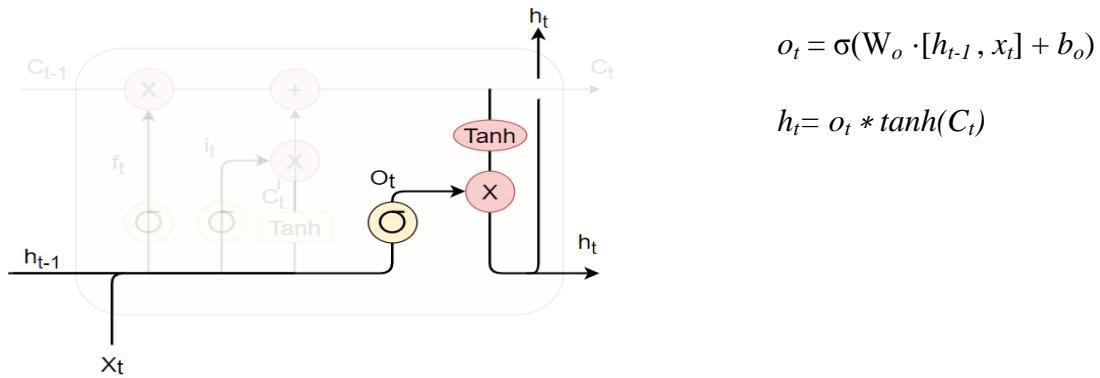
Next, the model filters out the information when entering the cell state through input gate. At this stage, there will be two processes taking place, with the purpose to see whether new data should be added or ignored and create a new vector called candidate values to evaluate the value of the added data. In the first process, h_{t-1} with x_t will be passed to i_t (using the sigmoid function) and will produce an output of 0 to 1, 0 meaning ignoring and vice versa. The second process uses a tanh activation function to produce the vector of candidate values, C'_t , this vector can evaluate the impact of the new input data through the oscillation from -1 to 1 of the tanh function. This layer also takes h_{t-1} and x_t as input.



After deciding what will be discarded in the old cell state and what can be added to the new cell state, C_t will be summarized with the following formula:



As a final step, the output gate layer determines the network outcome. In normal RNN, the output will be calculated via the sigmoid layer with hidden state h_{t-1} and new data x_t . But in LSTM, the output will be interacted with cell state C_t to give the more desired values. When multiplied by the output o_t , C_t will go through the tanh layer to put the values of C_t between -1 and 1 to regulate the final result.



III. IMPLEMENTATION AND COMPARISONS

1. Dataset:

The dataset used in this paper is a list of over 34,000 consumer reviews for products like Kindle, Fire TV Stick, and more from Amazon [6]. There are 21 features in total, but for the research purpose only three main features will be used, which are “title”, “text” and “rating”. The dataset includes 34000 instances but they are still raw data so data preprocessing is necessary before starting to train and test. Because in this dataset there are only 2,300 negative cases and there are 31,700 positive cases. Therefore, in order to avoid being overly biased on positive, only 2,500 positive cases were selected for research in order to balance with negative, which reduced the total number of instances to 5000.

After completing the data preprocessing step, this dataset will be divided into 3 main parts: train, test, validation to learn and create suitable model.

2. Implementation:

During the process of sentiment analysis, there will be two main objectives need to do, first is data preprocessing and then training under LSTM model. After training the model and getting results based on our prior research, the results will be compared with those of others, specifically with another LSTM and one based on the CNN model. This comparison aims to help evaluating whether the research process would yield better results compared to previous ones.

In the data preprocessing stage, the filter of stop words will be highly appreciated, because these words do not bring much valuable information when being put into the network and will dilute the important information of other words. After removing unnecessary words, the label will be initialized based on 'rating'. Because the rating ranges from 1-5, the value from 1 to 3 encode is 0 (negative) and from 4 to 5 encode is 1 (positive).

- - -

Review Sentiment		
28775	3rd kindle 3rd kindle continued move up on scr...	0
12393	adequate needed replace old kindle reader gave...	1
8158	never buy purchased item on sale alternative t...	0
12159	excellent tablet bought one son great tablets ...	1
8498	easy fun absolutely love kindle fire able ever...	1
3655	too many bells whistles just broke 3rd generat...	0

When the words have been refined to bring high information value, in the model training step, the parameters will be adjusted appropriately to give the perfect main model. Specifically, the LSTM network will consist of 40 layers, a magnetic array will have a length of 60, embedding size is 32 and run 10 times epochs.

The final result of the model is based on the image below.

```

Train on 2886 samples, validate on 962 samples
Epoch 1/10
2886/2886 [=====] - 6s 2ms/step - loss: 0.6920 - accuracy: 0.5243 - val_loss: 0.6922 - val_accuracy: 0.5229
Epoch 2/10
2886/2886 [=====] - 4s 1ms/step - loss: 0.6814 - accuracy: 0.5658 - val_loss: 0.6280 - val_accuracy: 0.6549
Epoch 3/10
2886/2886 [=====] - 4s 2ms/step - loss: 0.5184 - accuracy: 0.7779 - val_loss: 0.5375 - val_accuracy: 0.7703
Epoch 4/10
2886/2886 [=====] - 5s 2ms/step - loss: 0.4973 - accuracy: 0.7893 - val_loss: 0.5245 - val_accuracy: 0.7817
Epoch 5/10
2886/2886 [=====] - 4s 2ms/step - loss: 0.4918 - accuracy: 0.8049 - val_loss: 0.5844 - val_accuracy: 0.7765
Epoch 6/10
2886/2886 [=====] - 5s 2ms/step - loss: 0.5088 - accuracy: 0.7814 - val_loss: 0.5052 - val_accuracy: 0.7848
Epoch 7/10
2886/2886 [=====] - 5s 2ms/step - loss: 0.4351 - accuracy: 0.8365 - val_loss: 0.4778 - val_accuracy: 0.8160
Epoch 8/10
2886/2886 [=====] - 4s 2ms/step - loss: 0.4102 - accuracy: 0.8531 - val_loss: 0.4855 - val_accuracy: 0.8129
Epoch 9/10
2886/2886 [=====] - 4s 2ms/step - loss: 0.3949 - accuracy: 0.8586 - val_loss: 0.4984 - val_accuracy: 0.7931
Epoch 10/10
2886/2886 [=====] - 5s 2ms/step - loss: 0.3847 - accuracy: 0.8676 - val_loss: 0.5095 - val_accuracy: 0.8077

```

With a run time of 46 seconds, the model has train accuracy of 86.76% and test accuracy of 80.37%, which is a pretty great result. And this is the outcome of running an example sentence.

```

In [63]: X_test1 = ['the product is not good, it is very terrible, i am so disappointed about this product!']
y_test1 = [0]
test = pad_sequences(tk.texts_to_sequences(X_test1), maxlen=60, padding='post')
test

Out[63]: array([[ 16,    3,    6,   13,  521,   10,   171,   16,    0,    0,    0,    0,    0,
       0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,
       0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,
       0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,
       0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0],
[[ 0.09242114]],
1.0

```

When putting in a sentence that means negative, because this is supervised learning, it should add a label value of 0 (negative). The model returned with an exact result.

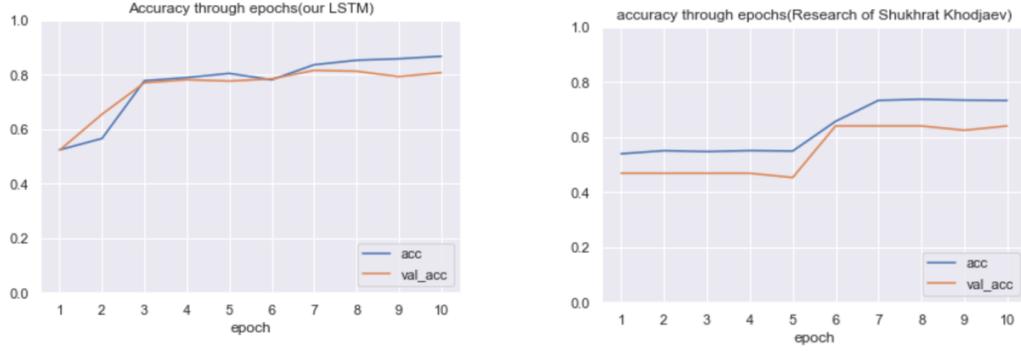
3. Comparison:

To evaluate the effectiveness of the application, there will be a comparison of results between this scientific research paper and others. Their coding will be kept the same and only replace dataset (after processing)

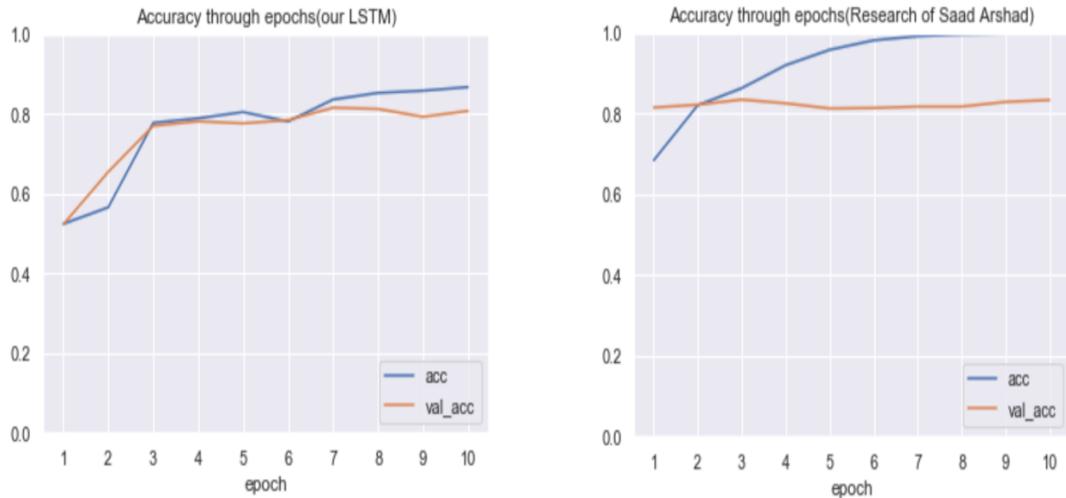
First, we will compare with the work of Shukhrat Khodjaev, the author also uses the LSTM model to handle the problem of sentiment analysis [7]. However, the data processing method will be a little different, their post does not separate stop words but in our research,

the separation of stop words is mentioned. Their application result returned with train accuracy is 73.33%, test accuracy is 75.14% and runtime is 306s, it is clearly see the difference in time and accuracy.

Although in their post, the result will be different due to dataset. However, in the process of researching, we found that the division of ratings was a bit 'biased'. Their article will be mentioned in the references section for everyone to refer.



Next is the article using the CNN model, by Saad Arshad [8]. In this article, they also handle stop words like ours, the difference here is that they use CNN and not RNN. Their application returned the result with train accuracy as 99.84% and test accuracy as 82.36% in a period of 290s. This is a high result, but nevertheless we can see their overfitting in the separation between accuracy and validation accuracy in the chart below.



After comparing with 2 research papers of other author, our application showed that the accuracy is very good. And the main strength of this application is the fast processing time.

IV CONCLUSION

In order to achieve this goal, a lot of things must be carefully calculated. Although there are good results, but when testing the examples is still wrong, this causes the process to start again to meet the main requirements of the problem. And to optimize the product, there are always links with the LSTM algorithm to find the right solution. Such as separating the stop words, dividing the number of words in the sentence by 60 to avoid diluting the information of words and so on. The application in this article not only checks a large number of new instances, but even individual instances can produce accurate results, which is something that other article applications cannot do (sometimes it is overfitting or bias). Analyzing emotions in other people's sentences is very normal in human life, and even the most ordinary things, artificial intelligence can do not just super things. This is the main inspiration for us to choose the topic Sentiment Analysis. With the successful analysis of emotions in English sentences, we have a solid foundation for future work, building an application for Vietnamese.

REFERENCES

- [1] Tom Young, Devamanyu Hazarika, Soujanya Poria and Erik Cambria, *Recent Trends in Deep Learning Based Natural Language Processing*, Singapore, 2018.
- [2] Raghav Prabhu, *Understanding of Convolutional Neural Network (CNN)- Deep Learning*, 2018.
- [3] John A. Bullinaria, *Recurrent Neural Networks*, pp.2-3, 2015.
- [4] Y. Wang, X. Zhang, X. Wang, R. Zhu, Z. Wang and L. Liu, "Text Sentiment Analysis Based on Parallel Recursive Constituency Tree-LSTM," *2019 IEEE Fourth International Conference on Data Science in Cyberspace (DSC)*, Hangzhou, China, 2019, pp. 156-161.
- [5] Christopher Olah, *Understanding LSTM Networks*, Colah's blog, Aug 27, 2015.
- [6] Datafiniti's Product Database, *Consumer review of Amazon Product*, Aug 14, 2017. Available: <https://www.kaggle.com/datafiniti/consumer-reviews-of-amazon-products/metadata>.
- [7] Shukhrat Khodjaev, *Application of RNN for customer review sentiment analysis*, Sep 26, 2018, Available: <https://towardsdatascience.com/application-of-rnn-for-customer-review-sentiment-analysis-178fa82e9aaf>.
- [8] Saad Arshad, *Sentiment Analysis / Text Classification Using CNN (Convolutional Neural Network)*, Sep 21, 2019, Available: <https://towardsdatascience.com/cnn-sentiment-analysis-1d16b7c5a0e7>.

MASK DETECTING PROGRAM

SVTH: Phạm Quang Chiến, Nguyễn Thu Hà,
Vương Khánh Linh, Trần Thu Hiền, Vũ Văn Tùng
GVHD: ThS Bùi Quốc Khanh

Tóm tắt: Mọi người đều được khuyến khích đeo khẩu trang trong suốt sự bùng phát của dịch corona để ngăn chặn sự lây lan của virus COVID19 một cách hiệu quả. Tuy nhiên, trên thực tế, không có nguyên mẫu nhận dạng mặt nạ công khai nào để giúp tất cả mọi người có thể phát hiện ai đó đeo khẩu trang khi tham gia vào nơi công cộng như trong nhà máy, công ty, trường học, v.v ... Điều này hoàn toàn làm cho công nghệ nhận dạng khuôn mặt hiện tại trong nhiều trường hợp không thành công. Do đó, dựa trên nghiên cứu thứ cấp, bài viết này thảo luận về việc phát hiện mặt nạ để nâng cao chất lượng nhận dạng trên khuôn mặt đeo mặt nạ từ các công nghệ nhận dạng khuôn mặt hiện tại, miêu tả bộ dữ liệu, thư viện, phương pháp, thuật toán được sử dụng để phát triển dự án. Tóm lại, bài viết rút ra kết luận rằng mục đích của ứng dụng này là vì sức khỏe của cộng đồng, góp phần vào tình hình khó khăn của thế giới hiện nay, để đẩy lùi dịch bệnh nhanh chóng.

Abstract: Everybody is encouraged to wear face mask during coronavirus outbreak to prevent the spread of COVID19 virus effectively. Actually, however, there is no public masked face recognition prototypes to help all people can detect someone wearing a mask when joining in society such as in factory, company, school, etc. This totally makes current facial recognition technology in many cases unsuccessful. Therefore, this paper, based on secondary research, discusses about mask detection enhancing identification quality on masked faces from current face recognition technologies, describes which dataset, library, methodology, algorithm used to develop project. All in all, the paper draws the conclusion that the purpose of this application is for the health of the community, contributing to the difficult situation of the world today, to repel the epidemic quickly.

Keywords: Convolutional Neural Network (CNN), Fully-connected layers (FC layers), fine-tuning, Keras, MobileNetV2, Convolution layers (CONV layers).

I. INTRODUCTION

A large and increasing outbreak of a novel coronavirus first emerged in Wuhan, China's Hubei province in December 2019 **Invalid source specified.** that also can lead to acute lung infections, including fatal acute respiratory distress syndrome **Invalid source specified..** The new coronavirus has been confirmed with transmission from human to human **Invalid source specified.**, because of the massive transport and high population freedom of movement before the Lunar New Year, this new coronavirus has spread very quickly to other areas of the world with fatal consequences. COVID-19 has become a dangerous pandemic threatening to life of people all over the world.

To avoid outbreaks of disease, WHO recommends some measures that people should do to protect themselves from virus and one of these is wearing a medical mask. Whenever and wherever they are, following the instructions for putting on, taking off, and disposing of medical masks, avoid crowded groups and enclosed areas **Invalid source specified..**

Nevertheless, opposite to widely spread news and announcements around the world, there are still some individuals who are unaware of protecting their own safety, not wearing masks as required by government, that affect the community.

Face detection has also been an important research field of the vision community over the past two decades. Marian found that in the video stream, the device automatically detects frontal faces and codes each frame in terms of dimensions: optimistic, furious, disgusting, afraid, happy, sorrow, and disappointment**Invalid source specified.** Sachin mentioned her approach follows some works but constructs a deeper CNN for the detection of the face **Invalid source specified.** Oliver's research showed that a quick face recognition scan algorithm was introduced and the model may be constructed using different methods: Neural Networks and Support Vector Machines,etc**Invalid source specified.** In other article, Shaumik also reported about this title with difference library which is Keras**Invalid source specified.** According to him, in Keras, he would use a pre-trained MTCNN model to detect faces in the images. Once he had extracted the faces from a picture, he will calculate a score of similarity between those faces to figure out if they belong to the same individual.

Ironically, up until now there have been no research written about mask detection, some journalists have tried to approach the article, but none of them reflect the focus of the topic. Face recognition - the most essential means of identification, had almost failed in the age of COVID19, such as public entry and exit, face attendance, face gates at train stations, face authentication based payment service**Invalid source specified.** Therefore, this paper, as the result of our research, aim to develop an application to detect who wearing mask or not , remind all people protect themselves carefully, also figure out dataset, methodology, library that used and learnt from the above researchers.

II. METHODOLOGY

A. Dataset

The dataset used in this paper consists of “with mask” and “without mask” images of user. We will use the dataset to build a COVID-19 face mask detector with computer vision and deep learning. Our goal is training a custom deep learning model to detect whether a person is wearing a mask or not. The dataset is then split into training set and test set to serve for the training and test phases, respectively. The data has total 1376 observation and is clean, which is suitable for research purpose.

B. Technology

The encoder-decoder architecture has been implemented using Googles framework for distributed machine learning TensorFlow. We used Keras on top of TensorFlow. Keras was chosen as it is designed for fast prototyping and experimentation with a simple API. It allows to configure NNs in a modular way by combining different layers, activation functions, loss functions, and optimizers, etc. Keras provides out of the box solutions for most of the standard deep learning building blocks. However, if someone wants to build a custom or

novel implementation, Keras API could be quite limited, and libraries like TensorFlow will be a better choice. Besides, we use OpenCV being an open source library for image and video analysis to implement our COVID-19 face mask detector in real-time video streams. For evaluation, different metrics have been used, that were partly taken from the scikit-learn library for Python.

C. Fine-tuning with Keras and Deep Learning

To train a classifier to automatically detect whether a person is wearing a mask or not, we are fine-tuning pre-trained CNN model in Keras and TensorFlow. In our project, we choose MobileNetV2 model for fine-tuning which is one of the highly efficient pre-trained CNN models.

Regarding to fine-tuning technique, fine-tuning is a strategy which we use to establish a baseline model while saving considerable time. We take a scalpel and cut off the final set of FC layers from a pre-trained CNN (MobileNetV2), then replace the head with a new set of FC layers with random initializations. From there, all layers below the head are frozen so their weights cannot be updated. We then train the network using a very small learning rate so the new set of FC layers can learn patterns from the previously learned CONV layers earlier in the network. The way how fine-tuning works is represented in the following figure 1.

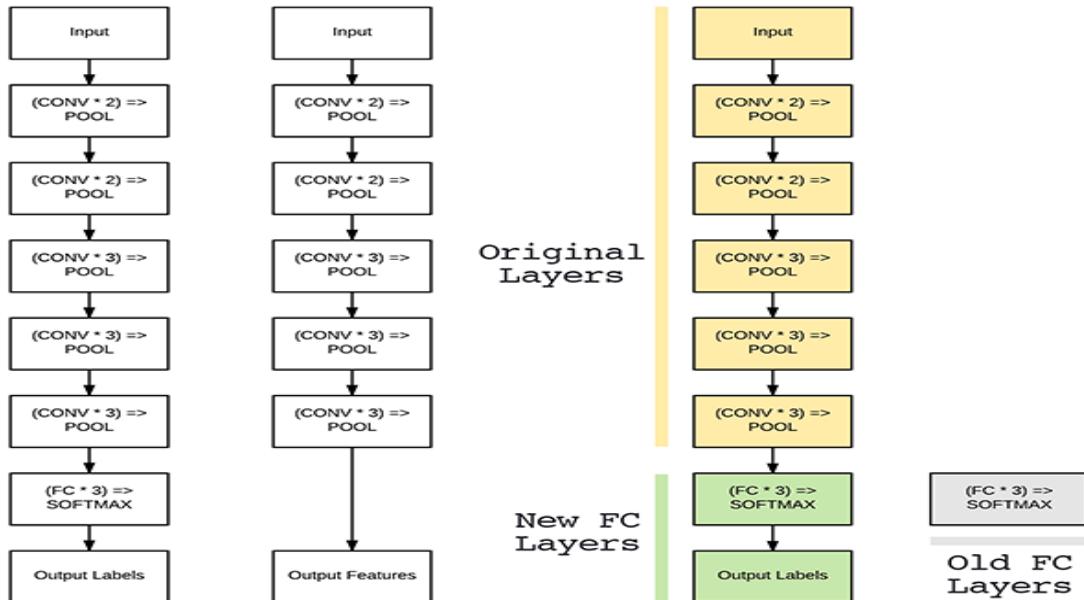


Figure 1: The way how fine-tuning works

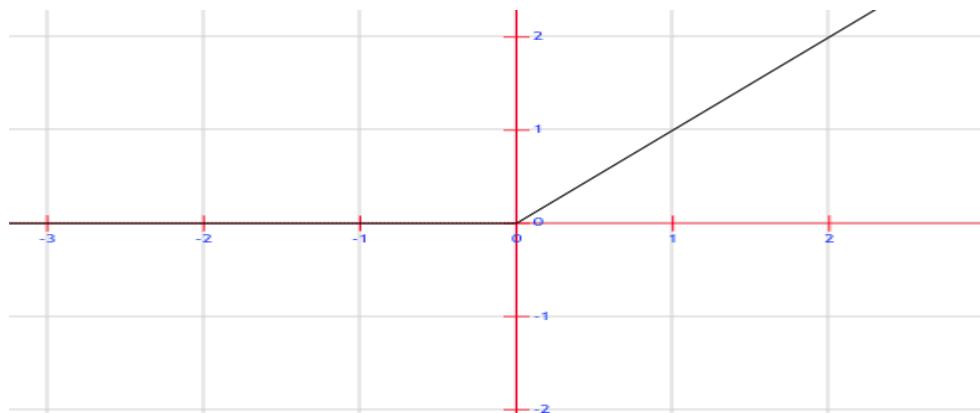
Looking at figure 1, the original FC Layers are removed and replaced with a brand-new FC head. These FC layers can then be fine-tuned to a specific dataset (the old FC Layers are no longer used). Especially, in fine-tuning process, one of the vital steps which is constructing a new FC head and appending it to the base in place of the old head is analyzed more thoroughly.

To creating FC layers, we need to declare Dense layer from tensorflow.keras.layers. With our project, we declare two parameters in Dense including number of units in a FC layer and activation. Activation determines the output of a deep learning model, its accuracy, and the computational efficiency of training a model.

```
headModel = Dense(128, activation="relu")(headModel)
```

In the above code script, the number of units of a FC layer is 128 units, and activation function is ReLU (Rectified Linear Unit). The ReLU is the most used activation function nowadays when training neural networks. ReLU simply filters values less than 0. ReLU is expressed by the following equation and graph.

$$f(x) = \max(0, x)$$



ReLU is more popular because this function brings many outstanding advantages such as efficient computation, the faster speed of convergence.

D. Model Evaluation Criteria

To evaluate the performance of forecasting models, we use classification metrics including accuracy, precision, recall, F1-score to compare predicted values with observed values. Accuracy is number of instances correctly classified over the total number of predictions. Precision is fraction of instances correctly classified. Recall is fraction of positive examples in the test set that have been correctly classified and also called coverage, sensitivity or True Positive Rate. F1-score is a metric to balance between Precision and Recall.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{N}} \text{ where } \text{N} = \text{TP} + \text{TN} + \text{FN} + \text{FP}$$

In the above equations, TP is true positive, FP is false positive, FN is false negative, TN is true negative. The above classification metrics values range from 0 to 1, and the larger these values are, the more effective the trained model is.

III. IMPLEMENTATION

To build a COVID-19 face mask detector, we implement main steps including data preprocessing, training MobileNetV2 model and detecting faces with and without masks. In terms of data preparation for image dataset, pre-processing steps include resizing images to 224×224 pixels, converting to array format, and scaling the pixel intensities in the input image to the range $[-1, 1]$ (via the `preprocess_input` function). After that, we append the pre-processed image and associated label to the data and labels lists, respectively and ensure that our training data is in NumPy array format.

```
57
58     image = load_img(imagePath, target_size=(224, 224))
59     image = img_to_array(image)
60     image = preprocess_input(image)
61
62
63     data.append(image)
64     labels.append(label)
65
66 data = np.array(data, dtype="float32")
67 labels = np.array(labels)
```

With regard to training pre-trained MobileNetV2 model, we need to prepare this model for fine-tuning. Fine-tuning is a strategy which we use to establish a baseline model while saving considerable time. Setting up fine-tuning goes through a three-step process including loading MobileNetV2 with pre-trained ImageNet weights, leaving off head of network and constructing a new FC head, and freezing the base layers of the network to not update the weights of these base layers during the process of backpropagation.

```

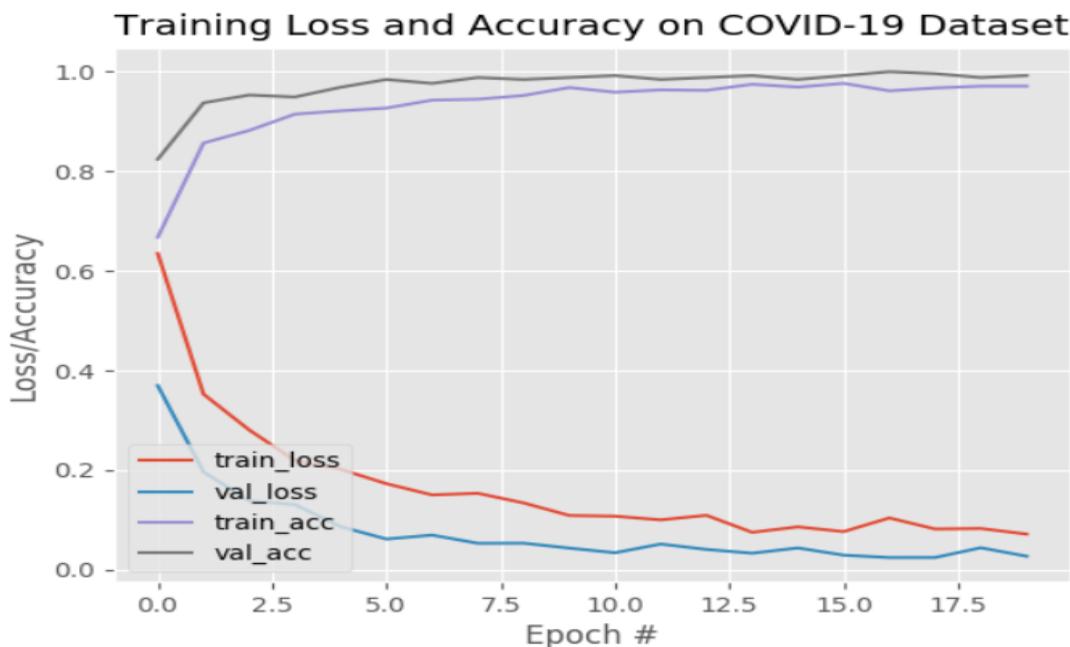
baseModel = MobileNetV2(weights="imagenet", include_top=False,
    input_tensor=Input(shape=(224, 224, 3)))

headModel = baseModel.output
headModel = AveragePooling2D(pool_size=(7, 7))(headModel)
headModel = Flatten(name="flatten")(headModel)
headModel = Dense(128, activation="relu")(headModel)
headModel = Dropout(0.5)(headModel) |
headModel = Dense(2, activation="softmax")(headModel)

model = Model(inputs=baseModel.input, outputs=headModel)

```

With our data prepared and model architecture in place for fine-tuning, we can compile and train our face mask detector network. Once the model is fit, we can estimate the performance of the model on the train and test datasets.



To present the visualization of the predicted values and the trained ones, we use matplotlib library to show the comparison among them.

Figure 2: Training loss and accuracy on COVID-19 database

Looking at Figure 2, we can see there are little signs of overfitting, with the validation loss lower than the training loss. Given such high accuracy, we are hopeful that our model will generalize well to images outside our training and testing set. With applying trained model and using OpenCV library, we can implement our COVID-19 face mask detector in real-time video streams.

IV. CONCLUSION AND FUTURE WORK

Generally speaking, masked face recognition technology can be used to recognize people wearing masks, but it is still not very effective compared to standard facial recognition technology that has already seen more than 99 percent accuracy. Another related task is the identification of face masks, that is to recognize whether a person wears a mask as necessary or not. Our research has contributed to the prevention and control of coronavirus epidemics and the development in industry through scientific and technological support. In fact, people still wear masks due to the regular incidence of hazy conditions, therefore the need for face recognition with masks may continue for a long time in future work.

REGERENCES

- [1] Y. W. X. L. L. R. J. Z. Y. H. L. Z. G. F. J. X. X. G. e. a. Chaolin Huang, "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *The Lancet*, pp. 497-506, 2020.
- [2] M. Z. X. D. J. Q. F. G. Y. H. Y. Q. J. Nanshan Chen, "Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study," *The Lancet*, pp. 507-513, 2020.
- [3] K. L. a. G. M. L. Joseph T Wu, "Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study," *The Lancet*, pp. 689-697, 2020.
- [4] W. Team, "World Health Organization," Advice on the use of masks in the context of COVID-19, 6 April 2020. [Online]. Available: [https://www.who.int/publications-detail/advice-on-the-use-of-masks-in-the-community-during-home-care-and-in-healthcare-settings-in-the-context-of-the-novel-coronavirus-\(2019-ncov\)-outbreak](https://www.who.int/publications-detail/advice-on-the-use-of-masks-in-the-community-during-home-care-and-in-healthcare-settings-in-the-context-of-the-novel-coronavirus-(2019-ncov)-outbreak).
- [5] G. L. C. L. I. F. J. M. Marian Stewart Bartlett, "Machine Learning Methods for Fully Automatic," in IEEE International Conference on Systems, Man and Cybernetics, San Diego , 2004.
- [6] M. S.-J. L. Sachin Sudhakar Farfade, "Multi-view Face Detection Using Deep Convolutional," 2015.
- [7] O. J. B. J.-E. V. a. M. C. RaphaeÈl FeÂraud, "A Fast and Accurate Face Detector," vol. 23, pp. 42-53, 2001.
- [8] S. Daityari, "sitepoint," Face Detection and Recognition with Keras, 7 November 2019. [Online]. Available: <https://www.sitepoint.com/keras-face-detection-recognition/>.
- [9] G. W. B. H. Z. X. Q. H. H. W. P. Y. K. J. Zhongyuan Wang, "Masked Face Recognition Dataset and Application," 2020.

ANALYZE INTERFACES AND SUGGEST IMPROVEMENT FOR ENTERTAINMENT WEBSITES IN VIETNAM

SVTH: Phạm Thanh Hà, Ngô Tuấn Anh, Nguyễn Sỹ Nhật, Đỗ Hồng Hà

GVHD: ThS Nguyễn Thùy Linh

Abstract—*Internet life is becoming more and more important to our lives, especially the youth. One of the most attractive purpose of using internet is entertainment. Millions of young people are accessing sites such as youtube, zing,... everyday, therefore, the interface design of these websites are crucial to build up its competitiveness to other competitors. In this paper, we analyze 4 sites that help young internet users to explore video, music, movie and comics, point out its both strengths and weaknesses, and suggest the way to make these site become more familiar to users. The role of HCI in design also applied.*

Keyword— copyrighted content, layout, UI/UX.

I. INTRODUCTION

A recent survey pointed out that Vietnam young people usually spend over 7 hours daily for entertaining purposes. We will not discuss on health and economical problems of this fact, in this research we focus on how to make these experience become more playful and better for their health in terms of technology and Human-computer interaction principles only. We have found out 4 examples, which can representative 4 kinds of entertainment sites in Vietnam, including:

Youtube.com for video broacasting

Mp3.zing.vn for music

Wattpad.com for reading comics

And galaxycine.vn for movie updates

From these examples, we identify its strong points and find out weak points which can be improved and suggest improvements.

II. RESULTS

A. Video site

YouTube is a website designed for sharing video, where user can watch, like, share, comment, download videos and upload their own videos.

Good and bad point in terms of UI/UX (User interface / User Experience) via: Use of color, layout decoration, well-placed necessary function (Are they placed in good position or not - why). Does webpage refresh every time you perform an action, indicate those similar actions.

Good points:

- YouTube has a simple and clear interface, which is easy to see, soft color, don't

distract the user.

- There is a navigation pane on the left hand side, which makes it's easy to manipulate and use website. There is a search bar at the top, which makes it's easy for user to search videos. The scroll bar if on the right because most of people use the right hand.
- Each specific symbol corresponds to the function of the website, which helps user easily and quickly manipulate.
- Webpage will refresh every time you perform an action.
- When user watches anything, it will automatically save data into the memory to recommend for later time and user can watch their history.
- When you see if there is a lag statue, it will show notice to support this problem.
- If your video contains copyrighted content that has been claimed by Content ID, YouTube will take down your video.

Bad points:

- The setting bar is not placed in a visible position (there are two positions: user must click on the user avatar or 3 dots if user don't login). It's difficult and unfamiliar for the new user to use.
- Video is not classified content and it doesn't have pagination on the interface for user to follow easily.

How do you tweak this website to be more familiar to users?

- We should put the setting bar outside the interface (can put in the left side and under other icons).
- We should have content classification and pagination

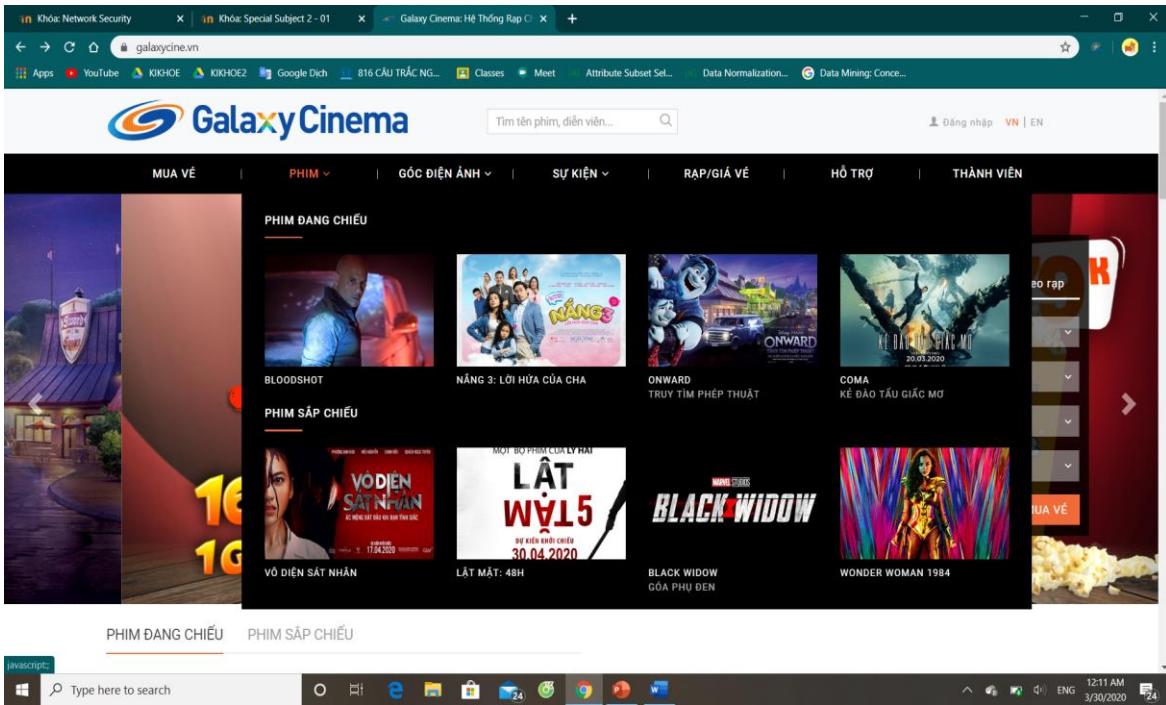
B. Movie site

This website is about booking films ticket and update new movies that are currently shown or about to be shown. In this website, it contains blog and news (review or comment) about the films for user to share their feelings.

Good and bad point in terms of UI/UX

Good points:

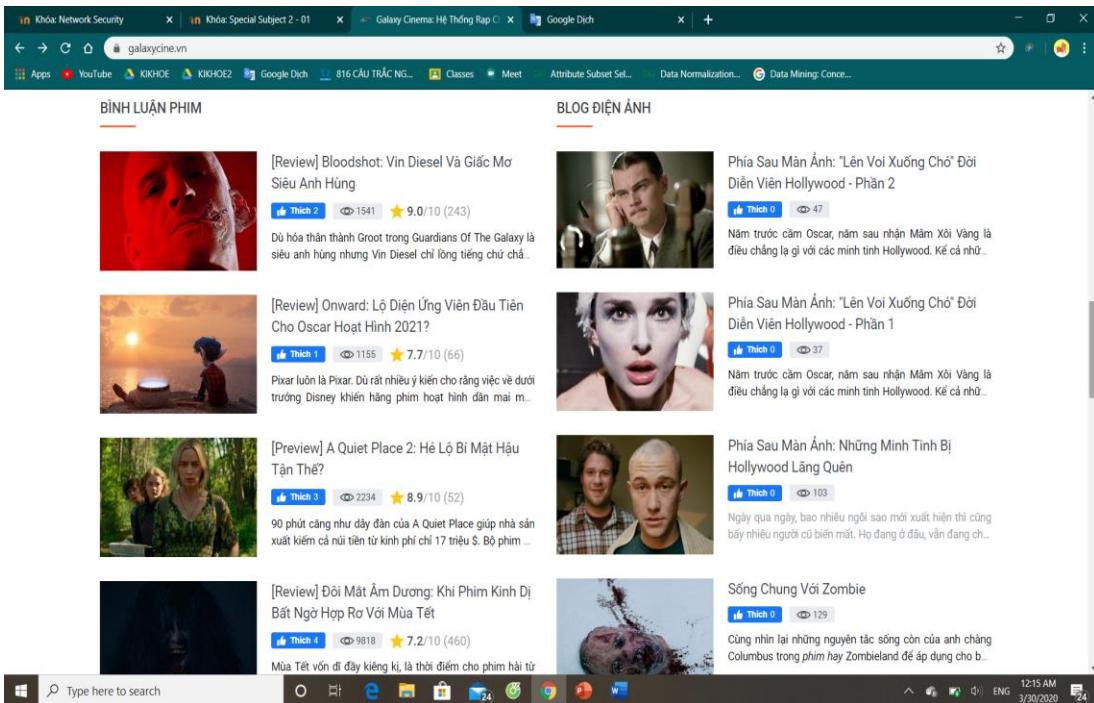
- Website uses white background that make all the information in every page because film's posters might be colorful.
- The layout and the order of buttons, navigations in this website are quite clear and easy for user to interact with. They put all of necessary navigations on top of the page like: booking tickets, prices, members, films, supports.



- The navigations on top bar contains pop-up mini table that helps user quickly reach their needs.

Bad points:

- Though this website has a blog and review categories but it is not highlighted separately. It uses the same interface with these two parts:



- Webpage refreshes every time users perform an action, indicate those similar actions.

How do you tweak this website to be more familiar to users

- I will separate the blog forum and news with the similar interface. Then I will put on navigations and let them on the top bar so that user can easily interact with them.
- Films category may be highlighted as a individual button. It makes user reach films that they want with less of time.
- The color may be change to black background to make on cinema effect for users.

C. Music site

The name of website is zingmp3.com. This website is for people who want to listen to some pop music or finding the music that they are looking for.

Good points about UI:

- Such as the website have good recommend in the center of the pages. Chart about top pop song from 3 main region that Vietnamese people use to hear (VN, USUK and K pop).
- The webpage use the bright color which very appering and easy to see.
- The UI is simple, friendly to use.

Bad point:

- The downside of this webpage is that the search bar is small and set in the top left of the page (witch usually be the account). Is hard to see in the first place

How do you tweak the site to become more familiar to user:

- I think they need to adjust the search bar to be more bigger and position it at the top center of the page.

D. Comic site

Wattpad.com is an online sharing website that is loved by the Vietnamese comic community. The archive on the site is extremely rich with all genres from Chinese language, detective stories, fanfic to modern Vietnamese stories. In addition to reading online, Wattpad also gives users an extremely useful application on mobile devices, you can download the whole story and read offline when there is no internet connection.

Good point:

- No annoying advertisement.
- Each section is separated, very clear.
- People can comment using Facebook account (social media that almost everyone use these days). The description of each serie is very informative, categorized. There is also version for mobile user.

Bad point:

- No option to scroll to top or search when user is at the middle or bottom of the page.
- Can't choose chapter/ issue when being at the middle of the page

How do you tweak the site to become more familiar to user:

- Create scroll to top button, search button and navigation button (between each chapter/ issue) that go along with the view of the users when they scroll down the page

III. CONCLUSION

HCI in design helps user understand how a system works faster and more effectively. Then the interaction between user and computer become more easier. Good design can help prevent many bugs/ unwanted situations/ questions. Good design comes along with better abilities for the user to control, manage the system. The better design is, the more chances happen to make the system be more recognized, client/ stakeholder like/ want it more and more people use it.

REFERENCES

- [1] A. Dix, J. Finlay, “Human Computer Interaction” Third Edition. Pearson; 3 edition (October 10, 2003)

HCI APPLICATION IN NUCLEAR REACTOR SYSTEM

INTERFACE DESIGN

SVTH: Lương Minh Hoàng, Đỗ Hồng Hà

GVHD: ThS Nguyễn Thùy Linh

Abstract—The use of nuclear power plants to produce electric energy is a safety-critical process where ultimate operational decisions still relies with the control room operators. Thus it is important to provide the best possible decision support through effective supervisory control interfaces. In nuclear power plant, the control room plays an extremely important role to ensure safety, especially in emergent cases and protect the plant as well as its surrounding. A human centered design approach, based on cognitive task analysis methods, was used to observe the operators training on the nuclear power plant simulator of the Human System Interface Laboratory (LABIHS). A new prototype of the interface including graphics, alarms and digital procedures was designed as an alternative to the current hardcopy procedure manuals. This paper will evaluate the prototype based on human-computer interaction evaluation techniques and suggest improvement in a stimulating emergency.

Keywords— nuclear reactor, control panel, alarm state.

I. INTRODUCTION

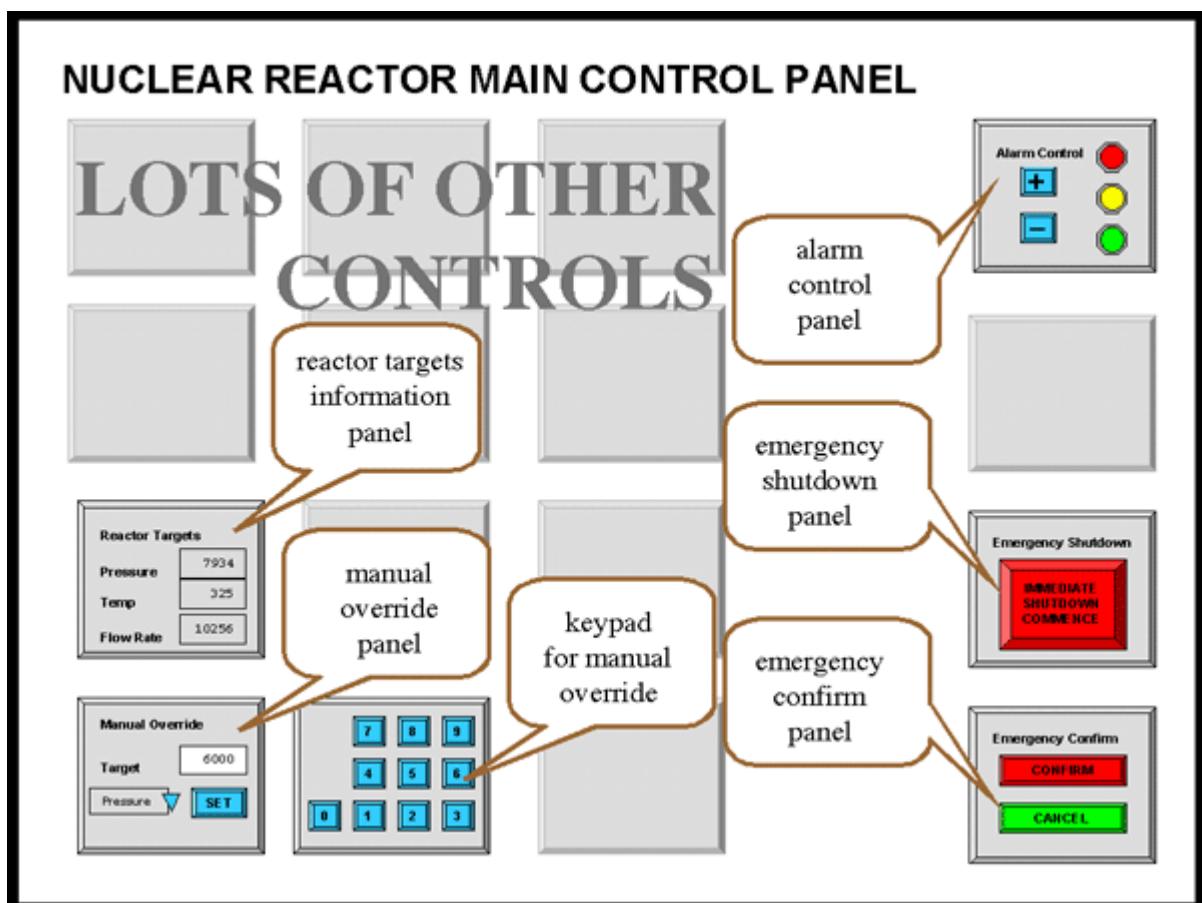


Figure CS.1 - nuclear reactor main control panel

Figure 1 shows a sketch of the control panel of a nuclear power plant. The actual panel is very large covering the whole wall of the control room and contains many sub-panels and controls. The locations of some controls at the two ends of the panel are shown in figure CS.1, although it should be noted that the panel is much wider than the illustration.

The system can be in one of three alarm states: GREEN, AMBER or RED.

(i) GREEN alarm state means everything is operating normally

(ii) AMBER alarm state is for when there is a minor problem with reactor operation. Workers in the reactor area are warned and take additional precautions, but no external services are involved.

(iii) RED alarm state is raised when the reactor is operating outside normal parameters and there is a possibility of external contamination. The police and other emergency services are alerted.

Typically AMBER state is raised once or twice a week and red state only a few times a year (so far only false alarms!). Raising a RED alarm unnecessarily causes significant inconvenience and cost both to the station staff and the external emergency services.

When the plant was commissioned, the alarm system controls worked as follows.

The current alarm state is indicated by which of the coloured lights on the **Alarm Control** panel is lit.. The '+' and '-' buttons on this panel increase or decrease the alarm state. Figure 4 shows a state transition network of the effects of the '+' and '-' buttons on the state as the system was initially installed.

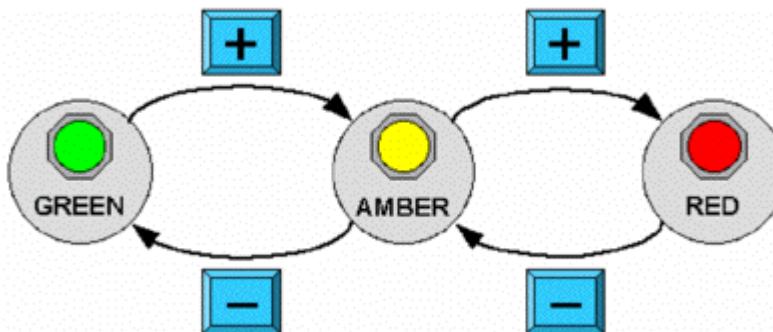


Figure CS.4 – STN for alarm state

When there is a very serious problem the operator can press the large red button labelled **IMMEDIATE SHUTDOWN COMMENCE** on the **Emergency Shutdown** panel, which initiates an emergency shutdown. This needs to be confirmed by pressing the **CONFIRM** button on the **Emergency Confirm** panel. (This is to prevent accidental shutdown of the plant.) The **CONFIRM** button is normally green, but glows red after the **IMMEDIATE SHUTDOWN COMMENCE** button has been pressed to remind the operator.

Emergency shutdown causes explosive bolts to blow that drive control rods into the

reactor completely stopping the nuclear reaction. Restarting the reactor after emergency shutdown may take several weeks and costs many millions of pounds in lost production and replacement of parts damaged during the shutdown procedure.

The **Reactor Targets** panel shows the current target state of several reactor operating parameters. These are normally set by an automatic control system to values that ensure optimal energy production.

In an extreme emergency the operator may need to control these targets. The **Manual Override** panel allows this.

Manual override is *only enabled* in RED alarm state.

To override a particular target the operator selects the desired target (Pressure, Temperature or Flow Rate) from a dropdown menu, types in the desired value using a numeric keypad and then confirms the value using the SET button. (The SET button is necessary to prevent part-typed numbers being treated as the new value.)

Some while after the plant was running a consultant suggested changing the operation of the **Alarm Control** panel and the software and hardware was revised in line with his recommendations. The current design works as follows.

Raising the alarm state from Green to Amber and back uses the '+' and '-' buttons as before. However now to raise the state from Amber to Red it is necessary to both press '+' and also confirm this by pressing the **CONFIRM** button on the **Emergency Confirm** panel.

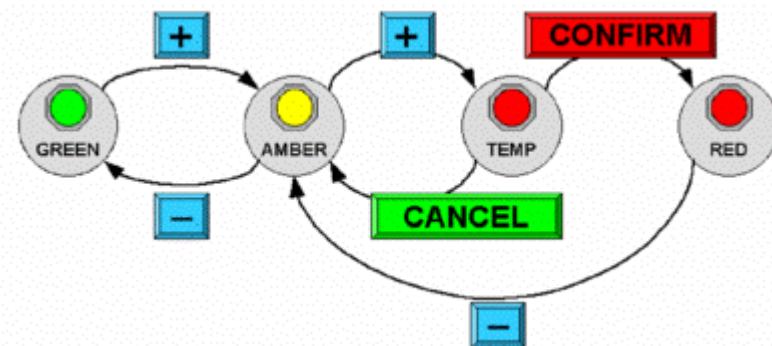


Figure CS.5 – STN for revised alarm state

II. PROBLEM

The problem here is, when an emergent case happen, if every step in safety procedure can ensure the meaningfulness and effectiveness to protect the whole system or not?

Firstly, we should consider use of color. Secondly, we analyze effectiveness of layout and other elements in the control panel including the way in which various visual elements support or hinder logical grouping and sequence. Finally, a stimulating emergency scenario will be taken place and from real experience, we identify weaknesses and suggest for improvement. These part will be presented in Findings part

III. FINDINGS

Use of color

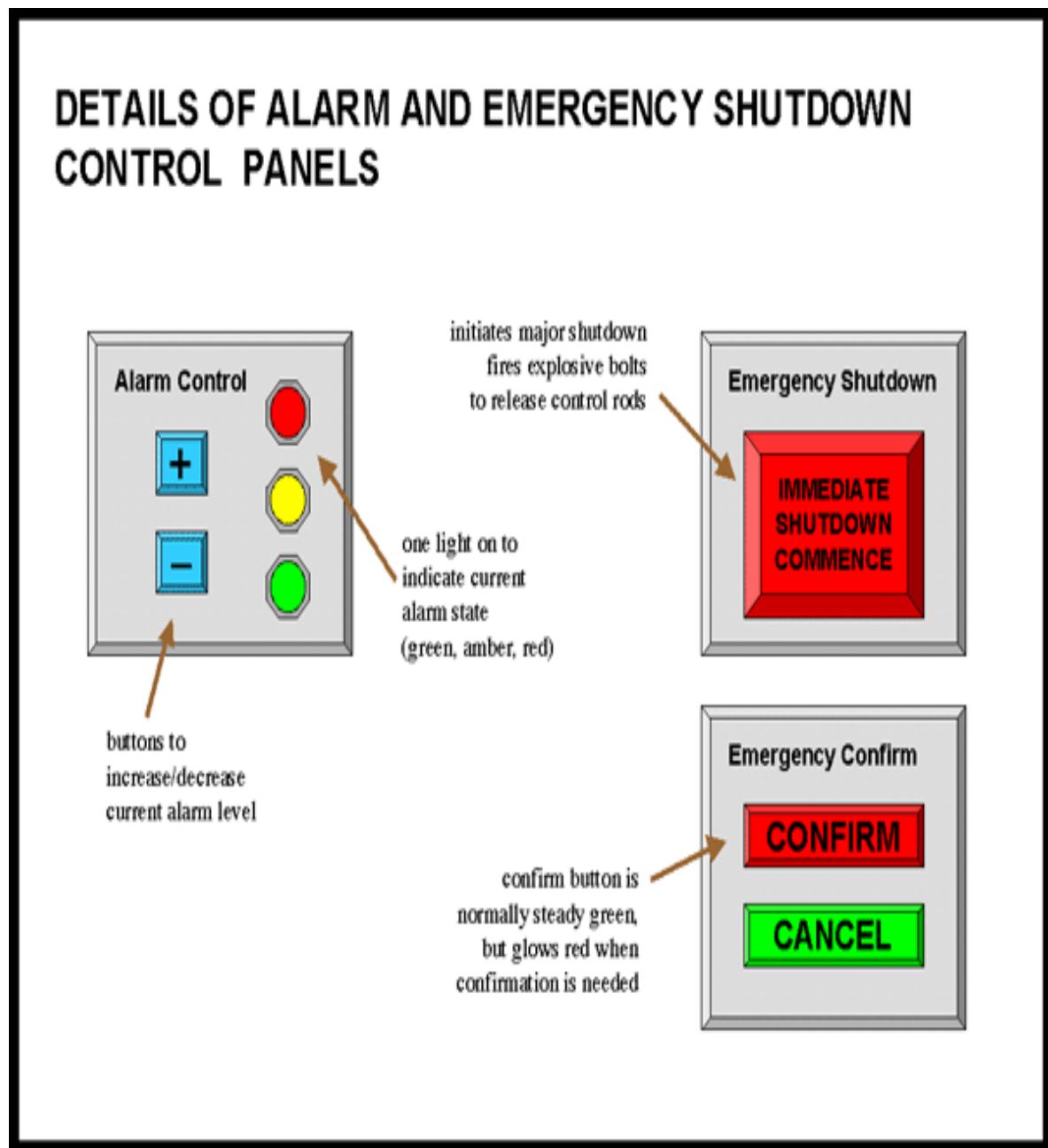


Figure CS.2 - alarm and emergency sub-panels

As we know, the tone is softer, it shows more safely. Because red is easy for people to recognize and to pay attention to dangerous situation. According to physics, RED, YELLOW and GREEN have descending wavelength, so we see red first, then yellow and finally green. In addition, red is like blood color, often symbolizes danger. Yellow is sunlight color, which expresses a reversed and respectful attitude, should be chosen to alter people more carefully. Green is a natural grass color, brings to relaxed and comfortable feeling.

Layout

NUCLEAR REACTOR MAIN CONTROL PANEL

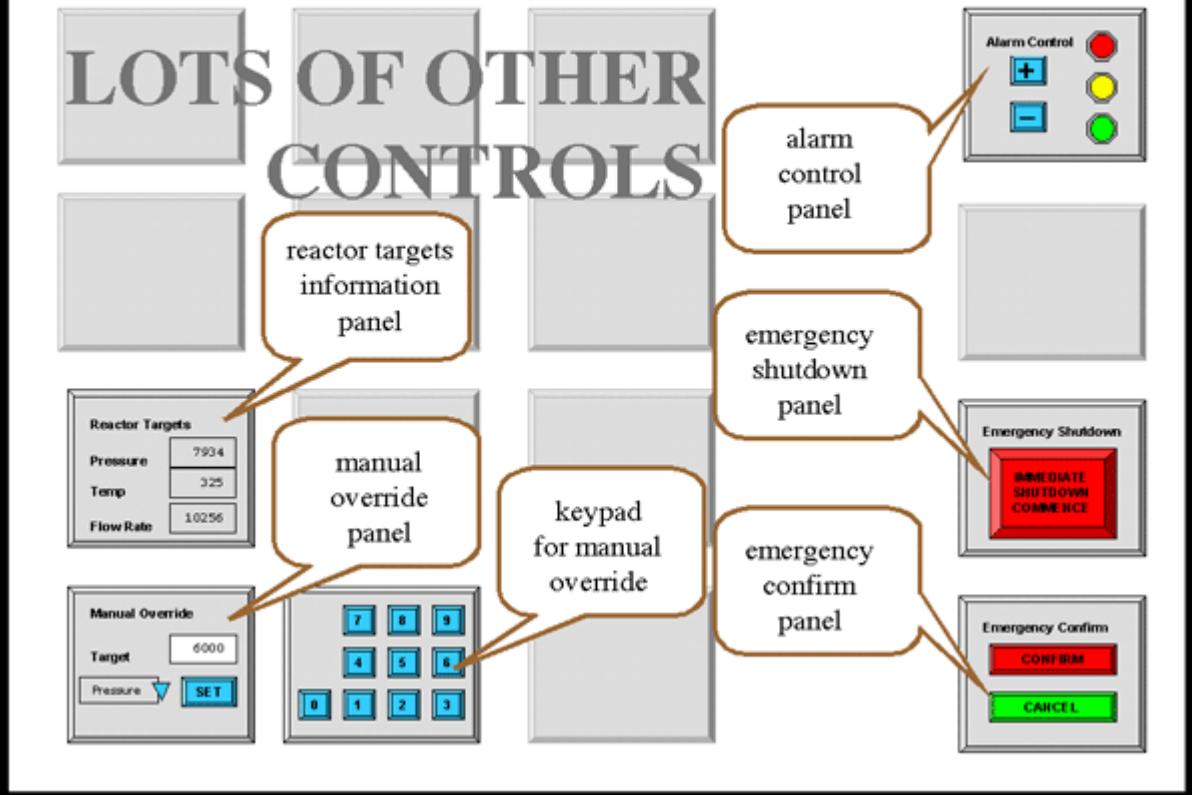


Figure CS.3 - reactor targets display and manual override

The manual override panel is reasonable. Because if the display is higher the panel, our hand will cover the display when we manipulate. Numeric Keypad is on the left side of Manual Override Panel, which is logical because most of people use the right hand to type numbers, similar to the numpad on the computer keyboard.

The size of Emergency Shutdown button is so big and red is reasonable because it causes great attention and vigilance for workers about serious consequence after pressing this button.

The alarm control panel is higher for necessary. It may be too high, which leads to out of the eyes and can't monitor the alarm state well.

Emergency scenario

Working through the accident scenario, we explored various problems arised and from those problem, suggestions have been drawn. The dangerous case happens through 25 steps as below:

1. “Jenny notices the core reaction rate has risen very rapidly”.
2. “she realizes she must immediately change the reactor target pressure to correct this”, she must interact with the KfMO and MOP in order to reduce the pressure level.
3. “she goes to the Alarm Control Panel on the far right of the main reactor control

panel and presses '+' twice (as it is starting off in green state)", ... if she presses the ‘+’ twice, the alarm level will increase; if the pressure is satiable, the alarm level should be low...

4. “the Emergency Confirm button glows red”, she should check the pressure level and alarm level carefully before confirm.

5.” she moves across to the Manual Override panel on the far left of the main reactor control panel”.

6. “she selects 'Pressure' from the pull down on the Manual Override panel”, It is easy to be mistaken 3 categories.

7. “she types the new value '6000' using the keypad”.

8. “she notices that the number on the Reactor Targets panel has not changed”, she may forget to press “set”, or she may choose wrong category.

9. “she realizes she forgot to press the SET button on the Manual Override panel”; Yes, she did; the “set” button may be bigger

10. “she presses the SET button”, ... OK

11. “the value still doesn't change”, ... Maybe 6000 is not enough?

12. “an automatic audio warning sounds "60 seconds to core meltdown"”, god blessing you.

13. “she presses the SET button repeatedly”, it won’t not effective.

14. “still the value doesn't change”; Yes, it is.

15. “she starts again, selects 'Pressure' from the pulldown, types 6000 and presses SET”, she finally calms down.

16. “still the value doesn't change”, ... God will cover you.

17. “the audio warning says "30 seconds to core meltdown"”

18. “Jenny runs across the room to the Emergency Shutdown panel”, I thought she will run away, she so brave. And I think the room is a bit far (for the elder).

19. ““20 seconds to core meltdown””.

20. “she presses "Immediate Emergency Commence" button”. Well-training

21. “the emergency confirms button glows red”, very easy to see!

22. ““10 seconds to core meltdown””. Good job, auto voice

23. “she presses the "Emergency Confirm" button”. Well-training.

24. “she hears the crash of the explosive bolts sending the control rods into the reactor”, ... What was that mean?

25. “the audio system announces "reactor shutdown successful"”, good job, girl. And

you too, auto voice

Besides, we also suggest potential ways of improving the interface to avoid a similar problem recurring.

- First, the alarm control panel should be in the bottom, the ESP should be in top and ECP below it.
- The “confirm” button should be green because “green mean go” (traffic light) and “cancel” should be red.
- The “set” button should be in the KfMO, and bigger than other.
- The chosen category should be different color to other category
- The KfMO and MOP should Switch places.
- The RTIP should be in the middle of this line.

IV. CONCLUSION

The human centered approach in complex industrial system design, evaluation and validation should be applied in the design process in which the system is produced, and in the system itself. In this research we investigate the human system interface of a nuclear power plant simulator to compare design solutions during the early design phase. The methodology used was based on observations of the operators’ performance in the LABIHS simulator. Performance evaluations based methods can be used considering the fact that the appropriateness of a given system expresses itself in the quality of the overall performance of the system is assessed. Normally, performance evaluation is something that is carried out towards the end of a given design process. However, within limitation of this study, we only give out some subjective suggestion to improve the current interface design for better use, especially in urgent cases.

REFERENCES

- [1] A. Dix, J. Finlay, “Human Computer Interaction” Third Edition. Pearson; 3 edition (October 10, 2003) Online resource, Available at URL [“https://www.hcibook.com/e3/scenario/nuclear/”](https://www.hcibook.com/e3/scenario/nuclear/)
- [2] Paulo V. R. Carvalho, Jose O. Gomes, Marcos R. S. Borges. “Human centered design for nuclear power plant control room modernization” CEUR Proceedings 4th Workshop HCP Human Centered Processes, February 10-11, 2011.

E-MARKETING TECHNIQUES APPLIED IN MOVIE LAUNCHING EVENT

SVTH: Đặng Cẩm Tú, Đào Văn Quang, Nguyễn Như Quỳnh

GVHD: ThS Nguyễn Thùy Linh

Abstract—Nowadays, the role of offline or traditional marketing strategies have been taken over by online one. It makes the role of e-business, especially e-marketing become more and more important. This project is the result of our cooperation with Gmovie for the second release of the series "Huong mat tua khoi suong" in Vietnam. During the project, we have applied both IT and non-IT knowledge in different e-marketing strategies that can boost the launching event to as many movie fans as possible. It is a case study then from that we can apply techniques in similar events or generalize to effective online marketing strategies.

Keyword— Instagram account, fanpage, post, schedule, ads.

I. PROBLEM

"Huong mat tua khoi suong" is a 2018 fairytale Chinese drama; it is based on the novel "Heavy Sweetness, Ash-Like Frost" and tells the story of the undying love between two immortals. Ever since the movie was released, it has achieved much success in the world. However in Vietnam, the film has not really attracted many fans because of the limited number of publishers and the quality of the translations is not really good. Therefore, Gmovie had decided to buy the copyright of the series "Huong mat tua khoi suong" to release at the website gmovie.vn. That promises to bring great movie viewing experience to millions of fans with higher quality. This project is the result of our cooperation with Gmovie for the second release of the series "Huong mat tua khoi suong" in Vietnam.

The company allow us to representative them to upload and deliver all copyrighted contents of the movie on social network channels. Time of campaign will last about 1.5 months, and take at least 1 million dong to run advertisement on social network.

II. OUR ACTIVITIES

A. E-marketing plan

1) Create and maintain Instagram account – Facebook's fanpage

- Create an account on Instagram
- Create a Facebook's fanpage
- Prepare image and text for attract the attention of young people
- Cut short video in each chapter and post
- Prepare content about series: director, characters, timeline, backstage,
- Design 1 post to run advertisement with budget: 900.000 VND for ads on Instagram and Facebook

2) Result

From those kinds of activities, we expect to see the following results:

Immediate effect measurement:

- Our customer has more details about series Huong Mat Tua Khoi Suong: Characters, director, backstage....
- Information about time and place our company release this series
- Watching series with higher quality (1080p)

Long-term effect measurement:

- Our customer will receive an Instagram account with 2000 followers and a Facebook's fanpage with approximate 500 likes.

B. Member workload breakdown

In order to control the progress of project, we have designed a workload breakdown as well as time table to manage teamwork effectively.

Date	Mission today	In charge	Time Duration
Starting day 28/10/2019	Create Instagram account	Nguyen Nhu Quynh	8:00am – 9:00am
28/10/2019	Upload posts	Dao Van Quang	11:00am -20:pm
28/10/2019	Attract Instagram's users	Dang Cam Tu	all day
29/10/2019	Follow followers (to attract reach)	all of member	all day
30/10/2019	Upload posts	all of members	3 times
31/10/2019	Upload content about male characters	Nguyen Nhu Quynh	8:00-9:00 am
31/10/2019	Upload posts about female characters	Dao Van Quang	11:00-12:00 pm
31/10/2019	Attract Instagram's users	Dang Cam Tu	all day

1/11/2019	Upload post	Nguyen Nhu Quynh	8:00 am- 11:00 am - 8:00 pm
2/11/2019	Upload posts	Dao Van Quang	8:00 am- 11:00 am - 8:00 pm
3/11/2019	Upload posts	Dang Cam Tu	8:00 am- 11:00 am - 8:00 pm
4/11/2019	Lost Instagram account		
5/11/2019	Use back up Instagram's account	Dang Cam Tu	all day
5/11/2019	Run ads on Instagram	Dang Cam Tu	all day
6/11/2019	Upload posts	Dao Van Quang	11:00am -20:pm
7/11/2019	Attract Instagram's users	Dang Cam Tu	all day
8/11/2019	Upload posts	Dang Cam Tu	11:00-12:00 pm
8/11/2019	Upload posts	Dao Van Quang	11:00am -20:pm
8/11/2019	Upload posts	Nguyen Nhu Quynh	8:00am – 9:00am
9/11/2019	Run ads on Instagram	Dang Cam Tu	all day
9/11/2019	Upload posts	Nguyen Nhu Quynh	8:00 am- 11:00 am - 8:00 pm
10/11/2019	Upload post	Dao Van Quang	8:00am – 9:00am
10/11/2019	Upload post	Nguyen Nhu Quynh	11:00-12:00 pm
10/11/2019	Upload posts	Dang Cam Tu	5:00-6:00 pm

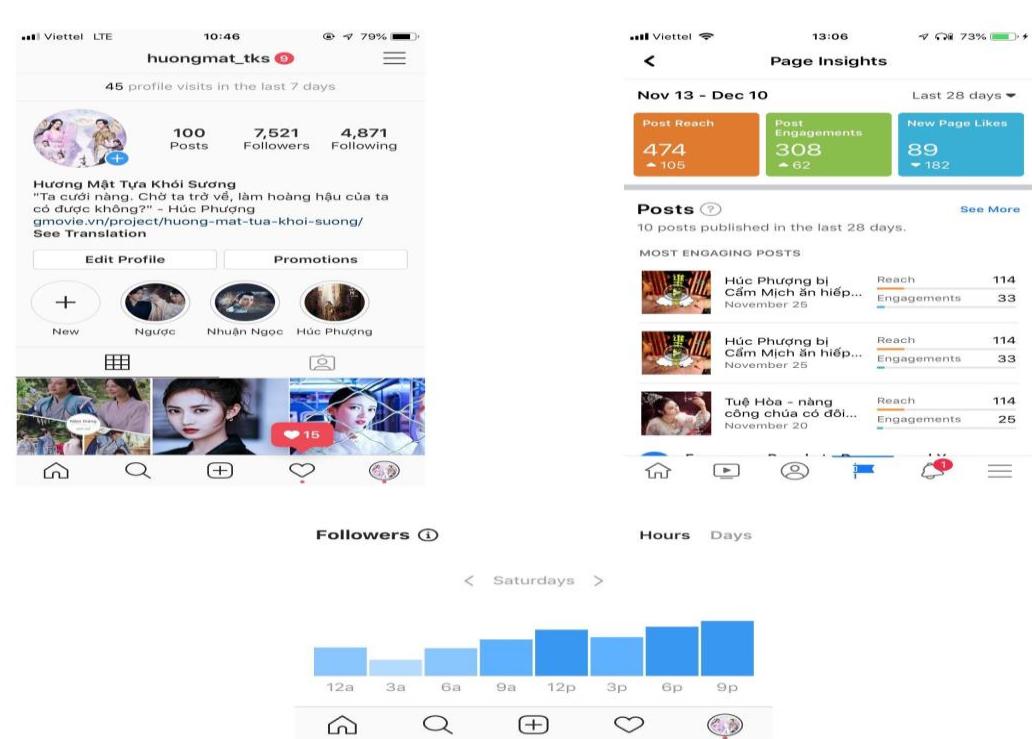
11/11/2019	Upload posts	Dao Van Quang	8:00am – 9:00am
11/11/2019	Upload posts	Dang Cam Tu	11:00-12:00 pm
Repeat until the end of project
Ending day	Write report	All member	2 days

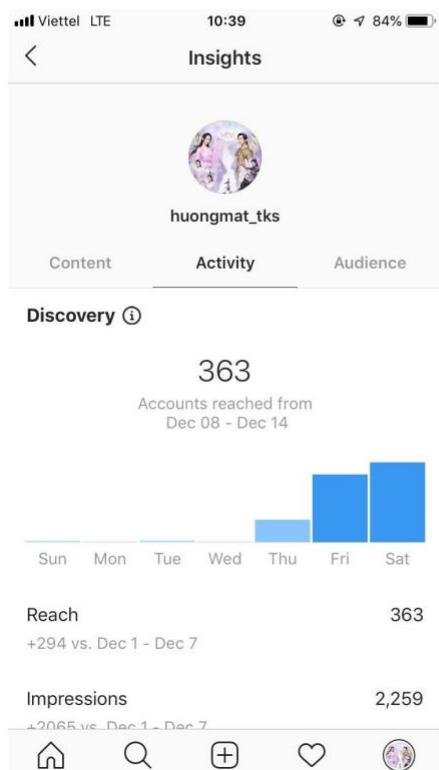
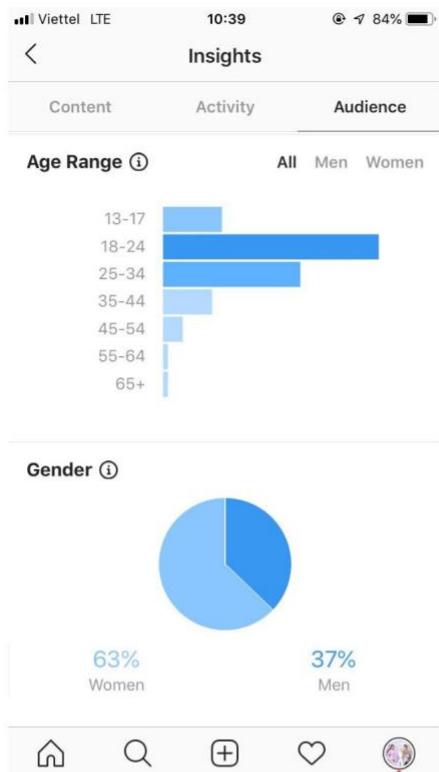
* Each job is one row. The number of row can range from 35 to 45 due to time duration of your project, or more if you have more jobs per day.

III. RESULT

C. Statistics result

We have achieved - 7500 followers on Instagram account and 400 followers on Facebook's fan page. The average reach of each post is over 100 people and about 30% of them have engaged to the content. Ho Chi Minh City contributed the highest number of fan, and people from 18 to 24 is majority of audience.





D.Applied theory and knowledge

We have some marketing tips on Instagram and Facebook and they helped us to succeed:

1) Switch to an Instagram Business Profile:

Besides visiting your website, your followers can click on the contact button to get in touch with you.

2) Cross-Promote on Other Social Media Channels:

One of the most effective ways to increase your reach is to promote your Instagram on other social media platforms. You may already have a fan/follower on Facebook, Twitter, LinkedIn, etc who know you and trust you. So why not invite them to follow you on Instagram as well?

3) Follow so many people who have the same interesting in your content:

By following them, they will follow back if they know you and see how interesting your contents

4) Use hashtags strategically:

Every social media marketer knows the value of hashtags and the role they play. But the key to finding success with them on Instagram is to use them strategically. This way you'll be able to get more people to see your content, connect to your brand and improve the overall reach

5) Start regularly posting “balanced” content:

- Publish original material 70 percent of the time.
- Post content relevant to the interest of your audience 20 percent of the time.
- Create self-promotional content only 10 percent of the time.

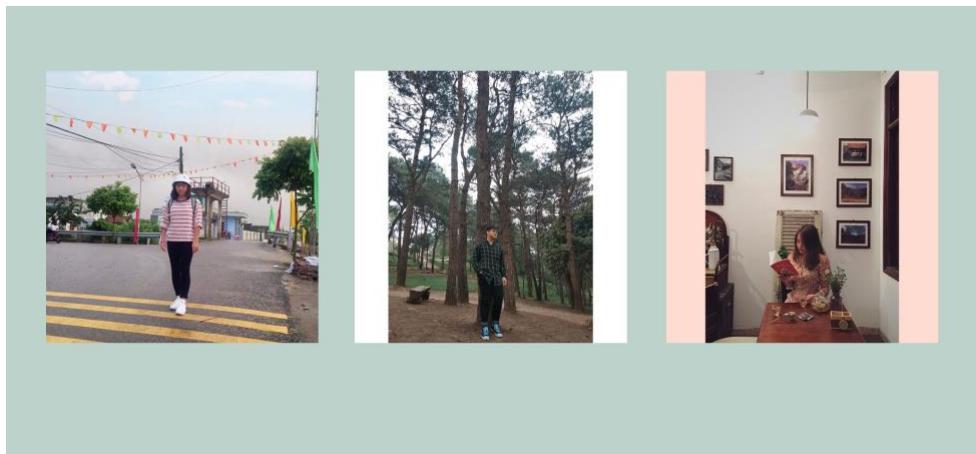
6) Post video more often:

- 85% of Facebook users prefer watching videos on mute. This implies that captioning your videos will be vital to engagement.
- 80% of users are irritated when videos auto-plays on their newsfeed, so make sure that you're setting up your videos correctly when you post.

7) Time your post correctly:

- We always post our content on time.
- The time when most of your Facebook followers are online.
- How well a particular piece of content does throughout the day.
- Information on which of your posts is creating the most engagement.

E. Teamwork revision



1) Member 1: Nguyen Nhu Quynh

Nguyen Nhu Quynh created Instagram and Facebook accounts for this project and maintained at the same time. Besides, Quynh has created a few photos for fan page to be livelier. She contributed in posting content on the fan page in a timely and complete manner. Besides, she attracts fan page likes from friends and colleagues very effectively (about 45%). However, she is not good at managing secure accounts.

2) Member 2: Dao Van Quang

Dao Van Quang is a designer, writer and a permanent management. He can use many tools like Photoshop, Illustrator, Sony Vegas and some mobile apps to create videos, banners and attractive posts. Besides, he always sends notices to other members so that each member can complete their work on time.

3) Member 3: Dang Cam Tu

Dang Cam Tu is the person who connects the group with the business and talks about the implementation of the project with gmovie.vn business. Cam Tu is responsible for planning the engagement as well as followers of both accounts. Besides, Cam Tu is also responsible for running ads for Instagram accounts. However, she is not good at creating content and images for fan page posts.

IV. CONCLUSION

The project has shown the effectiveness of communication on social networks instead of just using the traditional way of marketing in case of Gmovie. To build a complete marketing communication system, consistent with the business situation, market trends, target customers ... of each business sector of the company is not easy but the result of our timely efforts and recognition of the market and the employees of the Company. Thereby it is possible to maintain and maintain the efficiency of business operations, gradually expanding the customer base, attracting partners, to continue to achieve high and stable profit growth. Hopefully, in the near future, the company will grow strongly and achieve new successes, rise

to become one of the leading film publishing businesses in Vietnam! Thank Gmovie Company for providing enough documentation so that we can implement this project. We would like to express my sincere thanks to Ms. Nguyen Thuy Linh, who guided us to implement this project As a student, with our own efforts, we have tried to perfect this project in the best way, but cannot avoid the shortcomings. Therefore, we are looking forward to the input of Ms. Nguyen Thuy Linh who directly guides us to complete this project.

REFERENCES

- [1] Svedic, Z. “E-MARKETING STRATEGIES FOR E-BUSINESS,” Faculty of Business Administration SIMON FRASER UNIVERSITY. 2004.
- [2] NEILPATEL, “**10 Powerful Instagram Marketing Tips (That Actually Work)**” Online resource, Available at URL “<https://neilpatel.com/blog/instagram-marketing-tips/>”, Jan. 2018.

DIABETES PREDICTION ON WOMEN USING DATA MINING TECHNIQUE

SVTH: Bui Thi Anh Nguyet, Vu Thi Lieu, Le Thi Bich Ngoc

GVHD: ThS Bùi Quốc Khánh

Tóm tắt: Bệnh tiểu đường là một căn bệnh mãn tính gây ra bởi lượng đường trong máu quá cao và căn bệnh này gây ra những mối đe dọa nghiêm trọng cho sức khỏe con người. Với sự phát triển mạnh mẽ của lĩnh vực khoa học dữ liệu nói chung và lĩnh vực học máy nói riêng, rất nhiều những kỹ thuật của các lĩnh vực này được ứng dụng rộng rãi trong lĩnh vực y tế giúp thu được nhiều thông tin có giá trị. Mục tiêu nghiên cứu của chúng tôi là sử dụng các thuật toán khác nhau trong lĩnh vực học máy để dự đoán khả năng mắc bệnh tiểu đường ở một người phụ nữ. Dữ liệu được lấy từ bộ dữ liệu Pima Indians Diabetes và chúng tôi phân tích dữ liệu này với 3 thuật toán trong học máy: Logistic Regression, Decision Tree, Random Forest để so sánh hiệu năng của từng model. Kết quả phân tích cuối cùng đã chỉ rằng Random Forest đưa ra dự đoán chính xác nhất trong số ba thuật toán được thử nghiệm.]

Abstract: Diabetes is a chronic ailment characterized by the degree of blood sugar which poses incredible threats to human health. With the strong growth of machine learning, diverse techniques have been employed in many aspects of medical health which help gain valuable information. The objective of this study is to structure a model which can prognosticate the probability of diabetes in chosen diabetic women. The Pima Indian diabetes database (PIDD) was obtained from the UCI repository used for analysis. In this research, three machine learning classification algorithms, correspondingly Logistic Regression, Decision Tree and Random Forest have been accomplished. The performances of all three algorithms are estimated on various metrics. The final process has shown that Random Forest has the most perceptible results out of three algorithms when all the attributes were applied.

Keywords: Decision Tree, Google Colab, Logistic Regression, Pima Indian diabetes database (PIDD), Pycharm, Random Forest

I. INTRODUCTION

No longer confined to rich countries, diabetes has become a common disease worldwide. According to 2014 statistics from the World Health Organization (WHO), diabetes is affecting 422 million people globally. Without increased awareness and timely intervention, diabetes will become one of the seven leading causes of death by 2030. There are 3.7 million deaths caused by this disease each year. According to a study in the journal Annals of Internal Medicine, between 1971 and 2000, mortality rates among women with diabetes remain alarming. In addition, the difference in mortality between women with diabetes and non-diabetics more than doubled. Diabetes in women is different from men because of some reason. Firstly, women are often less likely to be treated for cardiovascular risk factors and diabetes-related disorders. Next, the complications of diabetes in women are harder to diagnose. Thirdly, women often have many different types of heart disease than

men. Moreover, hormones and inflammation also show very differently in women. Thus, our team decided to analyze a woman's diabetes data set to gain useful information to prevent these dangerous illnesses in women. There are a number of researches executed with different model classifiers and led to valuable results. In this paper, our team conducted using 3 classification algorithms such as Logistic Regression, Decision Tree, and Random Forest to training and testing for the diabetes data set to compare the strength of each one. In addition, an application would be created to visualize the results of research and help everyone diagnosed the diabetes disease by each classification. These results also would be useful for the classification of diabetes complication data. It would be discussed in the following sections.

II. METHODOLOGY

In this research, we used 3 kinds of models to predict the risk of diabetes on women such as Logistic Regression, Decision Tree, and Random Forest.

Logistic Regression Model: Logistic Regression is a popular statistical technique to predict categorical outcomes (binomial or multinomial). The predictions of Logistic Regression are the form of probabilities that could happen by an event. The purpose of logistic regression is to explore the best fitting model to show the relationship between the dichotomous characteristic of interest (dependent variable) and a set of independent (predictor or explanatory) variables. Moreover, Logistic Regression performance is good when the data set is separated linearly. It is less prone to over-fitting but can be over-fit in high dimensional data sets. To avoid over-fitting, Regulation can be used. In addition, Logistic regression could implement easily and train efficiently.

Decision Tree Model: Decision Tree is one of the most popular supervised learning algorithms in data mining classification because it is easy to use and understand. A decision tree is the building block of random forest and is an intuitive model. We can understand in a way that the Decision Tree is a collection of yes / no questions (considered nodes) about data until the algorithm obtains a model to make predictions similar data. However, Decision Tree is easy to get the over-fitting problem which affects the accuracy of the model. In Decision Tree, over-fitting occurs when a tree is trained and designed to match with the data in the training data set and gives high accuracy. However, when using the model obtained after training with the data set, the accuracy is not high with other data sets. There are two ways to avoid over-fitting in decision trees: pruning and random forest. Pruning is a regularization technique to avoid over-fitting for the decision tree in general. And the second way is we can combine many decision trees into a single ensemble model known as the random forest.

Random Forest Model: The random forest combines hundreds or thousands of decision trees, trains each one on a slightly different set of the observations, splitting nodes in each tree considering a limited number of the features. The final predictions of the random forest are made by averaging the predictions of each individual tree [1]. Random Forest has a different approach to Decision Tree. It is a collection of many decision trees and treats each decision tree as an independent voter (like a real election). The decision tree with the most votes will win - this means that the tree with the best results will become a model.

III. IMPLEMENTATION

A. Data Resource

The data was gathered and made accessible by "National Institute of Diabetes and Digestive and Kidney Diseases" as a major aspect of the Pima Indians Diabetes Database. A few limitations were put on the determination of these examples from a bigger database. Specifically, all patients here have a place with the Pima Indian legacy (a subgroup of Native Americans) and are females of ages 21 or more. There are totally 768 women in the dataset including 8 features such as Pregnancies, Glucose, BloodPressure, SkinThickness, Insulin, BMI, DiabetesPedigreeFunction, Age, and one target label called Outcome which decides if that woman is diabetes or not.

B. Data Preparation

The data has been processed cleanly and the total remaining after processing is 768 cases. In particular, there is a clear division of the results of the data: out of 768 cases, out of every 500 uninfected cases, 268 cases have diabetes. This division is shown in the figure below:

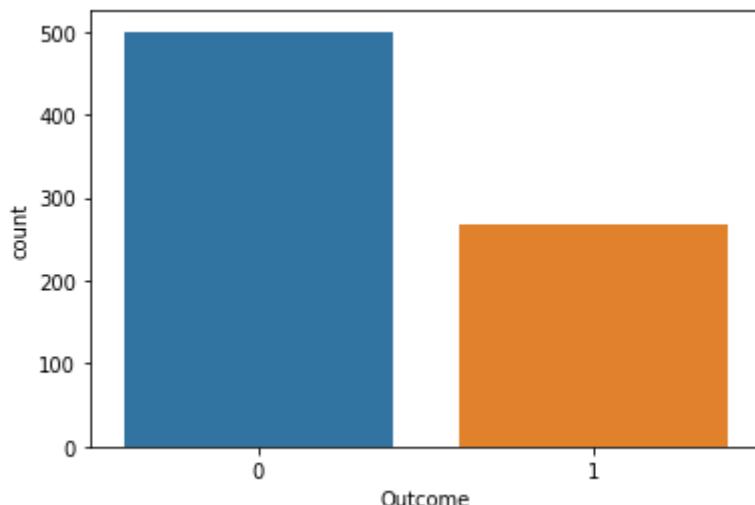


Figure 1: The difference between diabetes (1) and no diabetes (0)

After understanding about the data, start splitting the data into two sets: the training data

set and the test data set in a 7-3 ratio, meaning 70% of the data for training and the remaining 30% for testing.

C. Software Use

Our team used Google Colab and Pycharm as the main software for this research because of

their great benefits.

Google Colab: Google Colab is a free cloud service and now it supports free GPU [2]. For those who are studying and researching on deep learning, the use of GPU is extremely necessary. And of course, instead of spending money on a GPU, we can use it for free with Google Colab, which provides free GPU and is an application for the field of data science in general. Google Colab allows users to develop deep learning applications on popular libraries such as PyTorch, TensorFlow, Keras, or OpenCV. It is an incredible free browser that allows users to process large amounts of data, and train model on machine learning for free. It can also import a lot of libraries, use algorithms easily, and give fast and accurate results. With the undeniable benefits of Google Colab, so we use it as a tool to process data, train and test models, plot models using scikit-learn library to have the most intuitive and clear view of the results for each model.

Pycharm: To be able to code and run programs with the Python language, we need to install the Python programming environment for laptops and can install additional IDE to support us to

code faster. And Pycharm is the perfect choice - an IDE developed by JetBrains. Pycharm offers

a lot of smart features like code completion, indentation, error checking and suggestions for code

completion correctly. This IDE also supports extremely powerful libraries and frameworks, and

you can find almost libraries needed for Python programming. In this research, we created a simple website using Flask to predict a woman's risk of diabetes and Pycharm is the perfect choice to complete this website with their great features.

D. Data Analysis

After splitting the data set into training and test set, and selecting the model to use, we conducted training the training set with 3 models: Logistic Regression, Decision Tree, and Random Forest on diabetes data set. In addition, we used Grid Search algorithms to find the best parameters which make the model give the best results. To determine whether the model is really good or not, we take turns to calculate various metrics such as accuracy, confusion matrix, precision, recall, f1-score, and the results we got are quite high.

Accuracy: The simplest and most commonly used way is accuracy. This evaluation

method simply calculates the ratio between the number of correctly predicted points and the total number of points in the test set. In the diabetes data set, we have 230 cases separated as test data. Three models will use the algorithm to calculate the accuracy of the test data set and the result of Logistic Regression, Decision Tree, Random Forest obtained above are 0.788, 0.744, and 0.788 in the range of 0 to 1 respectively , which means that the accuracy of the model when testing data is 78.8%, 72.7%, and 78.8% respectively. Looking at to the statistic result, we can easily recognize that Random Forest and Logistic Regression have the same accuracy score and they're more accurate than Decision Tree.

Confusion Matrix: The calculation using accuracy as above just tells us what percentage of the data is classified correctly without specifying how each type is classified, which class is best classified, and the data of which class is often misclassified into another class. To be able to evaluate these values, we use a matrix called confusion matrix. Basically, confusion matrix shows how many data points actually belong to a class, and is predicted to fall into a class [3].

The table below shows a standardized confusion matrix that shows how many percentages of the instance are categorized into the correct diabetes class and non-diabetes class, how many percent of the instances are categorized into the wrong diabetes class and non-diabetes class.

Logistic Regression	Positive (1)	Negative (0)
Positive (1)	0.59	0.41
Negative (0)	0.11	0.89

Table 1: Confusion Matrix of Logistic Regression

Looking at the table above, with the non-diabetes class (0) -we can see that 89% of the instances are categorized into the non-diabetes class and 11% of the instances are assigned to the wrong diabetes class. As for the diabetic group (1) - only 59% of cases are assigned to the correct diabetes group, while up to 41% of the instance is assigned to the non-diabetic group - a fairly high percentage .

However, with Decision Tree, the percentage assigned to the wrong class is not too high but not too low. Those statistics are shown in the table below:

Decision Tree	Positive (1)	Negative (0)
Positive (1)	0.63	0.37
Negative (0)	0.19	0.81

Table 2: Confusion Matrix of Decision Tree

For non-diabetes class (0), the percentage of instances assigned to the correct class is 81%, while the remaining 19% is the percentage of instances assigned to the wrong diabetes class. As for diabetes class (1), only 63% of instances are assigned to the correct diabetes class, but the percentage assigned to the wrong non-diabetes class is 37% which is smaller than Logistic Regression.

Random Forest	Positive (1)	Negative (0)
Positive (1)	0.64	0.36
Negative (0)	0.13	0.87

Table 3: Confusion Matrix of Random Forest

Looking at Table 3 above, we can see that with the Random Forest, the percentage of instances assigned to the correct non-diabetes class and assigned to the wrong diabetes class seem to be between the Logistic Regression and Decision Tree models. For non-diabetes class (0), 87% of instances are assigned to the right class and 13% of instances are assigned to the wrong diabetes class. This is slightly better than the Decision Tree but slightly smaller than the Logistic Regression. As for diabetes class (1), 64% of the instance is assigned to the correct class while the remaining 36% is assigned to the non-diabetes class. Compared to the 41% instance assigned to the wrong class of Logistic Regression and 37% that Decision Tree brings, this number has actually improved insignificantly.

Precision-Recall: With the classification problem where the data set of class is very different, there is an effective magic that is often used as Precision-Recall. With diabetes data set, the difference in outcome between diabetes and non-diabetes is quite large (Figure 1), so we used Precision-Recall to analyze the data. For this data set, we consider diabetes is positive and non-diabetes is negative.

With a way of identifying a positive class, Precision is defined as the ratio of true positive points among those classified as positive ($TP + FP$). Recall is defined as the ratio of true positive points among those that are actually positive ($TP + FN$). High accuracy means that the accuracy of the points found is high. A high recall means that the True Positive Rate is high, meaning that the rate of missing positive points is low. Returning to the diabetes data set, 3 models calculated precision and recall data and gave different results. First, with Logistic Regression, the results of precision score and recall score are 0.776 and 0.743 respectively. However, by Decision Tree, these numbers have decreased significantly with 0.719 as precision score and 0.718 as recall score. With the last model being the Random Forest, these two numbers are a bit different. With the precision score, Random Forest gave us a figure of 0.769 - this score is slightly smaller than the Logistic Regression but much larger than the Decision Tree. As for Recall score, this model returns the highest result with 0.754 compared to Logistic Regression and Decision Tree. However, calculating the precision

score and recall score alone is not enough. A good classification model is one that has both Precision and Recall as high as possible, as close to 1 as possible. There are two ways to measure the quality of a classifier based on Precision and Recall: Precision-Recall curve and F1-score.

F1-score: The F1-score is defined as the weighted harmonic mean of the test's precision and recall [4]. This score is calculated according to the formula:

$$F_1 = 2 \frac{\frac{1}{\text{precision}} + \frac{1}{\text{recall}}}{\frac{1}{\text{precision}} + \frac{1}{\text{recall}}} = 2 \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

F1-score is directly proportional to the precision score and recall score. F1 is high when both precision score and recall are high, and the higher F1 is the better the classifier. With diabetes data set, all 3 models calculate f1-score and give different numbers together. However, the Random Forest has the highest f1-score of 0.76, while 0.753 is the f1-score obtained by the Logistic Regression, and finally the Decision Tree with 0.718 on f1-score.

IV. DISCUSSION

After using this algorithm, we found that Logistic Regression return the best results when ‘C’= 38, Decision Tree with ‘max_depth’= 2 , but Random Forest just got the best results when ‘random_state’=0 where we’re not used Grid Search. The table below compares the data on all 3 models and from there we have the basis to compare which model is the best among the 3 models and which model is most likely to accurately predict those 3 models:

Metrics	Logistic Regression	Decision Tree	Random Forest
Accuracy	0.788	0.744	0.788
Precision	0.776	0.719	0.769
Recall	0.743	0.718	0.754
F1-score	0.753	0.718	0.76

Table 4: Performance of proposed algorithm on testing data set

Looking at the table above, we can get the most intuitive view of the effectiveness of each model. With those figures, we can see that the Random Forest has the highest results, such as accuracy score, recall score, and F1-score. In contrast, Decision Tree is the model that brings the lowest results among the three models tested. Through the table above, we can somewhat know which model is more efficient classification and whether the model is reliable or not based on the metrics that we have researched.

Our team also tested the effectiveness of all three models with a simple website written in Python with the help of the Flask web framework. And the results are quite consistent with the analysis that we have previously analyzed.

The screenshot shows a web application with a teal header and footer and an orange central content area. At the top, it says "WELCOME, LET'S PREDICT DIABETES!". Below this is a form with eight input fields, each consisting of a small icon followed by a label: "Pregnancy Week", "Glucose Score", "Blood Pressure Index", "Skin Thickness Index", "Insulin Index", "BMI", "Heredity Index", and "Age". At the bottom of the form is a red "Predict" button.

G.

Figure 2: Diabetes Prediction Website

The figure above is our main website interface. This interface requires the user to input the necessary information corresponding to the features contained in the diabetes data set, and when the user clicks on the Predict button, the prediction results will show that she is likely to have diabetes or not. We tried with any row in the diabetes data set correspond with 8 features: *Pregnancies*, *Glucose*, *BloodPressure*, *SkinThickness*, *Insulin*, *BMI*, *DiabetesPedigreeFunction*, *Age* with the following statistics: 8, 124, 76, 24, 600, 28.7, 0.687, 52.

The results obtained after the prediction for each model are: Logistic Regression model returns the result that the person is not likely to have diabetes, Decision Tree returns the same result, while the Random Forest model returns the result that the person get risk of diabetes. When comparing the result with the actual value in the data set, only the Random Forest give the correct prediction. This is correspond with our analysis because following to Table 4, the Random Forest is the best model out of three models, so the result of the website is quite similar to what our team has studied. We continued testing three models with some other instances which show in the following figure:

Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Actual Outcome	Decision Tree	Logistic Regression	Random Forest
0	123	72	0	0	36.3	0.258	52	1	0	0	1
1	102	74	0	0	39.5	0.293	42	1	0	0	1
8	120	86	0	0	28.4	0.259	22	1	0	0	1
3	130	78	23	79	28.4	0.323	34	1	0	0	1

Figure 3: Testing instance

Looking at to the Figure 3 above, the Random Forest give the correct prediction for all 4 instances tested, that prove the most efficient of this model instead of Logistic Regression and Decision Tree. We also tested 3 models with lots of different instances and the average result was that the Random Forest gave the most correct predictions including both diabetes and non-diabetes. And this is perfectly consistent with the final result that we analyzed that the Random Forest gave the highest metrics of evaluating model, then the Logistic Regression and finally the Decision Tree.

V. CONCLUSION AND FUTURE WORK

Being a chronic disease, diabetes is considered as the culprit of numerous deaths which prompts the need of anticipating the probability of this ailment. Therefore, the use of three machine learning algorithms is conducive to predicting diabetes' risk. By comparing the results of three classifications, it can be concluded that Random Forest is the optimal solution among the three algorithms. It is hoped that this research will support doctors in diagnosing and treating diabetic women.

ACCEPTANCE TESTING TYPES AND PROCESSES

SVTH: Trần Thị Mai Hương, Trương Ngọc Quỳnh

GVHD: ThS Đặng Đình Quân

Tóm tắt— Nowadays, the software is applied in all aspects in order to gain better quality of life. It is essential that a problematic programming product must not be released, especially if it is expected to work in serious work such as medical or transportation. Correspondingly, testing is upgraded continuously to detect bugs and ensure software quality. Therefore, In this paper, we are going to discuss forms, processes, participators, and tools of a significant part of testing - Acceptance testing..

Từ khóa— Software testing, Red box Testing, Acceptance Testing

I. INTRODUCTION

In recent decades, the software industry grows rapidly, the number of products rockets up to meet social needs. Software products become an integral part of life, from electric games to business applications. Most people suffered disappointed when using a program that did not work as expected. Releasing a problematic software can lead to many problems, namely loss of money, time or business fame, and in the worst case, injury or death (ISTQB, 2018). To assess the quality of the software and reduce risk, software testing is engaged in the software development life cycle as a mandatory phase.

According to Guru99, in order to enhance testing systematic and easily identify all possible test cases at a particular level, tests are grouped based on where they are added in SDLC or by the level of detailing they contain. There are mainly four testing level: Unit Testing, Integration Testing, System Testing and Acceptance Testing.

This research aims at studying acceptance testing- one of 4 test levels, focusing on types, processes, participators, and tools.

II. OVERVIEW OF TESTING

I. Test level

According to Guru99, in order to enhance testing systematic and easily identify all possible test case at a particular level, tests are grouped together based on where they are added in SDLC or the by the level of detailing they contain. There are mainly four testing level: Component Testing, Integration Testing, System Testing and Acceptance Testing.

1) Component Testing

Component Testing, also known as module or unit testing, aims to test each part of the software separately, checks whether that component is fulfilling functionalities. Developers take responsibility for this kind of testing.

2) Integration Testing

Different software units are combined together and test as a whole to ensure the integrated system's readiness for system testing. The purpose of this level is checking the data

flow from one module to other modules, and this level is performed by testers.

3) System Testing

On a complete, integrated system, both functional and non-functional requirements are evaluated to verify that the system meets the specification. This test is performed by testers.

4) Acceptance Testing

Like system testing, acceptance testing typically focuses on the behavior and capabilities of a whole product or system to assess the system's readiness for deployment and use by end-user, basically done by the user or customer. However, other stockholders can joint in this process.

J. Testing's contribution to success

It is popular in the program industry that delivered software and system cause failures or under stakeholders' needs. However, using testing in the development life cycle in an appropriate way can reduce the frequency of such problematic deliveries.

1) Requirement

If requirement reviews or user story refinement has testers involved in, defects in these work products can be detected effectively, helps decrease the risks incorrect or untestable features being developed (ISTQB, 2018). Therefore, cost and time will not be wasted.

2) Design

If system designers and tester cooperate while design the system can increase each party's understand of the design and how to test it, thus reduce the risks of fundamentals design fault and enable tests to be identified at an early stage.

3) Development

Having testers support developers while coding can increase each party's understanding of the code and how to test it, reduce the risk of bugs within the code and the tests.

4) Verify and validate

Testers verify and validate the software before releasing can defect missed failures and support removing the defects. This increases the likelihood that the software satisfies requirements and meets stakeholder needs.

To sum up, in every phase of development life cycles, the involvement of testers also leaps the quality of the product and contributes to success. Therefore, despite acceptance testing basically done by the customer, the testing team should join acceptance testing.

III. ACCEPTANCE TESTING

A. Definition

As assigned in ProfessionalQA, the main goal behind acceptance testing is to check

whether the developed software product passes the acceptance norms defined on the basis of user and business requirements, so as to declare it acceptable or non-acceptable for its use by the users. Acceptance testing is one of the last types of software testing performed over software or application. It is conducted by a pool of targeted users to ensure the readiness and quality of the system from the user's perspective, which allows the team to meet their needs and expectations.

Acceptance testing is also referred to as red box testing. Additionally, the acceptance tests are derived from the user story and are based on the acceptance criteria.

B. Types of acceptance testing

1) Alpha Testing and Beta Testing

Criteria	Alpha Testing	Beta Testing
<i>Member</i>	Testers	Clients/ End users
<i>Test location</i>	Developer's site	Clients/ End users location
	Reliability and Security Testing are not performed	Reliability, Security, Robustness are checked
<i>Test techniques</i>	Both the white box and black box	Black Box Testing
<i>Test environment</i>	Lab environment/ testing environment	Real- time environment
<i>Time</i>	Long execution cycle	Only a few weeks of execution
<i>Fix bugs</i>	Immediately	In future versions of the product
<i>Objectives</i>	Alpha testing is to ensure the quality of the product before moving to Beta testing	Beta testing also concentrates on the quality of the product, but gathers users' input on the product and ensures that the product is ready for real- time users.

2) User Acceptance Testing (UAT)

User acceptance testing of the system is typically focused on validating the fitness for use of the system by intended users in a real or simulated operational environment. The main objective is building confidence that the users can use the system to meet their needs, fulfill requirements, and perform business processes with minimum difficulty, cost, and risk.

3) Operational Acceptance Testing (OAT)

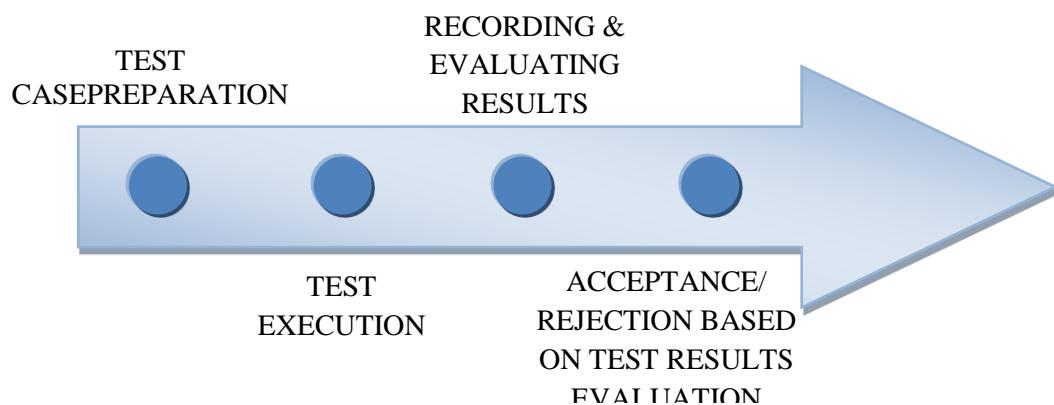
The acceptance testing of the system by operations or systems administration staff is usually performed in a (simulated) production environment. The tests focus on operational

aspects, and may include:

- Testing of backup and restore
- Installing, uninstalling and upgrading
- Disaster recovery
- User management
- Maintenance tasks

C. Testing in processes

1) Activities



- Test case preparation: Preparation of test cases, on the basis of combined requirements of both business and end-users.
- Execution: These test cases are executed on the software product.
- Documentation: Results are recorded and compared with the expected results, in accordance to specified acceptance criteria, for meeting the requirements.
- Evaluation: The number of pass or fail of these test cases decides the software suitability for getting accepted or rejected.

2) Acceptance Testing Criteria

There are three criteria:

- Functional Criteria.
- Non-functional Criteria.
- Performance Criteria.

D. Participants

1) The Users

a) Active Users

Active Users are the people who interact directly with the system, so they would have been key sources not only of functional requirements but also of performance, usability and other non-functional requirements as well. Much of what is built into the system is intended to meet these people's needs, so some (or many) of them should definitely be testers.

b) Passive Users

Passive users don't interact with the system directly but are impacted by the system none-the-less. For example, they may receive reports (e-mails or notifications from the system), or their jobs may include using the information the system produces, or they may supervise Active or Passive Users. The acceptability of the system for these people's purposes is no less important than for the Active Users, so some of these people should be testers as well.

2) Business Analyst

- BA responsibilities in acceptance testing are:
- Identifying Testers: Users (Active users/ Passive users), stakeholders...
- Planning Tests: Major focus of acceptance testing is to check how the system will actually be used in practice instead of finding bugs. Because of different purpose, planning for acceptance testing is significantly different from planning the other types of testing.
- Determining when to do acceptance testing: There is a mindset that acceptance testing must be performed at the end of the project, but doing this crucial evaluation too late might land up in significant risks. Issues found in acceptance testing can cause time delays or forcibly accepted products.
- Delivering business value through acceptance testing: Actively participating in UAT planning and execution is an important way for the BA to ensure that value is indeed delivered.

3) Quality analysis

- Review of acceptance criteria: The acceptance criteria must be clear, consistent, comprehensive, covers non-functional characteristics and provides measurable pass/fail criteria.
- Review of acceptance test cases: Previously defined acceptance criteria should be covered in the acceptance test cases.
- Traceability: Traceability between requirements/ user stories, acceptance criteria, test cases, and defects facilitates acceptance testing as it clarifies dependencies and provides simple access to related information.
- Coverage analysis based on traceability: If bi-directional traceability is established, it is possible to perform a systematic coverage analysis.
- Review of test report: Test reports should be clear, consistent and comprehensive.

They should contain all information provided by the tester to support decisions about the release.

E. Tool support

From both software testing domains and the business analysis, these tools for acceptance testing are derived (ISQTB, 2019).

Tool Type Usage for Acceptance Testing	Requirements
Requirements management tool	Description of acceptance criteria Traceability between tests and requirements Coverage analysis
Agile project management tool	Description of acceptance criteria Traceability between tests and user stories Coverage analysis
Business process management tool	Model business process and rules Analyze defect impact on business processes
Test management and automation tool	Manage acceptance tests and test execution campaign Manage test execution results
Model-based testing tool	Generate test cases from business process models Manage traceability between business process models, business rules, requirements and test cases
Defect / Incident management tool	Manage defect / incident lifecycle

IV. CONCLUSION

In conclusion, there are 4 levels of testing: Component, Integration, System and Acceptance Testing. Testing involvement in every phase all contributes to success of the application or system. This paper discusses 3 types of acceptance testing: alpha and beta testing, user acceptance testing (UAT) and operational acceptance testing (OAT). Testing process with activities, acceptance testing criteria, and participants also are studied. Finally, tools are reviewed. This research paper supply understandings that can be learned to upgrade acceptance testing.

PHP LOOSE COMPARISON & TYPE JUGGLING

SVTH: Phạm Đức Tùng, Đỗ Nguyễn Hoàng Ân

GVHD: ThS Đặng Đình Quân

Tóm tắt— PHP, one of the most prominent scripting programming languages along with JavaScript and Python, “especially suited to web development” [1] is “fast, flexible and pragmatic” [1]. Naturally, there is no type-safe restriction but loose comparison. In what was an effort to produce a more straightforward language to developer, variables are not meant to be declared with an explicit type. While this does compose a flawlessly coding experience, chances for application result in state of perplexing persists.

This paper aims to demonstrate the variables comparison mechanic and type coercion feature of PHP. Furthermore, this study enlightens one of the most crucial issues of rookie PHP programmers or even intermediate ones that undeviatingly runs the web application to be vulnerable.

Từ khóa— loose comparision, PHP, type juggling

I. INTRODUCTION

In programming, declaring a variable, which requires variable’s data type and the label or name of that variable, is a basic step for beginners to learn any programming languages. After that, variables will be initialized with a value with the same data type as declared in the variable declaration. If there are any problems, an exception will be announced to programmers. However, in PHP, there is no need to apply any type specification in the process of declaration, only the name of the variable and its value. Therefore, type juggling feature is created to deal with PHP variables. [2]

II. VALUES DIFFERENTIATION IN PHP

For developers familiar with compiled languages for instance Java, C#..., such languages provide type-safe flow, restrict programmers to use particular datatype at compilation time. The feature itself keeps the application a distance as far as possible away from bugs. However, considering the personal preferences, PHP comes with both loose comparison operator == and strict comparison operator ===.

A. Type checking

In order for your program to run, all the variable types must be verified. This Typing Checking may provide a static check (compile-time) or at dynamic check (run-time). Therefore, if a language lacks its typing rules strongly, one can refer to the process as Weakly typed, if not, as Strongly typed **Invalid source specified.** Type checking is to make sure that the program is “type-safe”, reduce the risk of type errors. For instance:

```
int AddTwoNumbers(int a, int b) {  
    return a + b;  
}
```

Normally, if programmers try to call that using:

```
int Sum = AddTwoNumbers(5, "5");
```

The compiler would throw an error, because programmers passing a *string*("5"), but the argument expects an integer.

Therefore, in a type-unsafe programming language like PHP, it is possible to use the function like this:

```
AddTwoNumbers(5, "5");
```

Although many languages can use the + operator for strings, integers and float together, but this often result in a type error because this expression is usually not meant to execute multiple data types.

Many programming languages are considered not to be type - safe, it means there are no action that prevent a type-error, and they are Dynamically typed languages.

1) Loose vs. Strict

a) Loose comparison

The tables **Invalid source specified.** below describe more other types, where anything in double quotes is string:

	TRUE	FALSE	1	0	-1	"1"	"0"	"-1"	NULL	array()	"php"	""
TRUE	TRUE	FALSE	TRUE	FALSE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE
FALSE	FALSE	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	FALSE	TRUE
1	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
0	FALSE	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	TRUE
-1	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE
"1"	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
"0"	FALSE	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE
"-1"	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE
NULL	FALSE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	FALSE	TRUE
array()	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE
"php"	TRUE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	FALSE
""	FALSE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE

Table 1 - Loose comparison

A loose comparison in PHP is an operation which use two equal signs (==). It can lead to some confusion:

```
echo gettype(0); // Outputs "integer"
echo gettype(FALSE); // Outputs "boolean"
echo gettype('guitar'); // Outputs "string"
```

Here, the confusion starts when comparing these types of data:

```
var_dump(0 == FALSE); // bool(true)
var_dump(0 == 'guitar'); // bool(true)
var_dump(FALSE == 'test'); // bool(false)
```

This is appropriate when *O* and *FALSE* are equal as they all represent “falsy” value. However, when comparing *O* with “*guitar*”, it is almost impossible to make any conclusion that the output returns *TRUE*. In fact, because PHP will manage to convert “*guitar*” to an integer, so it turns into *O* for the comparison.

In the last line of the example, the result for *FALSE* and “*test*” is *FALSE*. It just conflicts the previous example. If “*test*” become *O*, and *O == FALSE*, why it does not become *TRUE*?

Therefore, in majority cases, loose comparisons are not a problem. But it does cause really annoying bugs and provides weaknesses in security.

b) *Strict comparison*

	<i>TRUE</i>	<i>FALSE</i>	<i>1</i>	<i>θ</i>	<i>-1</i>	<i>"1"</i>	<i>"θ"</i>	<i>"-1"</i>	<i>NULL</i>	<i>array()</i>	<i>"php"</i>	<i>" "</i>
<i>TRUE</i>	<i>TRUE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>								
<i>FALSE</i>	<i>FALSE</i>	<i>TRUE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>							
<i>1</i>	<i>FALSE</i>	<i>FALSE</i>	<i>TRUE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>						
<i>θ</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>TRUE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>
<i>-1</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>TRUE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>
<i>"1"</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>TRUE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>
<i>"θ"</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>TRUE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>
<i>"-1"</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>TRUE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>
<i>NULL</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>TRUE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>
<i>array()</i>	<i>FALSE</i>	<i>FALSE</i>	<i>7</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>TRUE</i>	<i>FALSE</i>	<i>FALSE</i>
<i>"php"</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>TRUE</i>	<i>FALSE</i>	<i>FALSE</i>
<i>" "</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>FALSE</i>	<i>TRUE</i>	<i>TRUE</i>
	<pre> if (\$num_int === \$string_one) { // This will return False echo("Cannot print"); } if (\$num_int === \$string_two) { // This condition statement will return False echo("Cannot print"); } </pre>											

With this strict comparison, programmers can avoid many unexpected bugs, which makes the program become more robust.

B. Type juggling

The reason behind loose comparison in PHP is because of “type juggling”, which is sometime addressed as “type coercion”. An explanation for this term is that when executing an operation, PHP will have to convert them to the most appropriate type so that they are computable. For instance:

```

$test_int = 1;
$test_str = "1";
if ($test_int == $test_str) {
    echo("PHP can compare int and string");
} // no error
// Output: PHP can compare int and string

```

The statements will run with no problems and output “PHP can compare int and string.” This might seem helpful at first because programmers are flexible in handling different types of user inputs.

However, in real world situation, this “type juggling” can also contain considerable amounts of bugs and security vulnerabilities also. As example:

```
echo '40th' + 7; // Outputs 47
```

This unexpected result means that if a string starts with a number, PHP will try to extract that number and use it as a value for the mathematical operations. This problematic behavior also happens when programmers try to do an operation with a string which does not contain any integer:

```
echo 'fortieth' + 7; // Outputs 7
```

This should not be computed, because they are different in both data type and value. However, this PHP’s type coercion still tries to convert the string to 0, then execute the operations. Although this does not change the types of the variable themselves, the only change is in how the operands are evaluated and what type of the expression itself is**Invalid source specified**.

iii. VULNERABILITIES ARISE

C. Magic trick for the exploitation

Since == is the default comparison operator in other languages, bugs emerge in apprentice developer’s PHP program is reasonably ordinary. Notwithstanding, even though type juggling is not facile to be truly exploit because the input parse in HTML is always string after transferring to the program, still attacker can take advantage to dig out the application respectively

```
if ($_POST["password"] == "@dm1np@55w0rd") {
    login_with_role($this->ROLE_ADMIN);
}
```

The code clearly demonstrates the fact that if password was submitted as 0, application would have given admin permission undoubtedly. This is a dream for any cybercrime!

As been stated, HTML input parses the property to string, thus, \$_POST["password"] would be “0” instead of 0 and PHP performs no type conversion. Yet how about the application manages input by using famous *json_decode()* or *unserialize()*, which are trendy in modern RESTful API backed webservice:

```

// JSON blob looks like this:
//{
//  "username": "admin",
//  "password": 0
//}
$credential = json_decode($ POST['credential']);

// vulnerable
if ($credential["username"] == "admin" &&
    $credential["password"] == "@dm1np@55w0rd") {
    login with role($this->ROLE ADMIN);
} // the Laravel encryption payload
{
    "iv": "137f87545d8d2f994c65a6f336507747",
    "value":
    "c30fbe54e025b2a509db7a1fc174783c35d023199f9a0e24ae23a934277aec66",
interp    "mac":
manip "68f6611d14aa021a80c3fc09c638de6de12910486c0c82703315b5d83b8229bb"
place,
    // the neglected MAC check code
    $payload = json_decode(base64_decode($payload), true),
study
specifif ($payload['mac'] != hash_hmac('sha256', $payload['value'], $this-
>key))
    throw new DecryptException("MAC for payload is invalid.");

```

Obviously, as mentioned in **Section II - Division 2.3**, PHP compulsorily transform a string into a number while in loose comparison, hence:

"68f6...29bb" => int(68)

The MAC itself, in substance, is a result of hexadecimal string calculated by `hash_hmac()`, along these lines, if the encryption payload was provided with an integer, type juggling arises:

```
{
    "iv": "137f87545d8d2f994c65a6f336507747",
    "value":
    "c30fbe54e025b2a509db7a1fc174783c35d023199f9a0e24ae23a934277aec66",
    "mac": 68
}
```

This vulnerability result in “padding oracle”, ciphertext decryption, arbitrary text without capturing the encryption key, sensational repercussion of cryptography world. According to MWR InfoSecurity (2014), the indicated exploitation of the crypto flaw leads to impersonation via “Remember Me” cookie and RCE (the most dangerous) with possibly other bugs.

D. Example of vulnerabilities

The following challenge comes from a learning platform, hosted externally on the speculation that the source would be broken in the future, which can be found at <http://se2spring2020.ezyro.com/>. Assumes PHP knowledge acquired, the goal is to find the secret hidden behind *secret.php*, source code attached:

```
function gen_secured_random() { // cause random is the way
    $a = rand(1337,2600)*42;
    $b = rand(1879,1955)*42;

    $a < $b ? $a ^= $b ^= $a ^= $b : $a = $b;

    return $a+$b;
}

function secured_hash_function($plain) { // cause md5 is the best hash ever
    $secured_plain = sanitize_user_input($plain);
    return md5($secured_plain);
}

function sanitize_user_input($input) { // cause someone told me to never
trust user input
    $re = '/[^a-zA-Z0-9]/';
    $secured_input = preg_replace($re, "", $input);
    return $secured_input;
}

require_once "./secret.php";

if (isset($_POST['s']) && isset($_POST['h'])) {
    $seed = sanitize_user_input($_POST['s']);
    $hash = secured_hash_function($_POST['h']);
    $r = gen_secured_random();
    if ($seed != false && $hash != false) {
        if ($seed.$r == $hash) {
            echo "Well done! The flag is ${flag}";
        }
        else {
            echo "Fail...";
        }
    }
    else {
        echo "Hmmm...";
    }
} else {
    echo "input please";
}
```

This section might walk the challenge through the solution!

It is observable that the one and only way to get the flag is to make *\$seed* concatenates *\$r* looks like *\$hash*. However, *\$r* is absolutely randomized, unpredictable; thus, forcing these two strings to be total equal is impassable. Paying attention to *\$hash*, the input hashed with MD5 algorithm. As stated, PHP type juggling convert any string resembles number into number:

```
echo ('0e462...854' == '0efwa...1j2d') // true
```

The line of code is actually comparing *0* with *0*. Indeed, there is existence of a hash titled “magic hash” **Invalid source specified.**:

```
echo hash('md5', '240610708', false) // 0e462097431906509019562988736854
```

Similarly, *\$seed* concatenates *\$r*, as well as *\$hash*, can be converted to 0. Since *\$r* is a random number, the *\$seed* could take “0e” as input to produce a number-like string: “0e153...236”. For elucidation:

All credits are for root-me.org

Unbreakable Random

E. Type Juggling Evasion

Study of the research shows that developers have several approaches to stay away from this b
`$a = 7;
$b = "7_string";
===
only if echo($a === $b); // false - good`
`echo($a === (int) $b); // true - bad`

It is preservable that casting type immediately before comparing would cause an enigma
`strcmp(array(0 => ""), "0a37...c11"); // will return 0 - confused
strcmp((string) array(0 => ""), "0a37...c11"); // type-safed cast, error
should be thrown`

IV. KẾT LUẬN

According to the research, PHP loose comparison and type juggling cause issues for beginner in PHP developing or even intermediate programmers, therefore, the results do support the hypothesis. Moreover, this paper provided a way for developers to avoid aforementioned mistake and write code more precisely.

PHP is such a fantastic programming language that is powerful (blogging system WordPress), deep (social network Facebook) and simple for beginner. Despite having the flexibility, PHP sacrifices security for ease of use, numbers of features provoke grievous exposure. Such bugs, additionally, can create superfluous pitfalls into which even professional not on guard (MAC verification by Laravel in 2013 was broken again after 2 months **Invalid**

source specified.). PHP developers should always be cautious at comparing values in order to avoid type juggling.

REFERENCES

- [1] "PHP: TYPE JUGGLING - MANUAL", PHP.NET, 2020. [ONLINE]. AVAILABLE: [HTTPS://WWW.PHP.NET/MANUAL/EN/LANGUAGE.TYPES.TYPE-JUGGLING.PHP](https://www.php.net/manual/en/language.types.type-juggling.php).
- [2] "PHP TYPE JUGGLING - PHPPOT", PHPPOT, 2020. [ONLINE]. AVAILABLE: [HTTPS://PHPPOT.COM/PHP/PHP-TYPE-JUGGLING/](https://phppot.com/php/php-type-juggling/).
- [3] M. BHATNAGAR, "MAGIC LIES HERE - STATICALLY VS DYNAMICALLY TYPED LANGUAGES", MEDIUM, 2020. [ONLINE]. AVAILABLE: [HTTPS://ANDROID.JLELSE.EU/MAGIC-LIES-HERE-STATICALLY-TYPED-VS-DYNAMICALLY-TYPED-LANGUAGES-D151C7F95E2B](https://android.jlelse.eu/magic-lies-here-statically-typed-vs-dynamically-typed-languages-d151c7f95e2b). [ACCESSED: 24- APR- 2020].
- [4] "LARAVEL COOKIE FORGERY, DECRYPTION, AND RCE", F-SECURE LABS, 2020. [ONLINE]. AVAILABLE: [HTTPS://LABS.F-SECURE.COM/ARCHIVE/LARAVEL-COOKIE-FORGERY-DECRIPTION-AND-RCE/](https://labs.f-secure.com/archive/laravel-cookie-forgery-decryption-and-rce/). [ACCESSED: 30- APR- 2020].
- [5] R. HANSEN, "MAGIC HASHES", WHITEHATSEC.COM, 2020. [ONLINE]. AVAILABLE: [HTTPS://WWW.WHITEHATSEC.COM/BLOG/MAGIC-HASHES/](https://www.whitehatsec.com/blog/magic-hashes/). [ACCESSED: 30- APR- 2020].

APPLYING THREE-TIER ARCHITECTURE TO DEVELOP JAVA WEB APPLICATION FOR TRACKING CORONAVIRUS

SVTH: Do Hai Binh, Truong Ngoc Quynh

Nguyen Cong Quang, Quan Trong Tu

GVHD: ThS Nguyen Dinh Tran Long

Abstract:

Along with the rapid development of technology, the world has been demanding an increasing number of software products. Those products do not only need to be working, but also have to be fast and scalable. In this paper, the development team will discuss Three-Tier Architecture – an architecture facilitating scalability, along with the reasons to use the architecture and demonstrating the result of implementing the architecture in Coronavirus Disease Tracker web application.

Keywords: Multi-tier architecture, three-tier architecture, Java Servlet, Model-View-Controller, Single-page web application

A. INTRODUCTION

Several years ago, web applications were still relatively simple: they were mostly static pages written in Hypertext Markup Language (HTML) with a limited amount of dynamic content, navigation was lengthy, hard to scale, etc. Today, the world demands not only working software products but also systems that are scalable in order to appeal to an increasing number of users around the world. This is where good architecture comes in; there are a good number of architectures that can be used to maximize “separation of concerns” – the fact that components of a system can work independently without too much dependency on others. A popular way of improving a system’s scalability is to employ Three-Tier Architecture. In several organizations, the use of Three-Tier Architecture may also come along with distributing the system across multiple servers. The main content of this paper will explain Three-Tier Architecture along with the demonstration of the architecture in the Coronavirus disease tracker project. The topic was chosen at the greatest outbreak of the disease - when a good and easy-to-use web application to update the disease statistics was greatly in need.

B. MAIN CONTENT

I. Definition of Three-Tier Architecture.

Traditionally, a web application has been built upon Client-Server architecture, in which a client makes requests to the server, and in the server, everything else is done, from getting data to creating HTML content to serve to the client. That means, all necessary codes for managing data and serving markup content must be located within one server only, which makes the server itself hard to maintain, and should there be bugs in the system, it would take a long time to find and resolve the bug.

Then came Two-Tier Architecture, in which the data access node becomes a separate

server. In this approach, there is also the Application server which gets the data from the database and creates HTML content. While this approach is still relatively popular, it is by no means the optimal approach of building modern web applications, because the Application server still has to do multiple jobs. If a web application demands a new look, it can be hard to change only the view layer of the application without affecting other components of the system. It may also require much effort, and money, to add new features into such existing systems, which may not be appealing to corporations.

To facilitate the building of a highly scalable enterprise web application, Three-Tier Architecture was born. In this architecture, there are three separate tiers: Presentation tier, Application tier, and Data tier. It can be seen that the Presentation server has been separated from the Application server. Therefore, it is easier to manage source code for the HTML generation node, and also codes for the Application tier is cleaner since it does not have to take care of HTML generation. The number of tiers is not limited to 2 or 3; there are n-tier architectures available. However, it is not much more beneficial than using Three-Tier architecture.

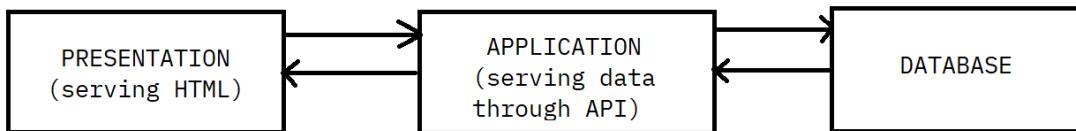


Figure 1: How the tiers of Three-Tier Architecture

The way Three-Tier Architecture works is not much different from that of Client-server and Two-Tier architectures. The client sends the request to the server, then the Presentation tier generates HTML markup as the User Interface to the user. Afterward, whenever a client makes any interaction with the User Interface that needs data, the Presentation tier sends a request to the Application server to fetch the data, then the Application server takes the data from the database and prepares it in the most usable form that the Presentation tier demands. Finally, the Presentation tier generates appropriate HTML content to be displayed on the client's browser.

II. Reasons for using Three-Tier Architecture.

First of all, separating the code of the Presentation tier to a separate server means that this server can be changed without affecting other tiers, provided that the Application Programming Interface (API) of the Application-server does not change. That means a simple Node.js template engine view server can be swapped to a full-fledge ReactJS server if the application needs to be changed into a reactive web application. Similarly, the Application server can change its implementation without affecting other tiers by keeping its API intact. Moreover, the Application-server can be reused as a web service by publicly exposing some of its functionalities.

Secondly, the web application can easily be scaled, which means, new features can be added at ease. For instance, it is possible to add social networking functionalities to an existing student management system in order to facilitate contact between students studying in the same courses, without affecting the operation of the system as a whole.

Thirdly, any of the three tiers can span multiple servers, thus increasing the availability of the system. An example is that, if there are 1 server in Presentation tier, 2 servers act as Application-server, then when 1 out of 2 Application servers is unavailable due to maintenance, the system still works as normal. Moreover, spreading the tiers across multiple servers is a good way of balancing overall system load – the system shall still be responsive even if there are a large number of users accessing it at a time.

III. Three-Tier Architecture in action: Coronavirus disease tracker project summary

Coronavirus Disease Tracker web application is a web application built during the largest outbreak of CoViD-19 disease around the world. This application provides the latest disease data from all over the world along with Vietnamese data of the disease in easy-to-understand visualized formats. Along with the ability to display visualized data, this application also allows the administrator users to update the disease data manually at ease.

Our Coronavirus Disease Tracker web application was built based on Three-Tier Architecture, with the Presentation tier being Node.JS server with EJS templating engine, the Application tier built upon Java Servlet technology and the Database being MySQL. All those three servers are configured to run locally: Presentation server running at port 4200, Application server running on port 8080, and MySQL running on port 3306.

To access the web application, we access <http://localhost:4200/>:



Figure 2: Coronavirus Disease Tracker Homepage

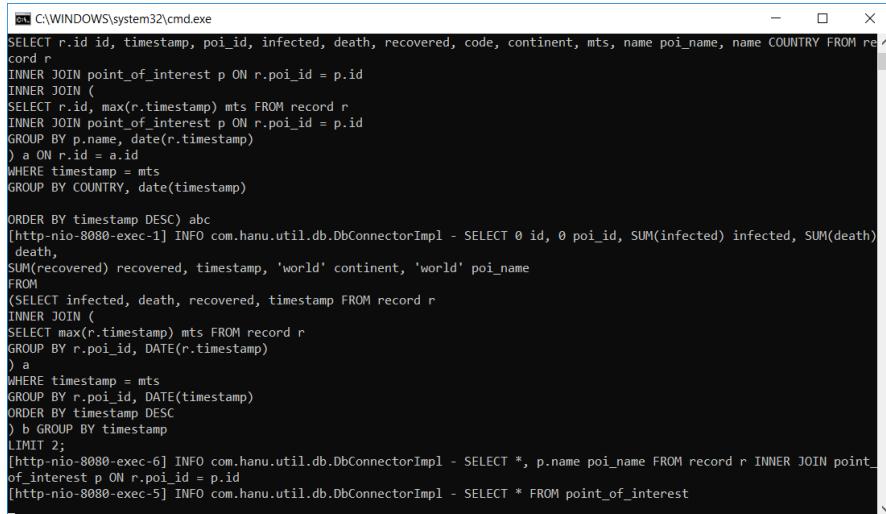
While the page is being loaded, the presentation server *makes requests to the application server* by *fetching* localhost:8080 using the *node-fetch* library to get data from MySQL database:

```
router.get('/', async (req, res) => {
  let page = req.cookies.page ? req.cookies.page : "overview";
  if (req.cookies.page === 'undefined') page = "overview"
  const authenticated = req.cookies.username !== undefined;

  if (page === 'dashboard' && authenticated) {
    if (!cache.get('_admin_stats')) {
      const response = await fetch(`http://localhost:8080/stats`);
      const statsJSON = await response.json();
      cache.put('_admin_stats', statsJSON, defaultTTL);
    }
  }
}
```

Figure 3: Code fragment calling from presentation tier to the Application layer.

Here, *hostname*= “*http://localhost:8080*”



The screenshot shows a Windows Command Prompt window titled 'C:\WINDOWS\system32\cmd.exe'. It contains the following text:

```
SELECT r.id id, timestamp, poi_id, infected, death, recovered, code, continent, mts, name poi_name, name COUNTRY FROM record r
INNER JOIN point_of_interest p ON r.poi_id = p.id
INNER JOIN (
  SELECT r.id, max(r.timestamp) mts FROM record r
  INNER JOIN point_of_interest p ON r.poi_id = p.id
  GROUP BY p.name, date(r.timestamp)
) a ON r.id = a.id
WHERE timestamp = mts
GROUP BY COUNTRY, date(timestamp)

ORDER BY timestamp DESC abc
[http-nio-8080-exec-1] INFO com.hanu.util.db.DbConnectorImpl - SELECT 0 id, 0 poi_id, SUM(infected) infected, SUM(death)
death,
SUM(recovered) recovered, timestamp, 'world' continent, 'world' poi_name
FROM
(SELECT infected, death, recovered, timestamp FROM record r
INNER JOIN (
  SELECT max(r.timestamp) mts FROM record r
  GROUP BY r.poi_id, DATE(r.timestamp)
) a
WHERE timestamp = mts
GROUP BY r.poi_id, DATE(timestamp)
ORDER BY timestamp DESC
) b GROUP BY timestamp
LIMIT 2;
[http-nio-8080-exec-6] INFO com.hanu.util.db.DbConnectorImpl - SELECT *, p.name poi_name FROM record r INNER JOIN point_of_interest p ON r.poi_id = p.id
[http-nio-8080-exec-5] INFO com.hanu.util.db.DbConnectorImpl - SELECT * FROM point_of_interest
```

Figure 4: Coronavirus Disease Tracker’s Application server (Tomcat Embedded 9.0.33)

When using this application, the administrator user may create or update the data of the disease. Under those circumstances, each time the administrator performs an operation on the data, the presentation server will make a call to the application server along with the JSON body so that the application server will process the request into database operation. Then, the application server returns a status code indicating success or failure, and the presentation layer will reload its cache to reflect changes in the database.

IV. Development of Coronavirus disease tracker application

Traditionally, Java web applications have been largely developed using Java Server Pages with Servlet. While this approach has been popular for a while, it has a severe limitation: to deploy the application on a server, Tomcat must be properly set up on the server, which may be time-consuming if the deployment server has not had Tomcat in advance. Moreover, as both the front-end code and back-end code are in the same place, it is difficult for the front-end and back-end team to work separately in parallel. In our project, we

have eliminated those disadvantages by using Tomcat Embedded as a back-end container, a separate server for front-end, and a consistent API for communication between front-end and back-end servers. The development process was divided into phases: Analysis, API design, back-end design, front-end design, back-end implementation, front-end implementation, and manual testing. Due to the length of this paper, only the most important and outstanding diagrams and processes will be attached and discussed; more information can be found in the project's repository [Appendix 1].

Analysis: From the requirements given by the instructor, the project team has listed the most important use cases for this system:

- **Normal user:** View the latest overall disease statistics, View a customized list of visualized statistics.
- **Administrator user:** Login/Logout, Manage disease data including Create, View, Remove, Update records of the disease.

Based on the listed use cases, the team had an agreement on the entities in this system. The entities include Admin representing administrator user, Record representing a disease record at an arbitrary point of time for one location, and Point of Interest representing a location that is having Coronavirus infected cases. Their relationships were modeled in the following Entity-Relationship Diagram:

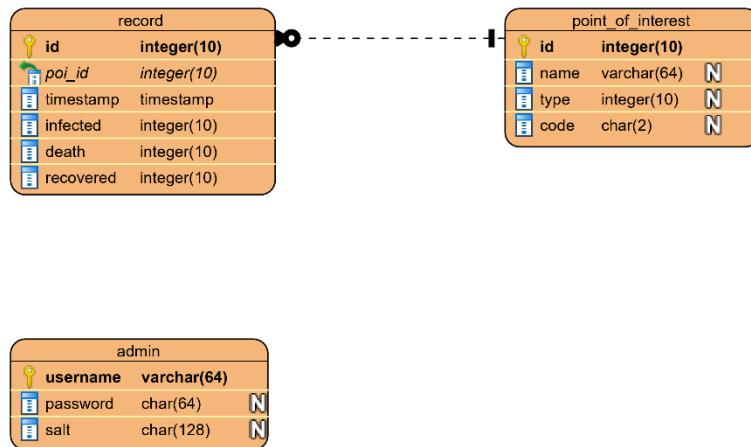


Figure 5: Coronavirus Tracker's Entity-Relationship Diagram. In this diagram, admin has no relationship with other entities, whereas one point_of_interest has zero or more disease records

API Design: The Application Programming Interface of the back-end was modeled based on REST API design, which is resource-based. The API design process was based on what the author Sanjay Patni mention in chapter 2 of his book, Pro RESTful APIs^[1]. First, we defined four resources for the system: stats which represent disease records, poi which represents a location affected by Coronavirus, articles representing news articles crawled from trusted sources, and session representing an administrator session. Each resource will be

represented as an endpoint in this system. Then, we defined appropriate HTTP methods (GET, POST, PUT, PATCH, DELETE) for the resources. Here are several examples of HTTP method usage from /stats endpoint:

- GET /stats: Return all the disease statistics from the outset of the disease.
- POST /stats: Create a new record and add it to the database.
- PUT /stats: Update individual records or mass update the database with the latest records.
- DELETE /stats: Remove a specified record from the database given its ID.

Finally, the team defined the formats which are accepted at each endpoint. The complete API design is available on the project's repository link [Appendix. 1] in YAML format.

Backend Design:

The backend was designed based on the Model-View-Controller (MVC) pattern with a few modifications. Firstly, the View in MVC is removed due to the Presentation layer being completely separated from the Application layer. Secondly, use-cases are separated from Controllers so that the flow of the system becomes easy to understand. Overall, the design is still MVC, and database operations are abstracted into Repository classes.

Frontend Design:

The frontend server only uses Node.js with the EJS template engine to serve dynamic HTML content, so this server doesn't need to have a fully-fledged Object-Oriented design. The views are divided into components in the same way as one would do in a frontend framework like React or Angular to improve modularity. More information on the frontend's code can be found in the project repository [Appendix.1]

Implementation:

Implementation is carried out at the same time for front-end and backend. The front-end server was developed using a mock API service provided by the Anypoint platform of Mulesoft. The backend server was developed in the same way as to how a normal Servlet application is created, the only difference is that our project has Tomcat container embedded into the system so that the whole application can run on the console like a normal Java application, in the same way as to how Spring Boot and ASP.NET Core web applications work.

V. Coronavirus disease tracker – work results

After a long time of working with this project, the project team has released the first major version of Coronavirus disease tracker app. This release contains a homepage for all users of the system as shown in the previous part, along with the functionality to log in as an administrator to make changes to the data. As the administrator, the user can create records of

the disease for a location manually, update records manually, or have the system get real-time data from trustworthy sources such as Worldometers or Vietnamese Ministry of Health.

The login form consists of two input fields: 'binhdh' in the username field and '*****' in the password field. Below the fields are 'Close' and 'Log in' buttons.

Figure 6: Login form for the system

The management view shows a table of COVID-19 data for various locations. The columns are: #, Time, Location, Continent, Infected, Death, Recovered, Edit button, and Delete button. The table includes rows for Papua New Guinea, Cambodia, Kazakhstan, Paraguay, Syria, Bahamas, Montserrat, Mali, and Panama.

#	Time	Location	Continent	Infected	Death	Recovered	Edit	Delete
276291	Fri, 01 May 2020 17:26:15	Papua New Guinea	Oceania	8	0	0		
276292	Fri, 01 May 2020 17:26:15	Cambodia	Asia	122	0	119		
276293	Fri, 01 May 2020 17:26:15	Kazakhstan	Europe	3551	25	866		
276294	Fri, 01 May 2020 17:26:15	Paraguay	South America	266	10	113		
276295	Fri, 01 May 2020 17:26:15	Syria	Asia	43	3	21		
276296	Fri, 01 May 2020 17:26:15	Bahamas	North America	81	11	25		
276297	Fri, 01 May 2020 17:26:15	Montserrat	North America	11	1	2		
276298	Fri, 01 May 2020 17:26:15	Mali	Africa	508	26	196		
276299	Fri, 01 May 2020 17:26:15	Panama	North America	6532	188	576		

Figure 7: Management view.

In this view the administrator can: Add a new record, Update or delete existing records, manage data for Vietnam or countries, manage data by date, and Automatically get latest data.

The search results show one record for Vietnam. The columns are: #, Time, Location, Continent, Infected, Death, Recovered, Edit button, and Delete button. The table includes rows for Vietnam.

#	Time	Location	Continent	Infected	Death	Recovered	Edit	Delete
276440	Fri, 01 May 2020 17:26:15	Vietnam	Asia	270	0	219		

Figure 8: Find a record for the current view date by location name

Add a new record

Location:	Vietnam	+
Timestamp:	Vietnam	
Infected:	271	
Death:	0	
Recovered:	224	

Close **+ Add**

Figure 9: Adding a new record for a location. This example will create a record for Vietnam, date is May 2nd, 2020, with the statistics as shown on the photo.

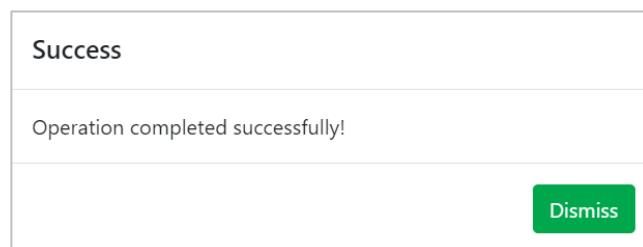


Figure 10: Success message

Manage CoViD-19 data						
Only stats for the latest date can be modified!						
#	Time	Location	Continent	Infected	Death	Recovered
276535	Sat, 02 May 2020 16:00:00	Vietnam	Asia	271	0	224

Figure 11: Record successfully inserted

Update a record

Location:	Vietnam
Timestamp:	2020-05-02 16:00:00 UTC
Infected:	271
Death:	0
Recovered:	224

Close **✓Confirm**

Figure 12: Updating a record. Note that Location and Timestamp are not allowed to change, so they were greyed out.

#	Time	Location	Continent	Infected	Death	Recovered	Edit
257661	Thu, 02 Apr 2020 16:00:00	Benin	Africa	13	0	1	<input checked="" type="checkbox"/> Edit
257761	Thu, 02 Apr 2020 16:00:00	Papua New Guinea	Oceania	1	0	0	<input checked="" type="checkbox"/> Edit
257861	Thu, 02 Apr 2020 16:00:00	Angola	Africa	8	2	1	<input checked="" type="checkbox"/> Edit
257961	Thu, 02 Apr 2020 16:00:00	Cambodia	Asia	110	0	34	<input checked="" type="checkbox"/> Edit
258061	Thu, 02 Apr 2020 16:00:00	Sudan	Africa	8	2	2	<input checked="" type="checkbox"/> Edit
258161	Thu, 02 Apr 2020 16:00:00	Kazakhstan	Europe	435	3	27	<input checked="" type="checkbox"/> Edit
258261	Thu, 02 Apr 2020 16:00:00	Paraguay	South America	77	3	2	<input checked="" type="checkbox"/> Edit
258361	Thu, 02 Apr 2020 16:00:00	Portugal	Europe	9034	209	68	<input checked="" type="checkbox"/> Edit
258461	Thu, 02 Apr 2020 16:00:00	Syria	Asia	16	2	0	<input checked="" type="checkbox"/> Edit

Figure 13: Manage data by date. Note that older records than the latest date cannot be deleted.

#	Location	Continent	Infected	Death	Recovered
1	US	North America	1111236	64876	157809
2	Spain	Europe	242988	24824	142450
3	Italy	Europe	207428	28236	78249
4	United Kingdom	Europe	177454	27510	0
5	France	Europe	167178	24376	49476
6	Germany	Europe	163542	6640	126900

Figure 14: Customized view of the disease data. All users can get access to this view.

In this view, user can include any of the pre-defined type of visualizations.

Custom visualizations are not yet supported.

VI. Evaluation

○ Advantages:

- The system is clearly separated into three tiers, each taking up one server and one port.
- The Application tier server, or the backend server, was designed with a consistent Application Programming Interface. The API specification in the YAML format is available on the project's GitHub repository [Appendix 1].

- We have implemented all the core functionalities: View world statistics, view Vietnamese statistics, administrator login, manage world statistics, manage Vietnamese statistics in a near single-page manner – which means the page still navigates while the URL address remains unchanged. We have also used Google Charts library to support the visualizations in a Customized View page. Moreover, getting a real-time update for the database is supported in the project.

○ **Limitations:**

- It is time-consuming to build the application, compared to building a plain JSP/Servlet application. This should only adversely affect inexperienced teams but will not hold back teams of experts.

- In our project, caching is not yet the main concern. We have only implemented a simple form of caching on the frontend server which monitors for data-modification requests and purges the cache on those requests.

○ **Future work:**

- A more sophisticated means of caching will be applied to reduce response time, especially for heavy processes such as loading multiple different visualizations that require different resource types.

- More demanding features will be added to the system, such as exporting report data to PDF or Excel format, or schedule automatic data update so that the data remains the latest even if administrators forget their job.

- Changing the Presentation tier technology to a reactive frontend framework such as ReactJS or Angular is in consideration, especially given that Coronavirus Disease Tracker application has been built to be a single-page web application.

- The development team also considers making the backend of this project available as a public Web API so that other developers can easily embed the latest disease information into their applications.

C. CONCLUSION

Software engineering today is no longer centralized on working software products but concentrates more on other metrics such as availability, ease of use, and most importantly scalability since large-scale systems are much in demand these days. In this paper, the authors have discussed Three-Tier Architecture, the upside and downside of this architectural choice, along with the application of the architecture to Coronavirus Disease Tracker web application. Three-Tier Architecture is a powerful architecture for building secured, scalable web applications with a guarantee for near-perfect availability; therefore, it has been popular in recent years among enterprises. Development of applications based on this architecture is beneficial as the back-end team and the front-end team can work in parallel at the same time, provided that the front-end has access to a mock API. Moreover, in recent years, there has

been a trend of building self-contained web applications following the advent of libraries and frameworks such as ASP.NET Core and Spring Boot. Tomcat Embedded was employed in our application in order to eliminate for an external Tomcat server container, making the application self-contained, thus can be executed by simply running the built JAR file. The resulting website – Coronavirus Disease Tracker web application – may not be the fastest web application to develop, and performance may be not yet good enough for a large number of users at the same time, but it demonstrated where the architecture is good at, which is complete separation between front-end service and back-end service. In the future, more work will be done on this project to ensure high-performance, and more useful functionalities will be added to increase the usability of the application even when the disease is completely under control.

REFERENCES

- [1] dzone.com. 2017. *What Is N-Tier Architecture?* - Dzone Devops. [online] Available at: <<https://dzone.com/articles/what-is-n-tier-architecture>> [Accessed 6 May 2020].
- [2] Patni, S., 2017. *Pro Restful APIs.*, Chapter 2: API Design and Modelling.

APPENDIX

- [3] Project repository: binhdoitsme/SE2_Spring2020_Project:
https://github.com/binhdoitsme/SE2_Spring2020_Project

RESEARCH ABOUT BLOCKCHAIN TECHNOLOGY, TYPE OF ATTACKS AND SOLUTIONS

SVTH: Lương Quỳnh Hương, Cao Thu Huyền, Lê Đức Anh

GVHD: ThS Đỗ Thị Phương Thảo

Abstract - In the 4.0 technology wave, Blockchain is considered by the technology world to be the most noteworthy advancement after the birth of Internet. With the ability to share data information transparently, saving storage space and high security. So, Blockchain is a wide application in many fields and trades. Especially, Blockchain's most well-known applications are celebrated cryptocurrencies (Bitcoin, Ethereum, Swell, ...) which are raging the advertise.

However, nothing is perfect, besides the advantages, blockchain also gets some problems. One of them is an attack vulnerability called Double spending. Double-spending is an issue where the same cryptocurrency can be utilized different times. Ensuring security for this issue is an extremely urgent task. So anticipating Double-spending attacks for Blockchain framework is greatly essential within the context of rapidly developing cryptocurrencies.

Therefore, the findings of this paper will find out about Blockchain and Bitcoin Cryptocurrency. Along with that, we will mention to a weak point of Blockchain- Double spending and research on security against attacks including Double spending hypothesis attacks and attack resistance in practice help Blockchain apply for Bitcoin cryptocurrency.

Key – Blockchain, attacks, double spending

A: OVERVIEW ABOUT BLOCKCHAIN AND CRYPTOCURRENCY

1: Overview of Blockchain

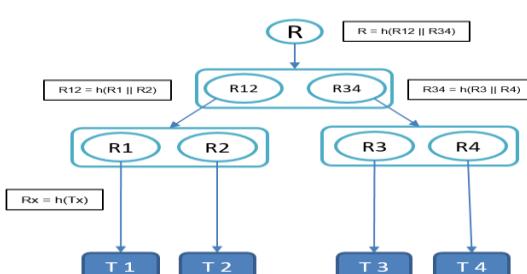
1.1: Definition

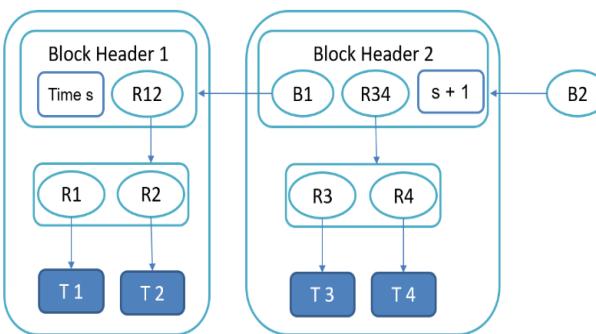
As defined, Blockchains is a decentralized database that stores information in information blocks linked together by coding and extending by time. Each information block contains information about the initialization time as well as linked to the previous block, along with a time code and transaction data.

“The blockchain is an incorruptible digital ledger of economic transactions that can be programmed to record not just financial transactions but virtually everything of value.”. According to Don and alex Tapscott, authors of Blockchain Revolution (2016).

1.2: Blockchain structure

Merkle tree: In encrypt data, Hash tree or Merkle tree is a model in which the treetop of these model are Hash's value of a data block, and the rest are the hash value of points lower than it. Root of Merkle is hash value of





all its intermediate values. (*Figure 1: Merkle Model*)

Chain link : A chain, or a list of chain in picture (*Figure 2: Merkle tree link together by Hash code*), created when an unit of data store a Hash value of an data unit before. Only need 1 change in the data before, Hash value will be changed and the chain will be broken.

A group of data include the trunk is Merkle tree that contain link to the exchange happen at the same time and the top contain root of merkle tree with time function. In picture, the other data blocks allow chronological order when created will contain with Hash link of the group data before. All of the structure will create a thing that called blockchain.

1.3: Characteristics of blockchain:

- Indestructible and indestructible
- Immutability
- Security
- Transparency
- Smart Contract

1.4: Applications of Blockchain

Blockchain is now revolutionizing most industries across the globe: Entertainment Technology (KickCity, Guts...); Social Networking dating (Matchpool,...); Retail industry (retail industry, warranteer,...); Driver (Bitcar,...); Health care (MedicalChain,...); Real Estate (BitProperty, Deedcoin,...).

According to the statistics of Infinity Blockchain Lab, although currently in Vietnam market, Blockchain technology, although not too strange, is still not widely available, mainly applied to the following areas:

Financial services (> 83%)

Agriculture / Food / Commodity Supply Chain (nearly 40%)

Public services (approximately 30%)

1.5: Main versions of Blockchain:

Blockchain technology currently has 4 major versions:

- Blockchain 1.0 - Currency and Payment
- Blockchain 2.0 - Finance and Markets
- Blockchain 3.0 - Design and Operation Monitoring

- Blockchain 4.0 - Real time platform

2. Overview of bitcoin cryptocurrency

2.1: Definition

Cryptocurrency, known as encrypted money, is designed to act like a means of transaction. This money uses code to secure and verify exchanges, also to control the way of creating new units of one specific encrypted money.

Bitcoin is the first electronic money to use peer-to-peer technology to have open source, the first electronic money using blockchain. "Bitcoin is really the first application of blockchain technology that allows us to send a digital asset via the internet to another user on the Internet, in addition to this secure transaction. Safe and confidential, everyone knows that this transfer has taken place and no one can intervene and interfere or modify this process" according to Marc Andreessen.

2.2: Popular cryptocurrency, bitcoin wallet, bitcoin mining

Common currency modes: Bitcoin (BTC, XBT); Ethereum (ETH); Ripple (XRP); Litecoin(LTC); Zcash (ZEC)

Bitcoin wallet

Everyone uses Bitcoin to create a self-wallet to send Bitcoin and store it. Most of the wallets store Bitcoin have high security through many classes or functions that allow people to print paper wallets for themselves to store in safe places.

Coin mining

Mining Bitcoin is a "miner" process of finding a solution to a difficult math problem. The difficulty of this problem is automatically corrected so that on average every 10 minutes a miner (or a group) scattered around the world solves a block and gets 12.5 BTC / block and 1 BTC is currently at \$ 2400.

Virtual coin mining in general and Bitcoin mining in particular are the form of using computers with high VGA configuration, or other specialized hardware to run mining software.

1) 2.3. Advantages, disadvantages and applications of Bitcoin cryptocurrency

a. Advantages:

- Convenience in transaction
- Bitcoin cannot be counterfeited
- High security and very safe
- Extremely low transaction costs
- Protect the environment
- The potential of developing e-commerce

b. Disadvantages of Bitcoin

- The number of users is not much
- It's not easy to use bitcoins
- Bitcoin prices often fluctuate
- Bitcoin prices often fluctuate

c. Applications

- Used for exchanging and trading other coins on the cryptocurrency trading floor
- Used to be an international remittance with super low cost
- Forms of payment in BTC help minimize costs, save time for businesses
- Use for short-term / medium-term investment / speculation, to earn additional income.

B: TYPE OF ATTACK TO BLOCKCHAIN SYSTEM

1: Risk when use blockchain technique

-Lack of privacy

Many blockchain doesn't have privacy. Account balance and trade can be seen by everyone.

- Privacy challenges

Many secure processes in Blockchain makes it more difficult to agree on a common problem and may be less secure than current methods because people will record a private key so they don't forget it.

- Inexistence of centralize management

None of single organization can be managed changes or trends in Blockchain, this make organizations that use it may have to deal with risks when trading since they can't control the adjustment in the system.

- Risk of 51% Attack

Many computer operating Blockchain globally are located where is unreliable due to criminal issue, law system and shortages of regulations. Low electricity costs and computer costs in many countries have created many blocks mining center in Blockchain. If data centrals linked together, they have ability to control more than 50% network and take control the management network.

- Unqualified technology

Blockchain technologies are practically unqualified and mostly used in virtual money. These limit companies or realistic system that apply blockchain technology to prove higher productivity than the current system.

- Expenses

Require high amount of electricity to operate. For every half an hour that blockchain consume equal to the electricity that households in America use in one year.

- Expanding challenges

The blockchain network hasn't been prove to have the efficiency compared to the current system. Bitcoin network can only process about 7 trade per second while Visa can

process upto 20.000 trades per sec.

- **Lack of knowledge about Blockchain technology**

Blockchain operating methodology and its benefit is complex to many people. They also concern about blockchain perspective such as their balances and transactions are public. Even when people understand those benefits, they still fond of current system.

- **Policy and combination**

Systems based on blockchain platform will have to deal with policy problems with the high-time consumption and high cost in cooperating with current systems. Government and banks stand against changes because of scale and the expense is too high to renovate current system.

2. Basic types of attacks

- **Attack Sybil**

The Sybil attack aims to damage the P2P network by forming some fake identities. Attackers can set up some fake buttons that seem real to their peers. These spoofed nodes are involved in damaging the network to authenticate unauthorized transactions and to change valid transactions. They may use some device, virtual machine or internet protocol (IP) address as a fake node for the attack. The P2P network assumes that each participating node consists of only one identity. Consequently, some pseudo-nodes provide an attacker the ability to reject transmission blocks and to bypass honest nodes. When an opponent owns a large number of network nodes, it will double the opportunity to spend double. However, having a large number of nodes in the bitcoin blockchain network leads to a very expensive attack.

- **Eclipse attack**

A fruitful Overshadow Assault empowers a would-be awful performing artist to disconnect and hence avoid their target from achieving a genuine picture of genuine organize movement and the current record state.

This assault is made conceivable since a decentralized organize does not let all hubs at the same time interface to all other hubs on the arrange. Instep, for proficiency, a hub interfaces to a select bunch of other hubs, who in turn are associated to a select bunch of their possess. For case, a Bitcoin hub has eight outgoing associations; Ethereum 13.

Other potential eclipse Attack effects on a PoW network such as Bitcoin, as noted in the 2015 paper by Heilman, Kendler, Zohar and Goldberg, include:

Engineering piece races: Two miners find a block simultaneously

Splitting mining power: In an effort to make it easier to launch 51% attacks a malicious actor could eliminate eclipsed miners from the network and in doing so reduce the amount of active miners

N-confirmation double spend: Most merchants wait until a number (N) of blocks have been confirmed before releasing their goods to prevent against double spend attacks. However, if enough miners have been eclipsed then an attacker can launch a double spend attack against an eclipsed merchant in spite of this. An attacker would be able to present their transaction to eclipsed miners, who add it to their ledger. This state is then shown to the merchant who cannot know better, given they too have been eclipsed and removed from the wider network. Once the merchant has sent the reciprocal goods to the attacker then the eclipsed miners' ledger is discarded and the merchant is left empty handed

- **51% attacks**

In blockchain bitcoin technique, Since the process of mining (in PoW-based systems) involves the investment of huge amounts of electricity and computational resources, a miner's performance is based on the amount of computational power he has, and this is usually referred to as hash power or hash rate. There are many mining nodes in various locations and they compete to be the next to find a valid block hash and be rewarded with newly generated Bitcoins.

In such a context, the mining power is distributed over different nodes across the world, which means the hash rate is not in the hands of a single entity.

But what happens when the hash rate is no longer distributed well enough? What happens if, for example, one single entity or organization is able to obtain more than 50% of the hashing power? One possible consequence of that is what we call a 51% attack, also known as majority attack.

Simply, a 51% attack occurs when malicious actors pick up control of more than 50% of a blockchain network's hash rate, hence the name. Since the hackers have at least 51% of the network's hash rate, they can force the rest of the network to erase their transactions.

- **Finney attack**

The Finney assault may be a variety of a double-spend assault. The aggressor makes two exchanges - one crediting the victim and one crediting themselves. They keep the primary exchange for presently and continue to undertake mining the moment one into a block. When they succeed (this may take a while), they quickly make a buy with the primary exchange, get the products they acquired, and after that discharge the pre-mined piece. This way the primary exchange will gotten to be nullified, indeed in case it is engendered through the entire network.

This variation of double-spend assault is harder to execute (mining a piece by oneself hasn't been simple for a long whereas), but is imperceptible until it has been completely executed. The as it were way to secure oneself against such an assault is to require at slightest one affirmation for exchange some time recently giving out obtained products, and requiring more affirmations for exchanges worth more.

- **Timewarp attack**

Timestamp faking was the key vulnerability used by the attackers in this hack. Since Verge is a decentralized network, there's no central authority to set the official time for everyone. As a result, enforcing perfect time synchronization within the network is also quite challenging, and consecutive blocks may have timestamps that are out of order (MihailSotnichek, 2018). It mean that Disturbing block timestamps to reduce network difficulty (recently used against cryptocurrency Verge)

3. Double spending attack

2) 3.1. Definition

- Double spending is a problem in which a kind of virtual money can be used many times. In the other words, Double spending is a case that a transaction uses the same input as another transaction which is broadcasted on the Network. This a only serious hole of digital currencies as digital information can be copied easily. Some virtual currencies such as Bitcoin can be considered as a digital file (Bisade Aslo,2018).

Looking at the example below

A goes to Vinmart and buys something, total orders is \$20. A pays by cash and receives bill. As soon as cashier puts \$20 in the register, A consumed \$20, A can only use that \$20 again only if A steals it.

Physical currencies do not have the same Double spending problem that is faced by digital currencies. Because each participants in the transaction of a physical currency already get right to visual access to the original physical currency.

3) 3.2 Operation principle

The hypothesis of Double spending is that a hacker gets control of more than half of the mining system in the Blockchain.

a. Task

Supposed that Linh owns 100 Bitcoin and use it to buy a laptop. After delivering 100 Bitcoin into laptop company's wallet. Linh will receive the laptop a few days later.

By carrying out a Double spending attack on Bitcoin'sBlockchain, the transaction can be reversed. In case of successful attack, Linh can own the laptop and 100 Bitcoin still remains in the Linh's wallet. From that, Linh goes on using 100 Bitcoin at another transaction.

When Linh carries out transaction, 100 Bitcoin will be transfer into laptop company's wallet. This transaction is encrypted and put into a place called "unconfirmed transactions". The miner selects these unconfirmed transactions and forms a block. Then the block is added to the network blockchain grid.

However, to do that, the miners need to solve a very complicated issue. Numerous

miners will fight with each other, whoever tackles the issue first will be united his square on the blockchain.

b. Mechanism of action

Before going into the working principle of the 51% attack, let's say once more about Blockchain. A blockchain governs a data ledger, for example transaction data. Blockchain replaces 3rd parties to do this, such as government or banking. This is what makes (most) decentralized blockchains. The Bitcoin blockchain protocol is based on democracy, which means that the majority of participants (miners) on the network will decide which version of the blockchain represents the truth.

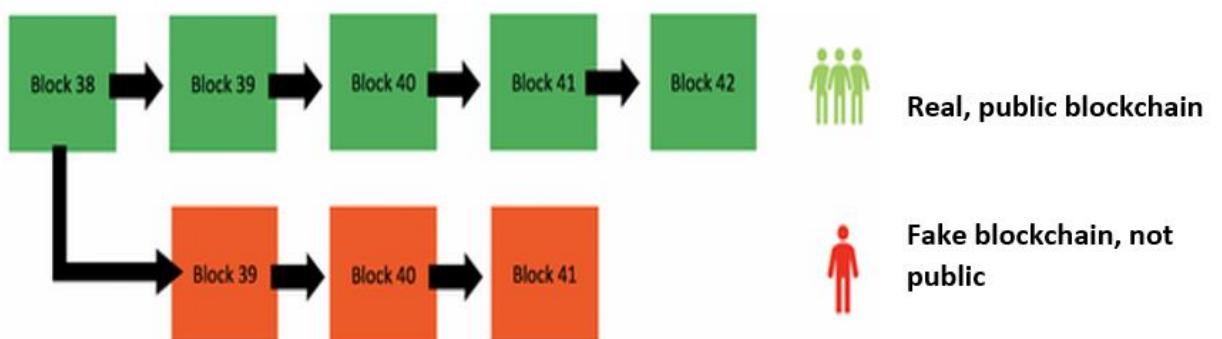


Figure 3: Creating a fake, undeclared Blockchain (The orange blockchain can be considered in 'stealth' mode.)

A miner be able to not public his block and create a branch blockchain which is not declared as Figure 3. In figure 4, when I complete a transaction of buying a laptop with 100 Bitcoin, I declare and confirm that transaction on green blockchain. The transaction will be confirmed by other miners and my wallet will be deducted 100 Bitcoin.

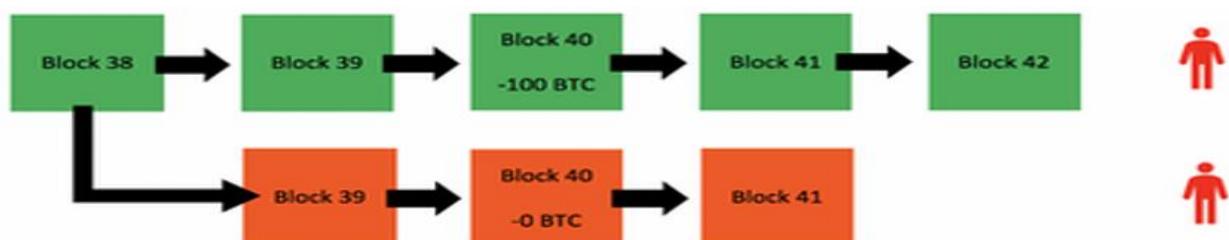


Figure 4: Making transactions both on the fake and real Blockchain

On the orange blockchain, I can not make this transaction and 100 Bitcoin be still in my wallet. This blockchain is not declared, so it is not possible for other miners to know

Blockchain operates under a democratic governance model, also known as a majority. Blockchain also confirms the longest blockchains. Normally, the real blockchain version with the participation of many miners around the world will have the fastest speed of adding new blocks. This is the way to blockchain determine which version is real. If I have enough power to run faster than the real blockchain, I can turn my fake blockchain into real blockchain (Whoever has the most hashing power will add blocks to their version of the chain faster- Figure 5)

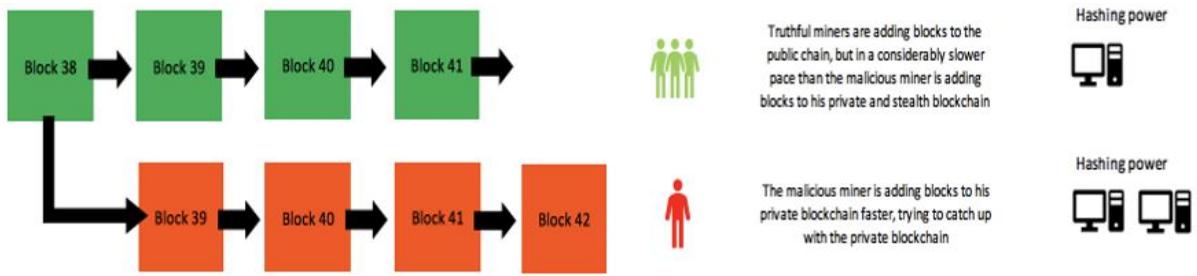


Figure 5: The corrupt miner is now adding blocks to his private chain faster because he has more hashing power.

Therefore, after creating the longer blockchain, I broadcast my orange blockchain to all other miners.

The rest of the network includes other miners, will discover that my fake blockchain is longer than the version which they are working on. These miners have to move through the new blockchain.

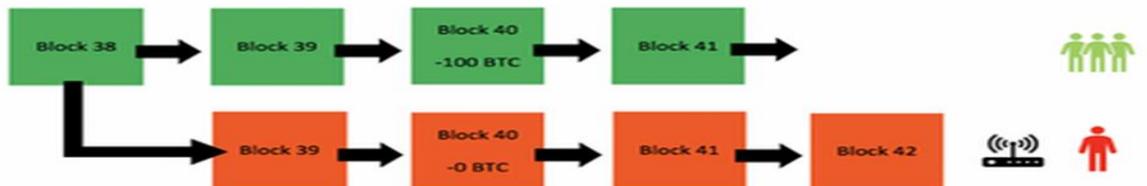


Figure 6: Creating the longer chains and broadcasting them to the miners

I fooled all other miners, turning my fake blockchain into real blockchain. In figure 7, at block 40 of the fake blockchain, I cannot make a Bitcoin transfer. So, I still have a laptop and I still have 100 Bitcoin in my wallet. Obviously, the wallet of company selling laptop will not have 100 Bitcoin of mine.



Figure 7: Occupy control and become the main chain

A flaw in the bitcoin system that can lead Double Spending is the 51% attack, because of getting a faster confirmation speed than the entire blockchain network, individually, I have to gain the computing power by 51% of total power of excavators in the world which joins to this network

C: Double spending resistance solution

1: State and probability of performing Double Spending and 51%

An attacker is trying to create an alternative string faster than honest chain. Even these happened, it also does not open the system for arbitrary changes, such as create value out of stream or take money never belongs to an attacker. Nodes will not accept an invalid transaction as payment and honest nodes will never accept a block contain them. An attacker

just can try to change one of the transactions of them to bring back the money that he pay recently.

After the researchers solved the problem on this issue, the probability of the attacker catching up is 0.0002428 (~ 0.03%). In the real Bitcoin network, it is very difficult for an attacker to spend double the amount they paid earlier. Although experts admit that it is theoretically possible, most people think that 51% attacks will be so difficult to execute that they have not posed a legitimate threat.

However, everything has changed. Last year, a number of successful 51% attacks, including a few on reputable blockchains, proved that the threat was real. 51% attacks are no longer an abstract concern. Every single blockchain project must put security before all else.

51% Attacks success in 2018 & 2019

There were a number of successful 51% attacks in 2018 and the trend is continuing in 2019. Here is a list of the attacks that have occurred

- April 4, 2018: Verge (XVG) 51% attacked for a loss of ~\$1.1 Million.
- May 14, 2018: Monacoin (MONA) 51% attacked for a loss of ~\$90,000.
- May 22, 2018: Verge (XVG) 51% attacked again for a loss of ~\$1.75 Million.
- May 29, 2018: Bitcoin Gold (BTG) 51% attacked for a loss of ~\$18 Million.
- June 2, 2018: ZenCash (ZEN) 51% attacked for a loss of ~\$550,000.
- June 4, 2018: Litecoin Cash (LCC) 51% attacked for unknown losses.
- September 8, 2018: FLO Blockchain (FLO) 51% attacked for a loss of ~\$27,500.
- November 8, 2018: Aurum Coin (AU) 51% attacked for a loss of ~\$500,000.
- December 2, 2018: Vertcoin (VTC) 51% attacked for a loss of ~\$100,000.
- January 7, 2019: Ethereum Classic 51% attacked for a loss of ~\$1.1 Million

It's important to note that these are only the 51% attacks that were revealed publicly. It's quite likely that many more attacks occurred away from the public eye and were never publicized

All of consequences make a 51% attack a catastrophic event for any blockchain project.

2. Solution to minimize attacks from Double spending

In reality, the attacks are troublesome to execute. Like it said recently, a miner will require more hashing control than the rest of the combined network to realize this. Consider the truth that maybe even hundreds thousands of miners on the Bitcoin Blockchain, one malicious miner will need to spend huge sums of money on mining equipment to fight with the rest of network. Even the strongest computers in the world cannot compete directly with the total power of calculation in this network

Avoiding double spending can frequently be taken care of in two ways: collective centralized or decentralized.

- With collective centralized, a central third party and trustworthy will usually be responsible for verifying that a digital currency has not been double spent.

- The decentralized nature of Bitcoin means interrelated issues related to the focus method mentioned above, such as a ventricular point defeat only and must believe that a third party is blocking the correct amounts doublespending, no longer exists. Bitcoin use a consensus mechanism is called proof of work to avoid the need of a focused party. Instead of requesting a trustworthy central party to validate the transactions are not double spending, a group of decentralized individuals called miners perform this task. All of the transactions Bitcoin also included in storage system called Blockchain, this guaranteed that it is able to prove that any party wants to spend Bitcoin really owns that bit. A transaction is invalid when a block is added to the chain and will be stored in blockchain. When many others block is added to Blockchain, the return and double spending will be more difficult

Besides, there are 5 most advanced defense techniques to reduce 51% of attacks that are: A Penalty system for delayed block submission, Delayed proof of Work (dPoW), PirlGuard, Chainlocks, Merged Mining.

Summary - Through time of research, research and implementation, we have introduced everyone to the general content of blockchain, cryptocurrencies and Bitcoin, outlined the basic types of attacks on the blockchain system. overview of Double Spending attacks and technical solutions to minimize Double Spending attacks today. This is a unique flaw for cryptocurrencies because digital information is something that can be copied quite easily. Ensuring safety for this is an extremely urgent task to ensure the transparency of Blockchian and ensure the value of digital currencies. So preventing Double Spending attacks for Blockchain systems is a prerequisite for cryptocurrency to be put into practical and safe applications.

REFERENCES

- [1] Bisade Asolo. (2018, December 21), *Double-Spending Explained*. Online: <https://www.mycryptopedia.com/double-spending-explained/>.
- [2] Jimi S. (2018, May 5), *Blockchain: how a 51% attack works (double spend attack)*. Online: <https://blog.goodaudience.com/what-is-a-51-attack-or-double-spend-attack-aa108db63474>
- [3] , Mark Gates, Book: *Blockchain – Bản chất của Blockchain, Bitcoin, Tiền điện tử, Hợp đồng thông minh và tương lai của tiền tệ*.
- [4] Satoshi Nakamoto. (2008), *Bitcoin: A peer-to-peer electronic cash system*. Online: <https://bitcoin.org/bitcoin.pdf>.
- [5] Daniel. (2018, July 20), *The Anatomy of A 51% Attack And How You Can Prevent*

One. Online: <https://komodoplatform.com/51-attack-how-komodo-can-help-prevent-one/>.

[6] Sarwar Sayeed and Hector Marco-Gisbert. (2019, April 29), *Assessing Blockchain Consensus and Security Mechanisms against the 51% Attack*. Online:

https://res.mdpi.com/appsci/appsci-09-01788/article_deploy/appsci-09-01788.pdf?.

[7] Satinder Grewal (2018, Dec 18), *Komodo's Delayed Proof of Work (dPoW) Security*. Online: <https://blog.komodoplatform.com/delayed-proof-of-work-explained-9a74250dbb86>.

[8] Fawkes. (2018, Nov 21) *PirlGuard—Innovative Solution against 51% Attacks*. Online: <https://medium.com/pirl/pirlguard-innovative-solution-against-51-attacks-87dd45aa1109>.

[9] Hà Chí Hiếu. (2017), *Khái niệm cơ bản về Blockchain*. Online: <https://kipalog.kaopiz.com/posts/Khai-niem-co-ban-ve-Blockchain>

[10] Margaret Rouse. (2019, May), *peer-to-peer (P2P)*. Online: <https://searchnetworking.techtarget.com/definition/peer-to-peer>

[11] Nguyễn Ngọc Khánh. (2018, August 06), *Blockchain: lịch sử, tính chất và ứng dụng*. Online:

<http://antoanthongtin.vn/Detail.aspx?CatID=c251d538-7a3c-4fc7-81df-44a2de35883f&NewsID=f25d2666-d463-4cec-a8d7-23c302b348f0>

[12] SolarWinds MSP. (2018, November 30), *Centralized Networks vs Decentralized Networks*. Online:

<https://www.solarwindsmsp.com/blog/centralized-vs-decentralized-network>

[13] Mi-A. (2018, May 23). *Những ứng dụng thiết thực của Blockchain trong cuộc sống*. Online:

<https://bigcoinvietnam.com/nhung-ung-dung-thiet-thuc-cua-blockchain-trong-cuoc-song>

[14] Thị trường tài chính Việt Nam. (2019, July 18), *Blockchain là gì? Thông tin về Blockchain người chơi tiền điện tử cần nắm được*. Online: <https://thebank.vn/blog/18009-blockchain-la-gi.html>

[15] PcDinh. (2015, April 23), *Tấn công mạo nhận (Sybil attack) trong mạng ngang hàng*. Online:

<https://thegioicoin.wordpress.com/2015/04/23/tan-cong-mao-nhan-sybil-attack-trong-mang-ngang-hang-p2p-network/>

[16] Goupadmin, (2018, June, 07), *What is an Eclipse Attack?*, Online:

<https://www.radixdlt.com/post/what-is-an-eclipse-attack/>

[17] Mihail Sotnichek. (2018, September 06), *Blockchain Vulnerabilities: Verge*

Network Mining Attack, Online: <https://www.apriorit.com/dev-blog/563-verge-mining-hack>

[18] Jimi S. (2018, May 3), *Blockchain: how mining works and transactions are processed in seven steps.* Online: <https://blog.goodaudience.com/how-a-miner-adds-transactions-to-the-blockchain-in-seven-steps-856053271476>

[19] Nohara. (2019, May 28) , *Blockchain 3.0 là gì, Những tính năng vượt trội.* Online: <https://vn.wacontre.com/vi/blockchain-3-0-la-gi-nhung-tinh-nang-vuot-troi/>

[20] Fernando Gutierrez. (2019, July 12), *ChainLocks and LLMQ-Based InstantSend mean Digital Cash.* Online: <https://blog.dash.org/chainlocks-and-llmq-based-instantsend-mean-digital-cash-efe2852da9ca>

[21] <https://www.binance.vision/blockchain/delayed-proof-of-work-explained>

HANU STUDENT CORE – AN ANDROID APPLICATION BASED ON EXISTING WEB SERVICE

SVTH: Đỗ Hải Bình

GVHD: ThS Luyện Thu Trang

Abstract— Smartphones have become increasingly popular around the world in recent years, therefore, smartphone applications, especially for Android and iOS, are greatly in need these days. For legacy system with outdated user interface that has little usability on handheld devices such as Hanoi University's Student Management system for Students, a smartphone application built on existing web service provides a convenient means of accessing student data on-the-fly. In this paper, the construction of Hanu Student Core application on Android platform will be examined.

Key words— Android application, student management, web service

I. INTRODUCTION

Hanoi University has been using an excellent student management system for student users which has a nice, simple to use user interface if viewed on desktop devices. However, in recent years, the main audience of the application has been students who mainly use smartphones to access the system, and the current system's user interface is not suitable for interaction on small-screen devices. As a regular user of the system, the author of this paper came up with an idea to build a web service exposing the current system's functionalities, along with an Android application to provide a mobile-friendly user interface to the user.

II. BACKGROUND KNOWLEDGE

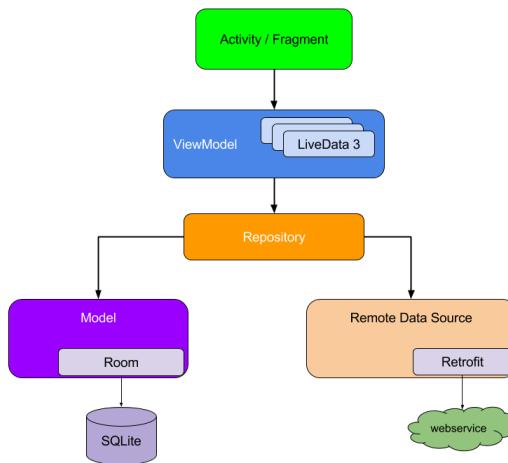
A. *Android User Interface's core components: Activities and Fragments*

In Android, Activities are among the most important components. Activities in Android represent an entry point to some of an application's functionalities; for instance: when users open the Facebook application from the Application drawer, the home page might be what they see, but if they use a “Share with Facebook” button, the Facebook app still launches but goes straight to the screen where user composes an article. However the application is structured, Activities remain one of the core building blocks.

Fragments represent a section of an activity that can be interacted with. Take Facebook application as an example again: In the main activity, there are sections like news feed, group feeds, profile, notifications and settings, each of which is a fragment. Fragments are different from Activities in the sense that Fragments cannot exist in isolation but have to be attached to an Activity host to function properly. Fragments' lifelines are the same as that of the Activity hosting them: When the activity is paused, hosted fragments are put in paused state; when the activity ends its lifetime, fragments inside it are also destroyed. When the activity is active, it can directly manipulate the lifecycle of its fragments.

B. *Android application recommended architecture: Model-View-View Model*

There are several ways in which an Android application can be built. The architecture can be Model-View-Controller, in which Views are layout files, Activities and Fragments act as Controllers, and the Models contain domain objects holding the data along with the mechanisms to interact with databases. However, Activities always have to do several view-related works regardless of the architecture being applied; therefore, Activities have more than one responsibility. This is a violation of Single Responsibility Rule which is a principle of Object-Oriented Design and should be avoided due to poor testability. Then came better architecture in terms of “Separation of Concerns” – the fact that software components have little dependence on each other so that the components can be reused without major modifications across different projects. They are: Clean Architecture – which is proposed by the author Uncle Bob in the book with the same name, based on Model-View-Presenter architecture pattern, and the recommended architecture by Google, Model-View-View Model. Both architecture patterns promote the separation of system components by defining several architectural layers and strict communication disciplines between them.



*Figure 1: Google's recommended Model-View-View Model architecture for Android
(Image courtesy of Android Developers)*

In Android's Model-View-View Model architecture, there are three layers:

- View layer: Comprises of Activities, Fragments and View components, the View layer exists to deliver the rendered graphical content to the device's screen and to listen for user interactions with the User Interface to trigger appropriate event handler.
- View Model layer: Consists of ViewModels that are used by Views to render content. ViewModels need to update their data regularly, which is facilitated by using implementations of LiveData interface.
- Data layer: This layer contains domain objects and their Repository to support data operations with both remote data source and local data source. Usage of this layer includes: getting data from a Web API, caching data to a local database provider such as the lightweight SQLite database through an interface like Room.

In this architecture, a layer may only have knowledge about the next layer. In other words: Views layer should only know what is in ViewModels, ViewModels should only know what the Repositories does and should never know what will be done in the Views layer. Therefore, the dependencies between system components are minimized. The author has applied Model-View-ViewModel architecture in Hanu Students Core application to demonstrate how the architecture works.

C. ASP.NET Core Web API

ASP.NET Core Web API is a web API technology based on Microsoft's .NET Core Framework. It provides enormous support on quickly developing a mock Web API that works right away. With its powerful Entity Framework, the Web API's connection to a database is greatly simplified and takes little time to set up. The framework also allows for automatic generation of REST API Controllers for given resources with a set of available HTTP methods to interact with. For this project, a web service providing data for display in the application is necessary, therefore I chose .NET Core to quickly draft a Web API application based on the identified resources.

III. HANU STUDENT CORE ANDROID APPLICATION

Hanu Students Core is a mobile application project with the aim to create an application that can replace the existing Student Management Web Application for Students for using on smartphone and small-screen devices. In the Core version, available functionalities include: Log in, log out, view profile, view student's related information including timetable, unpaid tuition fee, enrolled classes' information and marks, and make course registration. In this project, a mock web service written using ASP.NET Core framework was used locally in place of the university's real backend server, however, if the project catches the attention of Hanoi University's Student Management board, the web service will be adjusted to the university's current backend so that the API design of the web service stays intact.

D. Method

Due to the requirements being fixed from the beginning, the development of this project was based on Waterfall Software Development Lifecycle, with the following phases: Requirements definition, Analysis, Design, Implementation, Testing. In such a length-constrained paper, I will only discuss the most important phases in my development of this application.

1) Requirements definition

Based on the proposal, the system is only for Hanoi University students, therefore there is only one actor, Student. In this system, students can:

- Log in/log out
- View university's articles
- View personal profile

- View timetable for the current semester
- View information of an enrolled class
- View course marks
- View unpaid tuition fee
- Resolve unpaid tuition fee (mock feature, will work when the university allows the development of this feature)
- Make course registration choices during registration periods. This use case includes deletion of an existing class registration made during the same period.

2) Application design

The design of this application is derived from Android developers' recommended architecture – Model-View-ViewModel – discussed in Section II, part 2.b. Overall, the project is divided into main packages:

- base: Contains base interfaces and abstract class for implementation in other classes.
- model: Contains model classes corresponding to outputs from Web Service.
- repository: Contains repository classes with data operations, along with remote data source classes.
- ui: Contains fragments, activities and view components.
- .viewmodel: Contains ViewModel classes, each has access to corresponding repository.
- util: Container of helper classes that support operations of other classes in the application.

Due to the length restriction of this paper, class diagrams are not attached. They can be found in Hanu Students' Design Diagrams on Google Drive [Appendix. 4]

3) Implementation

a) Android application development

Due to the mobile application being more complex to build, the author made the decision to build it first. The Android application was developed feature-by-feature, with API connection to mock API service provided by Anypoint. In the development of each feature, the Remote Data Source is developed first, followed by Repository, then comes the development of individual view components within the fragment view, finally gluing them together with their Fragment. What takes the most time in developing this application is to build the User Interface using both XML markup and Java classes – which is much more difficult handle as compared to simply building HTML user interfaces for web applications.

b) Web API development

The Web API's development was simplified as much as possible. The author used one controller per resource, in which methods to handle specified HTTP methods in the API specification. To serve dynamic data, the Web API is connected to a reused database from a study project of mine in the previous semester. There are also Converters to handle the conversion from database format to the format that will be returned as specified in API specification.

```
[HttpGet]
public IActionResult GetTuitionFeesForStudent([FromQuery] string authToken)
{
    // authenticate
    string username = JwtUtils.ValidateJWT(authToken);
    if (username is null)
    {
        return Unauthorized(new MissingAuthTokenException());
    }

    int studentId = int.Parse(username);

    return new JsonResult(Context.Student
        .Include(s => s.Registration)
        .ThenInclude(r => r.FeeLine)
        .ThenInclude(f => f.Course)
        .Where(s => s.Id == studentId)
        .FirstOrDefault()
        ?.Registration
        .SelectMany(r => r.FeeLine)
        .Select(f => Converter.ForwardConverter(f))
        .ToList());
}
```

Figure 2: A normal .NET Core Web API Controller method with Entity Framework used.

This method corresponding to the endpoint method GET /tuition which returns the student's list of both paid and unpaid tuition fee up to the present. Note that students' usernames would always be their ID, which are integer numbers.

```
namespace HanuEdmsApi.Converter
{
    public class TuitionFeeConverter : OneWayConverter<FeeLine, TuitionFee>
    {
        public TuitionFeeConverter() : base(FromDatabase) { }

        private static TuitionFee FromDatabase(FeeLine feeLine)
        {
            var course = feeLine.Course;
            return new TuitionFee()
            {
                CourseName = course.CourseName,
                CreditCount = course.CreditCount,
                IsPaid = feeLine.Status,
                Value = (long)feeLine.LineSum
            };
        }
    }
}
```

Figure 3: A custom converter to handle the conversion from FeeLine entity to TuitionFee result as specified in the API design.

E. Results

After a long time working on this project, the complete Web API implementation was

done with data from the database of a previous project I have done on Student Management System. The Android application's alpha version was also completed with all listed use cases satisfied. Attached in this section are screenshots from the most outstanding use cases only. Use cases such as view profile or view course marks are skipped from this paper due to their simplicity: they only need to simply display the data in either textual form or tabular form.

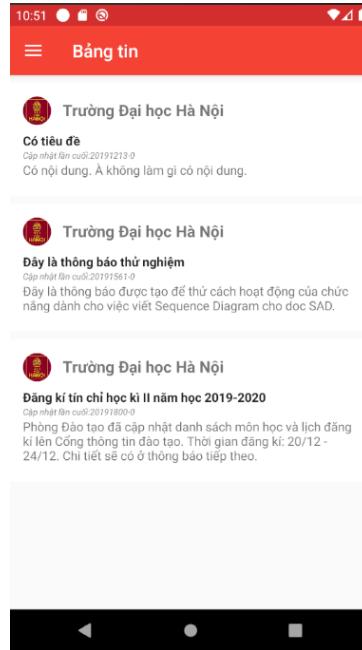


Figure 4: University articles. All university articles are displayed in this fragment.

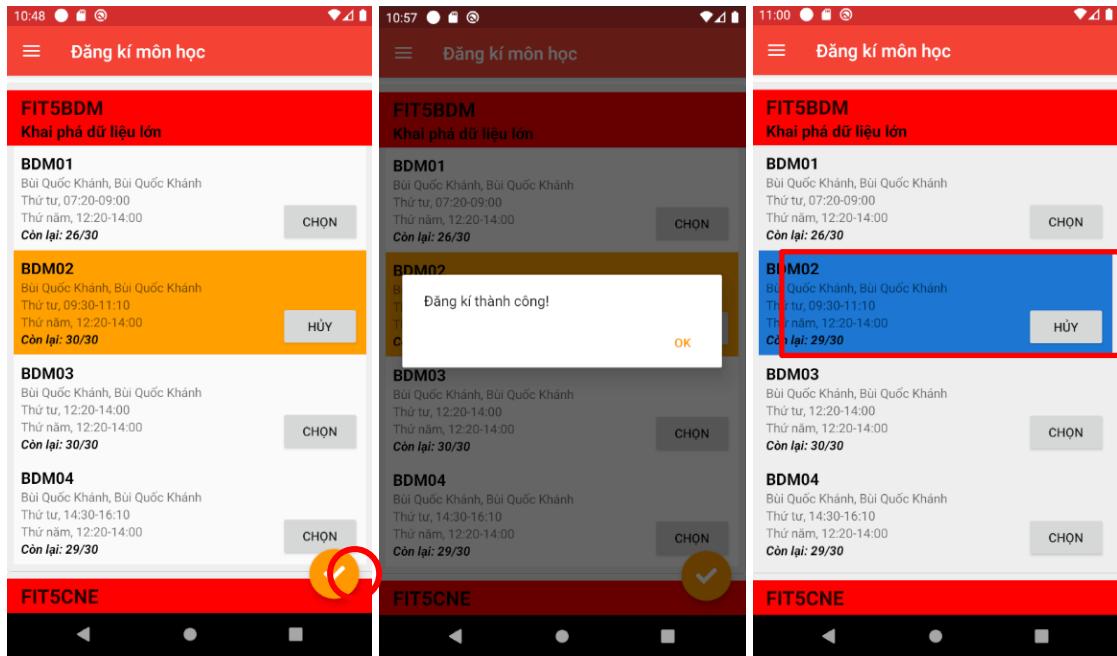


Figure 5: Register in course use case.

In this use case, students choose to enroll in several classes, then press the floating Save button to save their choices. If successful, the student is successfully enrolled and the class' number of remaining slots is automatically updated.

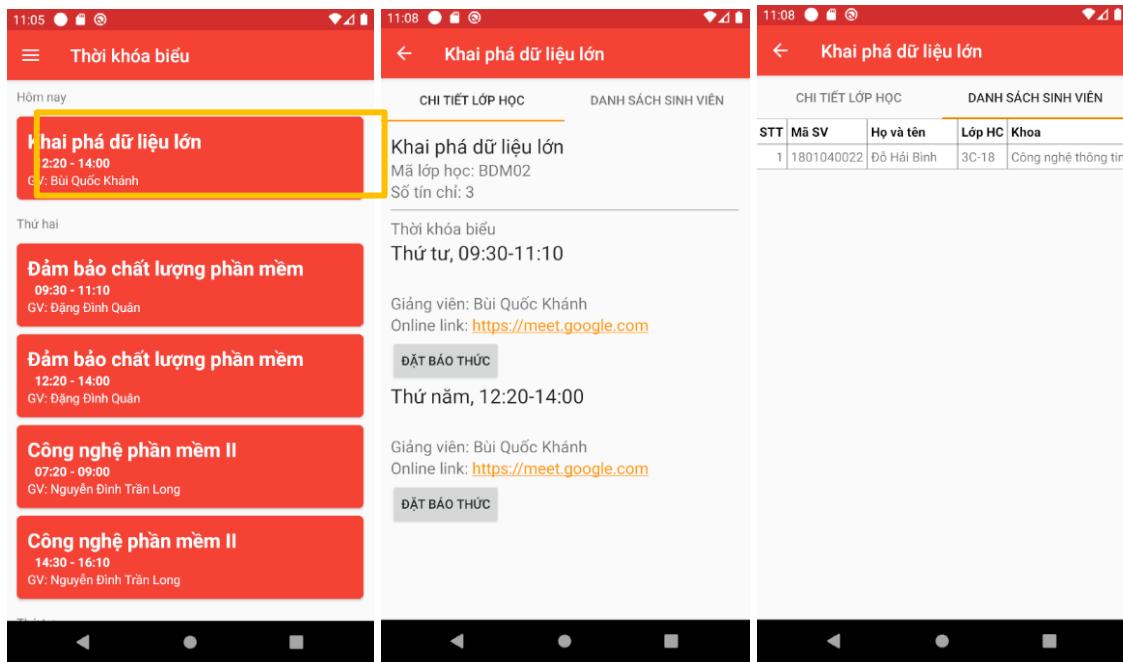


Figure 6: View Timetable use case.

In this use case, students are first presented with a list of all timetable periods of classes they enrolled in during course registration period. Then, students can tap on a course's timetable unit to view details about the course and the class, including the list of students who also enrolled in this class. In this screen, students can also create alarm for the class times of the class.

F. Evaluation

1) Advantages of the application

After a long and hard time spent on the development of this application, a stable working ALPHA version has been released, with all the features listed in Requirements definition section are implemented. The implementation strictly followed Google's recommended architecture for Android platform. The local web service is up and running, providing correct data format for use in the application. The user interface is mobile-friendly; features that are suitable for overall smartphone experience are implemented, for example: setting alarm for a class – in a very same way as how Google implemented in its Calendar application.

2) The application's shortcomings

However, such a development duration was not enough for an inexperienced Android application developer to justify the application in its finest; there are still a number of flaws ongoing within the application. Firstly, multilingual is not yet taken in consideration. Currently, the application only supports Vietnamese language, not yet having support for English. This can be resolved by using localized resource files in Android. Secondly, Android views are not optimized. The application can change to use better performing layouts and containers to improve performance, such as using Recycler

View in place of Linear Layout to display a list of items. Next, API calls are made every time the fragment is reload. Solution to this is to cache the fragment state in a local database such as SQLite and keep it for an amount of time. Moreover, force reload in the application may be implemented in the same way as applications like Facebook. Finally, support for real-time data update is not available. This should be done by employing a Cloud Messaging service like Google Cloud Messaging or Firebase to broadcast messages from the web service to subscribed devices. Using a cloud messaging service also makes it possible to make real-time push notifications, which are currently lacking in this application.

3) Future development

Based on what has been done and what had not been nailed after the development of the application, there are several jobs to be done in the future development of the application. Firstly, all the shortcomings of the application listed in the previous section need to be resolved, so that other aspects of the application can be continued. Next, other useful functionalities may be added to the application, such as: registering for dormitory, paying dormitory fee, and co-operate with a well-known banking service provider to finalize use cases concerning about paying money. In addition, the application can be expanded in such a way that a functional social network is added into the system to attract students to use this application. Specifically, the student can: find and add friends, make instant messaging with friends, join groups, publish articles in a similar way as Facebook. The backend of the application should then be able to automatically create student groups based on the list of enrolments in a course class. If those new functionalities are added into this application, Hanu Students application will provide a complete experience of the application to all students of Hanoi University.

IV. CONCLUSION

Building an Android application using a web service as data provider is becoming increasingly popular in recent years, due to the massive widespread of Internet usage around the world. While it is acceptable to develop a mobile-friendly application, it does not always provide the complete mobile experience to users, therefore, mobile applications play an important role in improving mobile user experience. In this paper, the creation of Hanu Students Core Application for Android platform along with the use of a local web service was examined. While the application is not the smoothest application to be used, it demonstrated what mobile applications can offer beyond the native user interface and experience. In the following time to come, the author will further develop the application so that its usability is improved, and it will attract more attention from students of Hanoi University, which are the target audience of this application.

REFERENCES

- [1] developer.android.com. *Introduction to Activities / Android Developers*. [online] Available at: <https://developer.android.com/guide/components/activities/intro-activities> [Accessed 5 May 2020]

- [2] developer.android.com. *Fragments / Android Developers*. [online] Available at: <https://developer.android.com/guide/components/fragments> [Accessed 5 May 2020].
- [3] developer.android.com. *Guide to app architecture / Android Developers*. [online] Available at: <https://developer.android.com/jetpack/docs/guide> [Accessed 5 May 2020]
- [4] guides.codepath.com. *Architecture of Android Apps / Codepath Android Cliffnotes* [online] Available at: <https://guides.codepath.com/android/Architecture-of-Android-Apps> [Accessed 5 May 2020]
- [5] Patni, S., 2017. *Pro Restful APIs*. Apress.

APPENDIX

- [1] Hanu Students application's repository | binhdoitsme/HanuStudentsAndroid, available on Github at: <https://github.com/binhdoitsme/HanuStudentsAndroid>
- [2] Hanu Students Web API's repository | binhdoitsme/HanuStudentsApi, available on Github at: <https://github.com/binhdoitsme/HanuStudentsApi>
- [3] Hanu Students (Core) API specification, available on Exchange at: https://anypoint.mulesoft.com/exchange/a4608c71-bee4-4869-a005-42360a72f385/hanustudents_core/minor/1.0/
- [4] Hanu Students' Design Diagrams, available on Google Drive at: https://drive.google.com/open?id=1qbmAJ84_T86D-p8CYEbSTeT6BuzDXG_C

EFFECTIVE VIRTUAL LEARNING ENVIRONMENT THROUGH MOODLE

SVTH: Nguyễn Khắc Hiếu

Nguyễn Thé An, Lâm Thị Thương Huyền

GVHD: ThS Luyện Thu Trang

Tóm tắt— Using the Internet to upgrade e-learning has become a pattern in present day advanced education foundations. E-learning systems are progressively turning into a significant part of the technique for conveying on the web and adaptable e-learning. The main advantage of e-learning is the opportunity for students to interact electronically with each other and their teachers during forums, on discussion boards, by e-mail and in chat rooms. In spite of the fact that perceiving that the world everywhere will keep on utilizing phrasing in various and regularly vague ways, the term of Virtual Learning conditions (VLE) is utilized to refer to the on-line connections of an assortment of sorts that occur among students and educators. There are numerous its of software accessible that give VLE systems, both business and Open Source Software (OSS). One such system that has been step by step picking up overall prominence is known as Moodle. This paper enters around this stage and on an examination between VLE (Moodle) and other VLE systems so as to find their qualities and confinements. The near examination is in two stages. The primary stage is in light of the highlights and capacities of VLE apparatuses and the subsequent stage depends on the specialized parts of the VLE stages.

Từ khóa— E-learning, Moodle

I. INTRODUCTION

Nowadays, people have become familiar with the so-called online learning. Although this concept has appeared and developed about 20 years ago, until now, online learning has affirmed its equal importance with direct learning. especially at present, when the COVID 19 pandemic is raging and has very serious consequences on all aspects of social life, and education are no exception. To ensure social distance, direct learning becomes impossible and the most effective solution now is online learning. Many online learning systems have emerged to meet the increasing learning needs. rising in the epidemic situation and Moodle emerged as a VLE with a lot of outstanding user-support features. Therefore, this paper, as the result of our research, aims confirm and prove the importance of Moodle on Virtual Learning Environment.

II. VIRTUAL LEARNING ENVIRONMENT THROUGH MOODLE

A. Literature Review

1) The learning environment

The literature has identified the importance of creating an appropriate e-learning environment. In particular, this environment needs to consider the issues of time, space and location as well as how e-Learning can be incorporated into classroom teaching and learning.

Creating such a learning environment has changed the way students or learners approach their learning. The documents note that students in e-Learning environments often control or take more responsibility for their learning.

2) What is Moodle?

Moodle is an open-source Learning Management System (LMS or otherwise called Course The executives' Framework or VLE - Virtual Learning Condition) (in this manner free and editable source code), to make seminars on the Web or web-based learning sites.



Moodle (short for Modular Object-Oriented Dynamic Learning Environment) was established in 1999 by Martin Dougiamas. Moodle stands apart as an education-oriented design for individuals working in the field of The Moodle learning management system is an open source, flexible and free to download as a learning management solution.

The Moodle learning management system is an open source, flexible and free to download as a learning management solution. With 68 million clients and 55,000 Moodle sites conveyed around the world, Moodle is user-friendly eLearning platform that considers the learning and preparing needs of individuals from a wide range of associations. In private or open Moodle locales, teachers, instructors, and businesses can make and convey online courses so understudies and their crowds can meet and surpass their learning objectives.

3) Moodle Architecture

Moodle runs on any web server that supports the PHP programming language and a database. It works best, and there is more support, when running on the Apache web server with a MySQL database. These requirements—Apache, PHP, and MySQL—are common to almost all commercial web hosts, even the lowest-cost ones.

While the Moodle data directory stores the files uploaded by students, the Moodle database stores most of the information in your Moodle site. The database stores objects that you create using Moodle.

Moodle has a modular architecture, shown in Fig. 1. In fact, Moodle is an acronym for Modular Object-Oriented Dynamic Learning Environment¹⁶, so it allows third party software to be installed and used on the platform. Therefore, the Forum participation plugin was made under Moodle architecture standards.

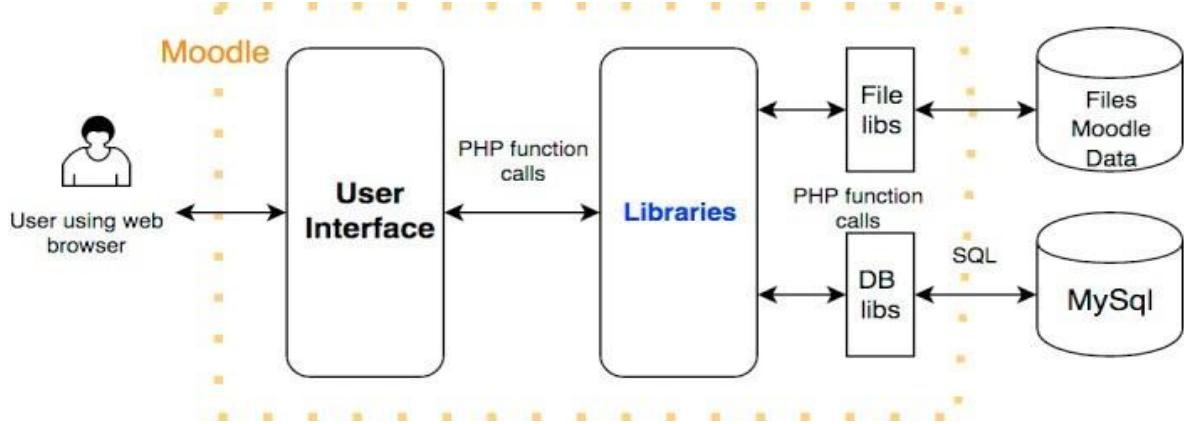


Figure1: Modular architecture of Moodle

The plugin architecture is based on three main parts. The first one is represented by all logical files that retrieve data from the database. The second part is formed by all working files to build matrices from the extracted data. The last one is formed by all the visual libraries that the plugin uses to present visual data to the user. Figure 2 shows all the plugin components, divided into two classes. The first layer displays all graphic libraries such as Bootstrap, Fixed-Header-Table, Jqplot, Jit, JQuery and J3V3. All of these libraries are used to display charts, such as the charts shown in the article. The second layer below is where the plugin business logic takes place. In this class, there is a main file called SNALIB; This PHP file has all the logic to do data retrieval and to build collaboration and matrix matrices. To visualize all the results, specific functions have been implemented in logic to construct each graph result. Then, a JSON object is created and passed as a parameter to the PHP entry page of each library mentioned earlier.

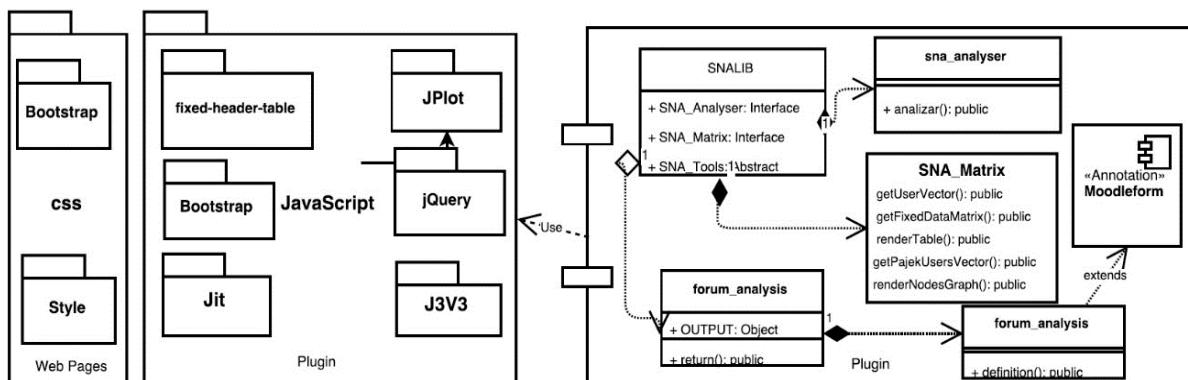


Figure2: Moodle Forum participation plugin components architecture.

B. Method

Data were collected in two stages using two groups of students in order specifically to

achieve objectives one and four of this research, respectively.

1) Stage One

A total of 80 undergraduate students who were registered on two courses related to human computer interaction and software testing at the Faculty of Information Technology at Zarqa University in Jordan were asked to participate. The students who participated in this study were given extra marks. However, their participation was voluntary.

A questionnaire was designed which aimed to gather data from the students regarding their experience of Moodle while using both desktop/laptop devices and mobile/tablet devices. It comprised two sections: Section 1 involved closed-questions which were designed to gather students' background information, Internet experience and Moodle experience while Section 2 included the following four open-ended requests:

- List the features which you are using in Moodle.
- List the features which you would like to use in Moodle but which are not supported by it.
- List the weak design features in Moodle which prevent you from interacting with it successfully.
- Suggest improvements to the design of Moodle to make it more usable.

Questionnaires are uploaded to Moodle as homework for students who have enrolled in two courses. The students were given two weeks as a deadline and they were asked to submit their answers to the questionnaire to Moodle. Students using only desktop / laptop devices to access Moodle were asked to answer questions once, while students used both desktop / laptop devices. hand and mobile / tablet device to access Moodle is required to respond twice to collect data about their experience with Moodle using different interfaces.

Data obtained from the questionnaire was translated into English from Arabic. Data is analyzed to explore student experience with Moodle features and usability. Descriptive analysis is used to describe the characteristics of students and their experience of the Internet and Moodle. Analyzing the answers for each of the four open questions follows the same procedure; The answers to each question have been checked and classified. They have created popular subtopics (subcategories) that suggest corresponding topics (categories).

2) Stage Two

A total of 240 undergraduate students from four faculties: Faculty of Information Technology, Faculty of Business, Faculty of Nursing, and Faculty of Engineering at Zarqa University were asked to participate. Their participation was voluntary.

A questionnaire was developed based on the students' responses to the questionnaire developed in Stage One to confirm the data collected during that period. It consists of two parts: Part 1 involves closed-ended questions designed to collect basic student information, Internet experience, and Moodle experience. However, Section 2 involves the four sections

corresponding to the four open-ended questions given to students in Stage One. Each section includes questions that are sub-topics created from the analysis of student responses for each of the four open questions collected in Stage one. Students are asked to rate the sub-topics that are determined based on their relevance to each statement based on the seven-point scale (Likert scale).

In order to collect data from students about their experience with Moodle about the type of interface they used, the questionnaire was designed to include Section 1 and two copies of Section 2. Students should use only desktop / laptop devices to access Moodle are required to answer Section 1 and Section 2 of the questionnaire, while students use both desktop / laptop devices and Mobile / tablet devices to access Moodle are required to answer Part 1 and two copies of Section 2 to collect data about their experience with Moodle using two interfaces. It is worth mentioning that the questionnaire was tested using five college students from the same university where the study was conducted. The pilot study identified some ambiguities in the questionnaire so that the results from the pilot trial were reviewed and the changes were made. The data is then collected over a period of three weeks.

Descriptive analysis is used for Section 1 of the questionnaire to describe the characteristics of students and their experience of the Internet and Moodle. Student characteristics are presented in the Results Section. The Likert score is calculated for each statement in Section 2 to describe student confirmation of sub-topics identified in phase one of the study. It is worth mentioning that, for the purpose of analyzing using the Likert scale in this study, the Likert 1-3 score is considered to be a negative reaction, 5-7 positive reactions and 4 neutral points. Likert scores for statements (subtopics) and their explanations are presented in the Results Section. The version of Moodle LMS used by the university is 2.9. However, some plug-in updates have been installed by college administrators for updates.

C. Results

1) Student's Characteristics

A total of 71 students out of 80 (88.75%) participated in Phase One of this study and uploaded the answers to the questionnaire; 35 of them answered Part 2 of the questionnaire twice and expressed their experience with Moodle while using both desktop / laptop devices and mobile / tablet devices. The majority of students are between the ages of 18 and 22. Most of the students (70%) are male while women make up 30% of the participants. Half of the students come from the Faculty of Software Engineering (65%). Students from other faculties also participated, namely from: Computer Science (16%); Computer information systems (12%) and Internet Technology (7%). Students also participate in the second year (18%) in the third (52%) and the fourth (30%) of the school year. The majority (92%) have more than three years of experience using the Internet and all students use the Internet daily. Regarding the student experience with Moodle, most (82%) have more than three semesters. Experience using Moodle and most (72%) using Moodle daily.

In addition, a total of 225 students out of 240 (90.75%) participated in Phase Two of

this study. The majority of them are between the ages of 18 and 22. Men represent (60%) of the students while women make up 40%. Regarding their Faculty, 40% of students come from the Faculty of Business; 30% of students come from the Faculty of Engineering; 15% is from the Faculty of Nursing and the rest (15%) is from the Faculty of Information Technology. Students also participate in the second year (12%) in the third (33%) and the fourth (55%) in the school year. The majority (96%) had more than three years of experience using the Internet and all students use the Internet daily. Regarding the student experience with Moodle, most (95%) have more than three semesters of experience using Moodle and most (83%) use Moodle daily.

2) Students' Experience in Moodle

This section presents the results obtained from analysing the data that were collected from the two groups of students concerning their experience of the features and usability of Moodle on desktop/ laptop devices and mobile/tablet interfaces. It includes four subsections related to: features supported by Moodle; features not supported by the local instance of Moodle; usability problems in Moodle; and suggested improvements to the design of Moodle. Each subsection presents the results obtained from Group one of the students followed by the results obtained from Group two.

VLEs, as e-learning systems, have many features and capabilities but in order to simplify and clarify the comparison, we have divided these features and capabilities into three phases, which are Learner Tools, Support Tools and Technical Tools, as in Tables 1, 2 and 3. Also, this comparison has two kinds of answers Yes (Y) or No (N).

No	1	2	3	4	5	6
Product name Tools	Design2 Learn 8.1	KEWL	ANGEL (7.1)	eCollege	The Blackboard System	Moodle 1.8
1. Learner Tools						
1.1. Communication Tools						
Discussion	Y	Y	Y	Y	Y	Y
Forums						
Discussion	Y	Y	Y	Y	Y	Y
Management						
File Exchange	Y	Y	Y	Y	Y	Y
Internal Email	Y	Y	Y	Y	Y	Y
On-line Journal	Y	Y	Y	N	Y	Y

Real-time Chat	Y	Y	Y	Y	Y	Y
Video Services	N	N	N	N	N	Y
Whiteboard	Y	N	Y	Y	Y	Y
1.2. Productivity Tools						
Bookmarks	Y	Y	Y	Y	N	N
Calendar	Y	Y	Y	Y	Y	Y
Orientation	Y	Y	Y	Y	Y	Y
Searching Course	Y	Y	Y	Y	Y	Y
Work Off-line	Y	Y	Y	Y	Y	Y
Group work	Y	Y	Y	Y	Y	Y
Community	Y	Y	Y	Y	Y	Y
Student Portfolios	Y	Y	Y	Y	Y	Y
Total Features	16	16	16	16	16	16
Total Available	15	14	15	14	14	15
Total Missing	1	2	1	2	2	1

Table 1: The Comparison between the Selected VLE Products based on Learner Tools.

As we can see in Table 1, the comparison between the VLE products is based on Learner Tools. Four products are shown to be the best with almost the maximum number of features - 15 out of 16 features or capabilities of Learner Tools. These products are Moodle, Desire2Learn, ANGEL. KEWL, eCollege and The Blackboard

Learning System platforms have missed 2 out of 16. Moodle is the best with three products missing only one feature. Overall the best OSSs are Moodle and respectively, which missed 1 out of 16 Learner Tools.

As we can see in Table 2, the comparison between the VLE products is based on Support Tools.

2. Support Tools						
2.1. Administration Tools						
Authentication	Y	Y	Y	Y	Y	Y
Authorization	Y	Y	Y	Y	Y	Y
File exchange	Y	Y	Y	Y	Y	Y
Registration	Y	Y	Y	Y	Y	Y
Integration						
2.2. Course Delivery Tools						
Test Types	Y	Y	Y	Y	Y	Y
Automated Management	Y	Y	Y	Y	Y	Y
Automated Support	Y	Y	Y	Y	Y	Y
Course Management	Y	Y	Y	Y	Y	Y
On-time grading	Y	Y	Y	Y	Y	Y
Student Tracking	Y	Y	Y	Y	Y	Y
2.3. Content Development Tools						
Accessibility	Y	Y	Y	N	Y	Y
Content sharing	Y	Y	Y	N	Y	Y
Course Template	Y	Y	Y	Y	Y	Y
Look and Feel	Y	Y	Y	Y	Y	Y
Design	Y	Y	Y	Y	Y	Y
Instructional Standards	Y	Y	Y	Y	Y	Y
Total Features	16	16	16	16	16	16
Total Available	16	16	15	15	16	16
Total Missing	0	0	1	1	0	0

Table 2: The Comparison between the Selected VLE Products based on Support Tools.

In this phase, all products have all features and capabilities except eCollege and The Blackboard Learning System (V.7). This means that Moodle and the other remaining products are strong on Support Tools.

As we can see in Table 3, the comparison is based on Technical Specifications Tools.

3. Technical Specifications						
3.1. Hardware/Software Tools						
Client Required	Y	Y	Y	Y	Y	Y
Database Requirements	Y	Y	Y	Y	Y	Y
Unix Server	N	N	N	N	Y	Y
Windows	Y	Y	Y	N	Y	Y
3.2. Pricing/Licensing Tools						
Company Profile	Y	Y	Y	Y	Y	N
Costs	N	N	N	N	N	Y
Open Source	N	N	N	N	N	Y
Optional Extras	Y	Y	Y	Y	Y	N
Total features	8	8	8	8	8	8
Total available	6	5	5	4	6	7
Total missing	2	2	3	3	4	2

Table 3: The Comparison between the Selected VLE Systems based on Technical Specifications Tools.

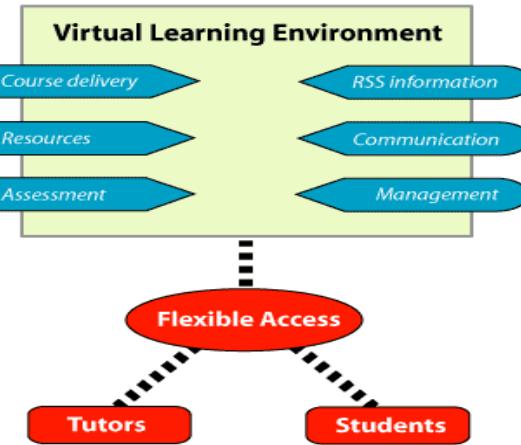
In this phase, the best products are Moodle 1.8, which have missed only 1 out of 8 Technical Specifications Tools, and then Desire2Learn 8.1, The Blackboard which missed 2 out of 8. The weakest products are KEWL which missed 4 out of 8 Technical Specifications Tools.

By some simple comparisons, it is possible to show the optimization of moodle compared to the rest of the learning tools. Therefore, it is reasonable to confirm the importance of Moodle to virtual learning environment

D.DISCUSSION

1) Virtual Learning Environment

A virtual learning environment is an online-based platform that offers students and professors digital solutions that enhance the learning experience. Unlike a virtual classroom, which is meant to replicate and replace the physical classroom environment for distance learners, a virtual learning environment (or VLE) harnesses technology to supplement an in-class experience, with, for example, digital communication, interaction and quizzes or polls run through the VLE.



The virtual learning environment has been linked to formal learning and to the relationships between teachers, students and schools. There is a growing interest in virtual learning environments supported by the internet, in particular between educational institutions, students and teachers. The concept of virtual learning environment (VLE) can be considered a dynamic concept due to the continuous development of digital technologies, its features and potential and the importance of environments in learning progress.

Web-based education systems are being used by more and more universities, schools and companies, not only to incorporate web technology into their courses, but also to complement their courses. face to face with their traditions. These systems collect large amounts of valuable data for course content analysis and student use. The learning environment based on the use of technology and digital resources are intermediaries in the learning process through the activities they allow. This is due to the fact that they facilitate interaction and interaction in a continuous communication process, thereby increasing the building and refactoring of knowledge and meaning as well as habit formation. and attitude within the general framework for all those involved education process.

The use of VLE in each context implies recognition of their main features and potentials. The learning environment and context are dynamic and multi-dimensional concepts emerging from new educational concepts and practices in the digital society.

The features and potentials of VLE turn them into spaces that allow for testing, promotion and support of well-planned and oriented teaching and learning strategies. Observations of constant dynamism are encouraged in the use of resources and in the changes witnessed around them, as this will allow them to be considered as the context for building. build learning processes.

The virtual learning environment allows learning to take place according to the elements in the learning environment, based on a continuous scale from the specified elements in the environment to the emerging elements when used.

2) The Reasons for Choosing Moodle

The importance of Moodle is that it is well evaluated in many reports, has a high level of acceptance in the community and in some organizations, and has many active courses,

available in many languages. It gives users the ability to post news articles, assignments, ezines and resources, and to collect assignments, etc. Moodle's greatest strength is the community that has grown around the project; Both developers and users participate in Moodle's active discussion forums, share tips, post code, help new users, share resources, and debate new ideas. Therefore, we have chosen Moodle's software as our field of research and analysis. It is important to understand the Moodle environment, and explore its functions and limitations to develop practical examples for the use of VLE in Qassim University. We list here the most important reasons to choose Moodle:

1. Moodle can be used on almost all servers that can use PHP. Users can download and use it on any computer and can easily upgrade it from one version to the next.

2. The key to Moodle is that is developed with both pedagogy and technology in mind. One of the main advantages of Moodle over other systems is its strong grounding in social constructionist pedagogy and good educational tools.

3. The Moodle software is used all over the world by independent teachers, schools, universities and companies. The credibility of Moodle is very high. Currently, there are 3324 web sites from 175 countries that have registered with it, and it has 75 languages.

4. Moodle runs without modification on any system that supports PHP such as Unix, Linux and Windows. It uses MySQL, PostgreSQL and Oracle databases, and others are also supported.

5. It has many features useful to potential students such as easy installation, customization of options and settings, good support/help, and good educational tools. Moreover, it has excellent documentation, and strong support for security and administration.

6. It is an OSS, which means users are free to download it, use it, modify it and even distribute it under the terms of the GNU license.

7. It is a CMS & VLE that lets teachers provide and share documents, graded assignments, discussion forums, etc. with their students in an easy-to-learn fashion, and in high quality on-line courses.

3) The Limitations of Moodle

Moodle's low cost, flexibility and ease of use helps bring VLE technology within the reach of those with limited technical or financial resources. On the other hand, Moodle has some limitations as follows:

1. Moodle is only for IT experts. It is complex for normal users to use and more than 66% of them are teachers, researchers and administrators.

2. It is difficult for beginner technicians to install and use Moodle, because there are many technical word lists in installation instructions.

3. Moodle will work, but not by itself. If there is not a course administrator that can work with both teachers and technicians in creating on-line materials, then Moodle will remain an empty shell, like a good aircraft but with no pilot;

4. Lack of simple-to-obtain support. Forums carry a great deal of information, but nearly all forums are in the English language.

III. CONCLUSION

In the conclusions, we present a summary of the main results obtained from a study conducted in the academic year of 2014-2015, involving 6347 undergraduates from a Portuguese public higher education institution. This study aimed to assess the frequency of students' access to the virtual learning environment (VLE) adopted by the institution, as well as the relation between the frequency of access to the VLE and the students' performance. The nature of the study was quantitative and the data was obtained from databases associated with the VLE.

Considering the great variability of the number of accesses to the VLE, groups of students were formed according to the number of their accesses to the VLE so as to assess the relations between the groups' accesses to the VLE and their performance. The results obtained regarding students' performance enable us to infer that the higher the mean of the group's accesses to the VLE is, the higher are: the number of course units in which the student is registered, the number of units they passed, the percentage of units they passed relatively to the units they are registered in, and the percentage of course units the student passed. Also, the higher the mean of the group's accesses to the VLE is, the lower the percentage of students who failed all the course units is.

The results concern only one higher education institution and therefore, cannot be generalised. However, these results show relatively positive indicators regarding students' access to a virtual learning environment and the relation between such access and their performance.

This paper has made a comparative study between Moodle and other VLE systems, and this was based on two kinds of comparison. The first phase was based on the features and capabilities of VLE tools, and the second one was based on the technical aspects of VLE systems. From this paper, we aimed to discover the best and most suitable choice of VLE systems that would meet the requirements of Qassim University. In this, our initial assessment, we have succeeded in finding that optimal VLE platform, and it is Moodle.

REFERENCES

- [1] Valsamidisa, S., Kazanidisa, I., Petasakisa, I., Kontogiannisb, S., & Kolokithaa, E. E-Learning Activity Analysis. Procedia Economics and Finance. 9. pp. 511 – 518. 2014.
- [2] Becerra, E., García, M. C., & Chávez, R. Ambiente de aprendizaje con uso de tecnología en la formación docente inicial y las habilidades intelectuales. In R. E. Navarro, M. J. Pacheco, Y. N. Rangel, & M. S. Montoya (Eds.), Foro interregional de investigación sobre entornos virtuales de aprendizaje: Integración de redes académicas y tecnológicas. México: Redtic. pp. 154–164. 2011.
- [3] Morais, C., Alves, P., & Miranda, L. Valorização dos ambientes virtuais de aprendizagem por professores do ensino superior. In Á. Rocha, L. Reis, M. Cota, M. Painho, & M. Neto (Eds.), Sistemas e Tecnologias de Informação. 8^a Conferência Ibérica de Sistemas

e Tecnologias de Informação. Lisboa: AISTI/ISEGI. pp. 289-294. 2013.

[4] Zitter, I., Bruijn, E., Simons, P., & Cate, T. Adding a design perspective to study learning environments in higher professional education. *High Educ* 2011. 61. pp. 371-386.

[5] Dahlstrom, E., Brooks, C., & Bichsel, J. The Current Ecosystem of Learning Management Systems in Higher Education: Student, Faculty, and IT Perspectives. Research report. Louisville, CO: ECAR, September 2014. Available online: <http://www.educause.edu/ecar>.

[6] Alves, P., Miranda, L., & Morais, C. Open Educational Resources: Higher Education Students' Knowledge and use. In Rikke Ørnsgreen & Karin Tweddell Levinsen (Eds.), *Proceedings of the 13th European Conference on e-learning ECEL 2014*. Reading, UK: Academic Conferences and Publishing International Limited. 2014. pp. 11- 18

[7] Agudo-Peregrina, A., Iglesias-Pradas, S., Conde-González, M., & Hernández-García, A. Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning. *Computers in Human Behavior* 2014. 31. pp. 542–550.

[8] Gašević, D. & Siemens, G. Let's not forget: Learning analytics are about learning. *TechTrends* 2015. 59(1). pp. 64-71

[9] Siemens, G., & Gašević, D. Guest Editorial - Learning and Knowledge Analytics. *Educational Technology & Society* 2012. 15(3). pp. 1-2.

[10] Ferguson, R. The state of learning analytics in 2012: A review and future challenges. Technical report KMI-12-01. UK: Knowledge Media Institute, The Open University. 2012.

[11] Dawson, S., Gašević, D., Siemens, G., & Joksimovic, S. Current State and Future Trends: A Citation Network Analysis of the Learning Analytics Field. In *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*. New York, NY, USA: ACM. 2014. pp. 231–240.

[12] Greller, W., & Drachsler, H. Translating learning into numbers: A generic framework for learning analytics. *Educational Technology & Society* 2012. 15(3). pp. 42–57. Available online:

<https://www.semanticscholar.org/paper/Translating-Learning-into-Numbers%3A-A-Generic-for-Greller-Drachsler/d1bd219962defaeb326c3b51fb4fb1086c5b7b28>

[13] ECAR Working Group. The Predictive Learning Analytics Revolution Leveraging Learning Data for Student Success. Available online: <http://net.educause.edu/ir/library/pdf/ewg1510.pdf>.

SOCIAL NETWORK PLATFORM

SVTH: Nguyễn Thu Hà, Phạm Quang Chiến

Vương Khánh Linh, Trần Thu Hiền

GVHD: ThS Luyện Thu Trang

Tóm tắt: Mục đích của nghiên cứu này là xem xét kỹ lưỡng lý do, lợi ích, vấn đề và giải pháp gặp phải trong việc sử dụng mạng xã hội dựa trên ý kiến của sinh viên Đại học Hà Nội, từ đó, chúng ta có thể tạo ra một mạng xã hội hữu ích. Nghiên cứu này được thực hiện bằng cách sử dụng kỹ thuật bảng câu hỏi để thu thập dữ liệu. Do đó, bảng câu hỏi được tổ chức với 100 sinh viên ngẫu nhiên. Dựa trên kết quả nghiên cứu, người ta thấy rằng có nhiều lý do khiến sinh viên sử dụng mạng xã hội như kết nối và giữ liên lạc với bạn bè và gia đình một cách dễ dàng, chia sẻ cảm xúc hoặc câu chuyện với người khác và giảm căng thẳng. Một điểm quan trọng khác trong cuộc khảo sát này là ý kiến của sinh viên về mạng xã hội, liệu là nó tốt hay không. Nếu không, giải pháp cần thiết để cải thiện mạng xã hội là gì. Do đó, người ta tin rằng báo cáo này có thể phục vụ như một nguyên mẫu cũng như một khuyến khích tiềm năng cho các nghiên cứu lớn hơn và xa hơn trong tương lai.

Từ khóa: Social network platform, REST Architecture Style, Software Development Life Cycle (SDLC)

Abstract: The purpose of this study is to scrutinize the reasons, benefits, problems and solutions encountered in the usage of social networking based on opinions of Hanoi University's students, thence, we can create a useful social media web. This study was conducted using the questionnaire technique to collect data. Therefore, questionnaire was held with 100 random students. Based on the study's results, it was found that there are various reasons why students use social media such as connecting and keeping in touch with friends and family easily, sharing emotion or story with others and reducing stress. Another key point in this survey is the students' opinions about social networking, whether it is good or not. If it does not, what is the solution needed to improve social networking. As such, it is believed that this report can serve as a prototype as well as a potential stimulus for larger and further studies in the future.

Keywords: Social network platform, REST Architecture Style, Software Development Life Cycle (SDLC)

I. INTRODUCTION

With the explosion of internet, social networks are becoming increasingly popular in our society. Thanks to social media, we can communicate easily with others regardless of geographical distance and time. However, social network platforms also bring many troubles for users as well as administrators such as security errors, annoying advertisements, slow processing speed and difficult interface to use. Thus, based on collected users' requirements, we aim to conduct Social Media Web Project which creates a simple and convenient social media web and focuses more on restful architecture.

II. METHODS AND RESULTS

Software Development Life Cycle (SDLC) is the development approach for our project which is a process used by the software industry to design, develop and test high quality softwares. The SDLC aims to produce a high-quality software that meets or exceeds customer expectations, reaches completion within times and cost estimates**Invalid source specified.** SDLC divides development processes into the following five phases: analysis, design, implementation, testing, and maintenance.

A. Analysis and Requirement

1. Methods

The data used in this report is acquired from a survey carried out within Hanoi University. At the beginning, a questionnaire consisting of 8 questions was designed to capture requirements of users with a social media web. In the next step, about 100 students of different ages and academic years were randomly selected, and each of them respectively asked the 8 questions. Therefore, we were confident that the data obtained achieved a fairly high level of genuineness.

2. Results

This section outlines the results of the survey conducted to discover HANU students' opinions of the usage of social networking. With a view to determining the kinds of functions of social networking which are popular among the students of Hanoi University, a great deal of the typical features were listed in the questionnaire. In order to simplify the data interpretation, the different requirements have been combined and labeled under the headings "user interface", "security" and "communication and interaction tool".

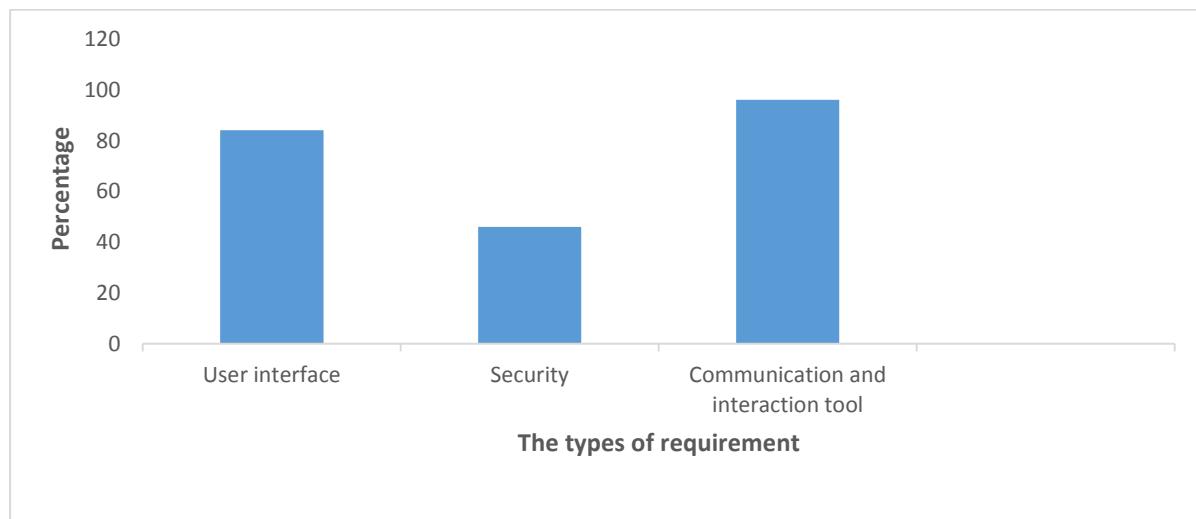


Figure 14: The percentage of the students setting the requirements

As can be seen from Figure 1, the number of the students desiring good communication and interaction tools(such as chat box, status post) occupies the first place, reaching 96% of

the total. Meanwhile, the figure for the students requiring user interface ranks the second, with 84%, is significantly more than that of security (only at 46%).

Especially, the required user interface is clear, convenient and easy to use and includes different tabs with title according to its function. With regard to security, every user has a distinct account and individual page, which helps users to avoid being hacked individual information and account.

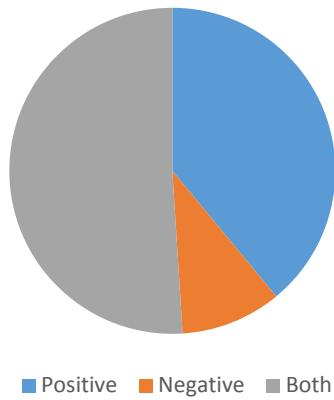


Figure 2: The percentage of the students' attitudes towards social media' effects

Turning to the social network platforms' effects, more than a half (51%) of those who responded were of the opinion that social media has both negative and positive impacts on them. Whereas, 39% of the people replied that social networking is positive, and, only 10% believed that it has negative influences.

Therefore, we believe that creating a social network platform is really useful, which may meet the essential needs of people.

B. Design

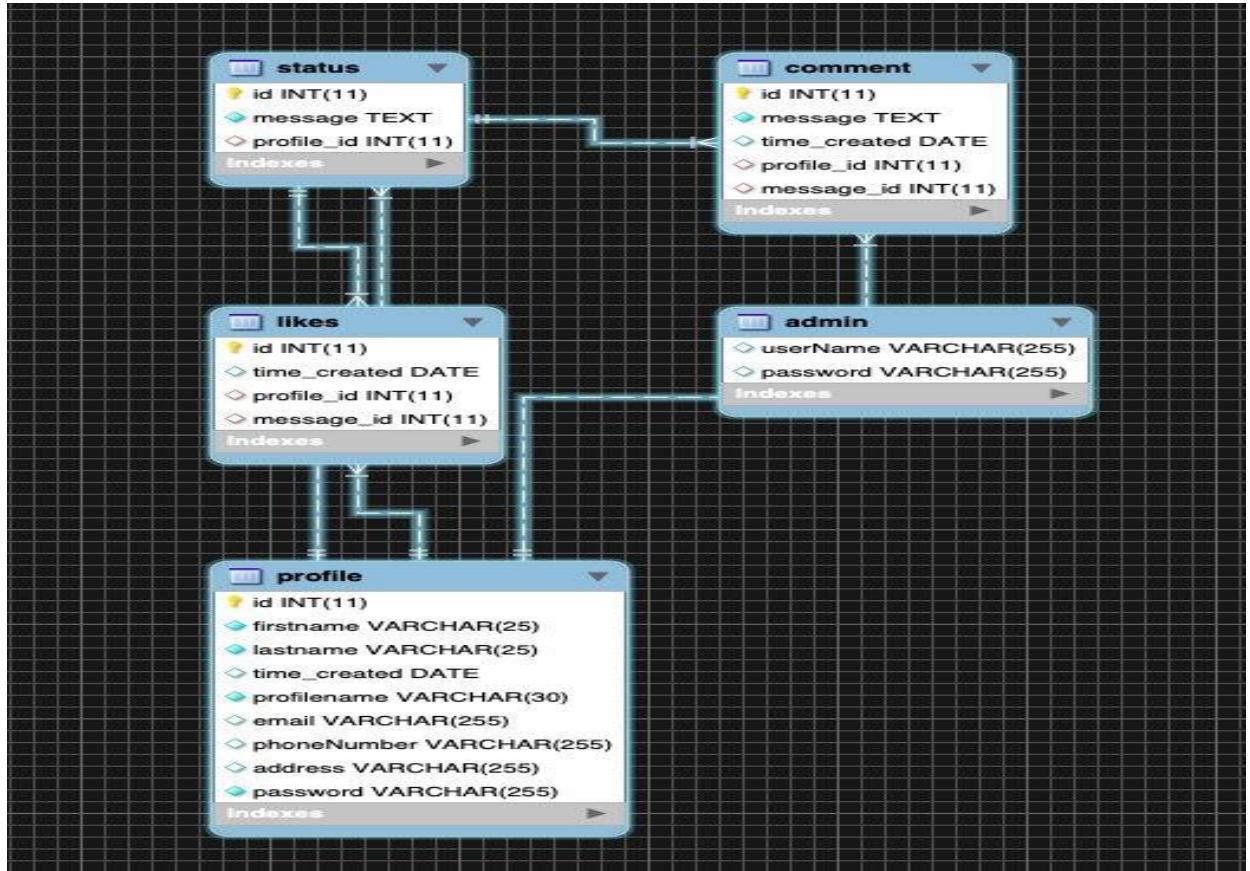
1. Methods

Based on these above requirements, design provides a comprehensive design overview of the system, using a number of different design views to depict different aspects of the system, such as Entity Relationship Diagram and Data Flow Diagram. It is intended to capture and convey the significant design decisions which have been made on the system.

2. Results

The below entity relationship diagram (ERD) shows the relationships of entity sets stored in our database and sketches out the design of database**Invalid source**

specified. Creating the necessary entities(such as comment, status, profile, like) helps us to develop useful communication and interaction functions for users.



Our project has two main processes including chatting and posting status, and the following data flow diagrams (DFDs) map out the flow of information for these two processes

Invalid source specified.. Chatting and posting status are two vital processes which assist us in creating simple and user-friendly social media web.

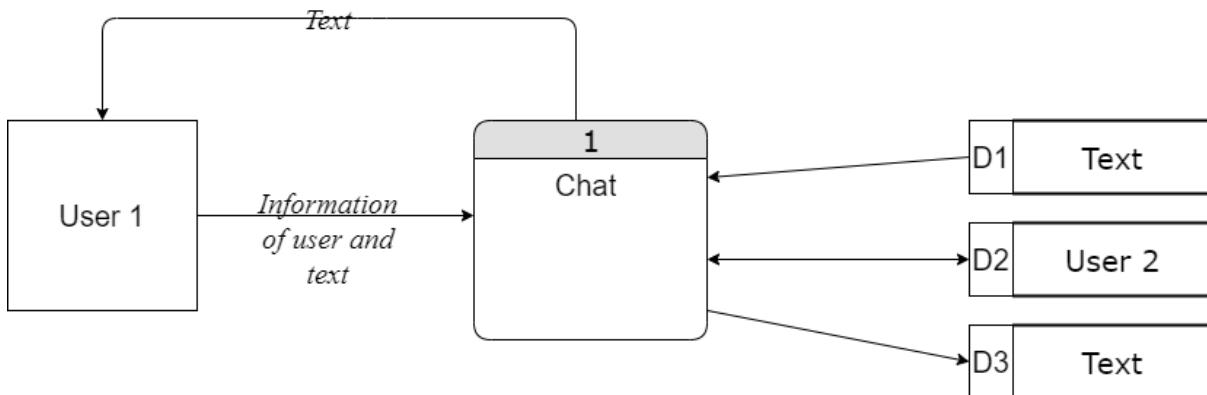


Figure 3: Chatting data flow chart

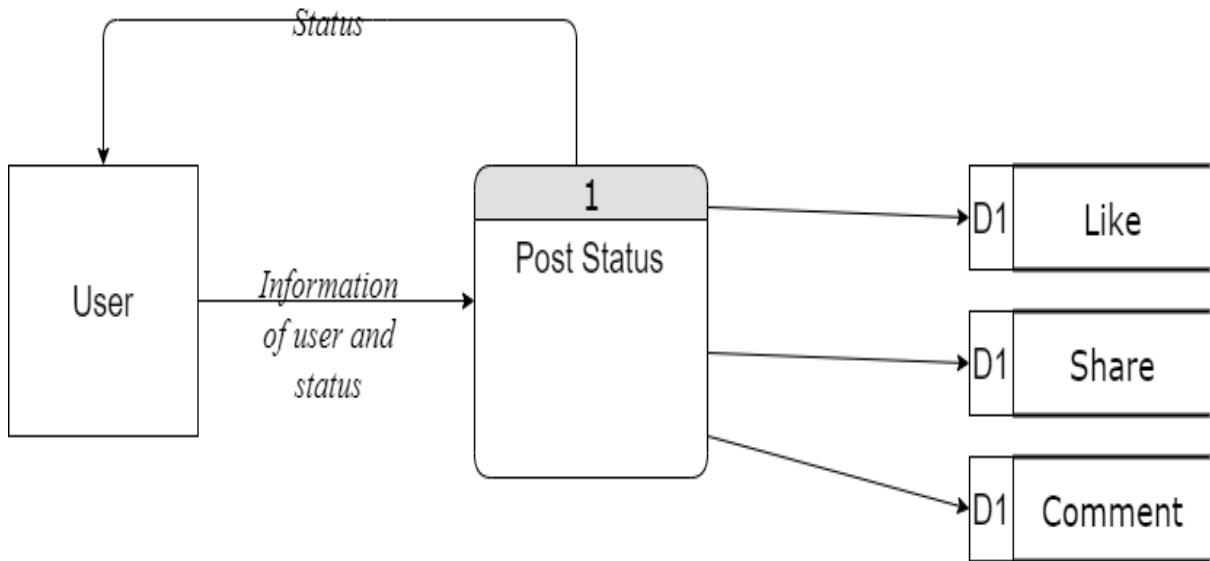


Figure 4: Posting Status data flow chart

C. Implementation

1. Methods

To develop our social media web, we use Tomcat version 9.0 web server, MySQL database server and Java programming language. Besides, we also attach special importance to using Restful architecture that defines the set of rules to be used for creating web services. Web services which follow the REST architectural style are known as RESTful web services. It allows requesting systems to access and manipulate web resources by using a uniform and predefined set of rules. Interaction in REST based systems happen through Internet's Hypertext Transfer Protocol (HTTP). In our project, we use jax-rs interface and jersey which are implementation of jax-rs to create Rest APIs.

2. Results

To concretize REST architecture style, we exemplify the following profile rest. The profile model helps to declare essential attributes for profile of user. The profile DAO roles play as data layer for profile of user in database. Profile DAO conforms to CRUD(create, read, update and delete). Next, we provide profile service handling business logic(post, get, update and delete data from data layer). Moreover, profile service also helps get expected data range or data by year. The profile resource implements sufficiently get, post, update and delete method. Profile resource returns both JSON data and XML data, conforms to REST Architecture Style. Finally, we use postman to get JSON data from our social media web.

```

1 {
2   "created": "2020-05-24Z",
3   "firstName": "Ha",
4   "id": "ThuHa219",
5   "lastName": "Nguyen",
6   "links": [
7     {
8       "rel": "self",
9       "uri": "http://localhost:8080/social-media-platform/webapi/profiles/ThuHa219"
10    }
11  ],
12  "password": "123456",
13  "profileName": "ThuHa219"
14 }

```

In terms of social media web for users, we represent user interface of two main pages including chatting and posting status.

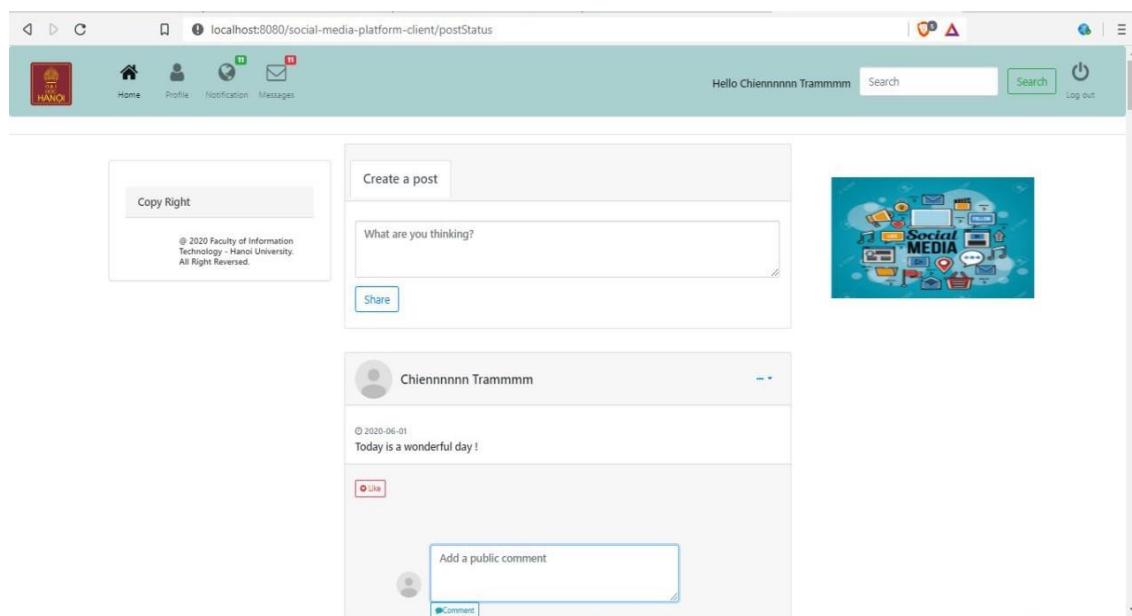


Figure 5: Posting status

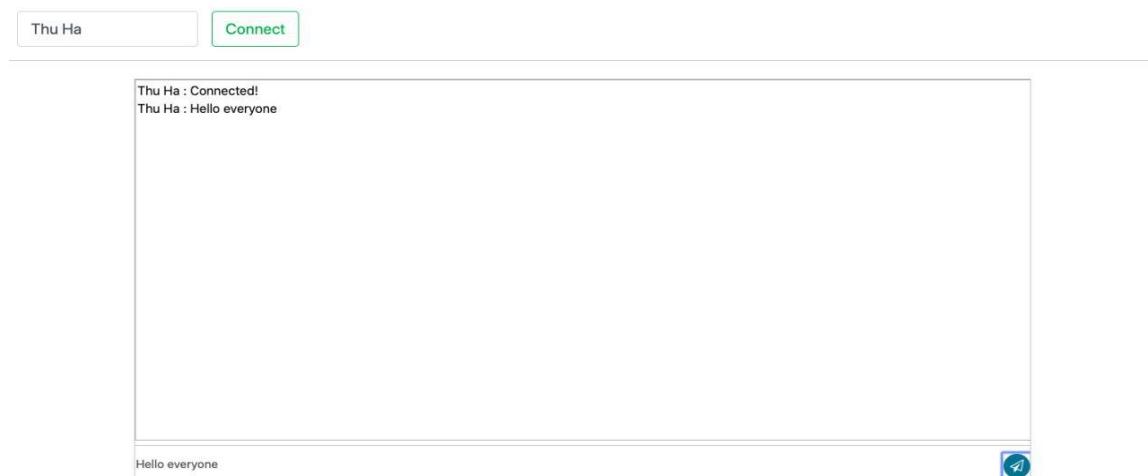


Figure 6: Chatting

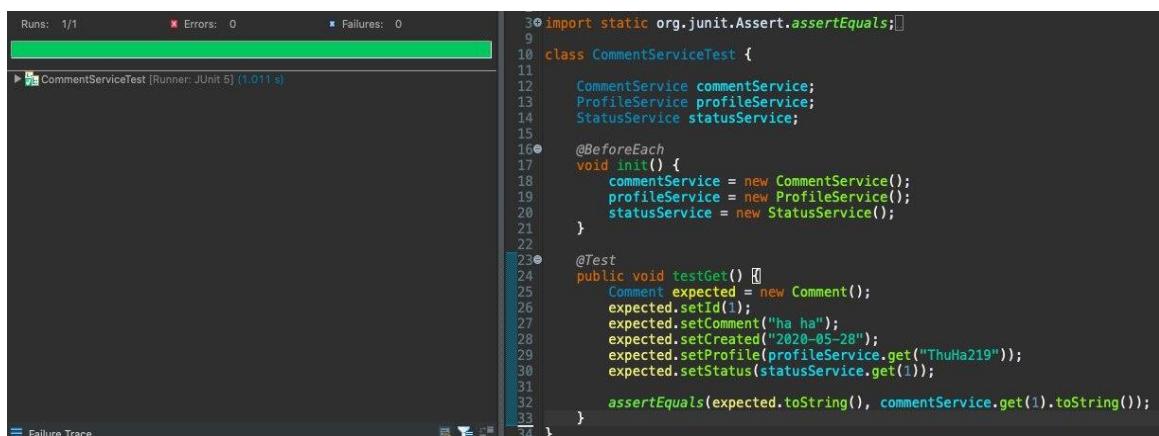
D. Testing

1. Methods

To test software, we use Junit which is a Regression Testing Framework used by developers to implement unit testing in Java, and accelerate programming speed and increase the quality of code**Invalid source specified.** JUnit Framework can be easily integrated with Eclipse.

2. Results

The below figure 7 shows our Junit Test demo for Comment Service with testGet method to check selecting comment from database, which assists in increasing the quality of code.



```
Runs: 1/1      Errors: 0      Failures: 0
CommentServiceTest [Runner: JUnit 6] (1.01 s)

import static org.junit.Assert.assertEquals;
class CommentServiceTest {
    CommentService commentService;
    ProfileService profileService;
    StatusService statusService;
    @BeforeEach
    void init() {
        commentService = new CommentService();
        profileService = new ProfileService();
        statusService = new StatusService();
    }
    @Test
    public void testGet() {
        Comment expected = new Comment();
        expected.setId(1);
        expected.setComment("ha ha");
        expected.setCreated("2020-05-28");
        expected.setProfile(profileService.get("ThuHa219"));
        expected.setStatus(statusService.get(1));
        assertEquals(expected.toString(), commentService.get(1).toString());
    }
}
```

Figure 7: Comment Service Junit Test

III. DISCUSSION

This study has proposed the simple and user-friendly social network platform which helps users can communicate and interact with others more conveniently. Moreover, our project applies REST Architecture Style to build API easily. However, our project is still limited, social media web is quite simple and has limitation in functions. The potential of application of social networking platforms is that they assist not only in communicating and interacting better but also in growing marketing. Increasing numbers of businesses in various industries have already integrated or plan to integrate social media applications into their marketing programs. Besides, we will develope more about maintance pharse by updating the code according to the changes taking place in user end environment or technology. This topic is of interest for further research and implementation.

IV. CONCLUSION

In short, while studying users' usage of social media, we concluded that social network is a double - edged sword, therefore, in order to make the benefits outweigh the dissavantages,

creating a convenient and friendly social media web for users are really essential. This paper has discussed about the development approach - Software Development Life Cycle (SDLC) which helps us to design, develop and test high-quality softwares, that meets or exceeds customer expectations. It is hoped that in the future, there would be many more further projects about social media so as to help users to have better social network platforms.

APPLYING MULTIMEDIA TO SUPPORT ACTIVE LEARNING AT HANOI UNIVERSITY

SVTH: Nguyễn Thị Diễm, Lê Đức Lâm
Nguyễn Dương Bích Quỳnh, Phạm Thị Quỳnh
GVHD: ThS Đinh Thị Minh Nguyệt

Abstract - Research paper on the application of multimedia to positive learning methods. Active learning is an activity that allows students to participate in the learning processes in the lecture so that teachers and students can identify mastery of knowledge and adjust the guidelines to facilitate learning. Long-term training. Multimedia is a combination of information of various forms and that information can be shown simultaneously. We have searched documents and gathered information to have a survey at HANU about learning by actively applying multimedia methods. The results showed that the active learning method has certain positive points such as students have a more flexible way to learn and improve some skills such as debate, information search, presentation, teamwork. , ... Moreover, with the application of multimedia in active learning, students are more interested in self-study. So a lot of software has been designed to help with learning and they have been effectively exploited. Besides, we also noticed some negative results. Learning tools do not fully meet the needs of students. Most students who use multimedia are neglected in their studies.

Key words - **Multimedia, active learning, PBL**

I. INTRODUCTION

While lecturing is an important aspect of university instruction, it is not necessarily the only or best way of engaging students in the ideas and information we are presenting. Lecturing induces passivity of thought, even in the best of students. This method of teaching is being seen as teacher-centric wherein the role of the student is reduced to mere listeners. Active learning is often contrasted to the traditional lecture where students passively receive information from the instructor. Therefore, for students to be more active in learning, schools began to innovate teaching methods. Multimedia is heavily applied in active learning methods. In this paper, there is an introduction to active learning, multimedia, and a comparison between traditional learning methods and positive learning methods. There are many difficulties students also have some challenges in active learning with multimedia.

II. LITERATURE REVIEW

A. ACTIVE LEARNING

1. *Definition*

Active learning is an important method for students to acquire knowledge. This process is student-centered and a pedagogic approach that encourages students to actively obtain knowledge. It encourages students to do on their own understand the knowledge that teachers

provide instead of sitting and listening and writing in a notebook. Active learning is derived from an individual or small group activities which relate to the learning processes done by all students in class and by taking turns with intervals led by an instructor, in which students' responses are processed and new information is collected as the learning result is being presented (Felder & Brent, 2009). The advantages of active learning for students are: being more interested in learning materials, improving intrinsic motivation, improving their comprehension due to a decrease in resistance to learning materials, developing their willingness to be lifelong learners, improving communication skills, improving interpersonal relationships, improving problem-solving skills, improving critical analysis skills and higher thinking (Brickner & Etter, 2008). When joining a passive learning class, students are quite lazy to think and lack focus because the boring teaching methods did not inspire them. However, by applying active learning methods, students will explore and understand more deeply. Students who actively answer teachers' questions, offer options to improve lessons or have many breakthrough ideas when group discussion are active learners. It is not necessary to be as active as extracurricular activities, but it must be dynamic enough to promote thinking, self-study ability, and problem-solving. Active learning does not mean to be completely isolated from teachers, students rely on modern teaching activities in the classroom such as the use of powerpoint, presentations, games, group study, debate, ... to find a lesson for themselves with the guide from their teacher. During the self-study process, students can discuss the problems they encounter with their teachers for help. Students may not need a notebook or pen but are required to have a spirit of self-study, understanding and practicing what is taught in class.

2. Compare active learning with traditional.

In their paper, Raja and Najmonnisa (2018) were conducted to analyze strategy teaching between the traditional method and experiential method for the study of business courses. To clarify the definition of the experiential method, it is said that students are more creative in studying with the new method which is similar to active learning. The mean score of the test among 60 students has been clearly shown that the effectiveness of the experiential strategy. In conclusion, they found that not only teachers but also students are excited with advanced techniques that have a greater impact on improving student communication skills than the old educational methods. The academic community has extensively explored the benefits of fancy educational methods. Following Powerful Schools (2018), in traditional classes, students passively listen to the teacher and they only focus on reading and writing what the teacher says. On the opposite, in an active learning class, the students are allowed to join in different activities and use various techniques to discover knowledge. They have chances to require portions within the learning procedure effectively. Until now this methodology has been applied wisely in other papers. In 2019, Del Campo et al has extended the tremendous advantages and disadvantages between traditional education and modern education. The teaching methods have been gradually changed from the way of using blackboard, and overhead projectors in the old-fashioned classes to video projects, and the Internet websites

presently. In the literature, it shows to be strongly supported for modern techniques. Because of the comfort, while using cyberspace in teaching, teachers can provide detailed lectures as well as show the recently updated knowledge. Besides, students remember the lesson quickly through active activities. The evidence from these studies points towards that active learning is superior to conventional instructive strategies.

B. WHAT IS MULTIMEDIA?

Multimedia is a common term in the world. People could understand multimedia with many definitions in each situation. Almost, multimedia could be combined and presented text, audio, and video conveyed to you by computer or other electronic devices (Vaughan, 2014). In addition, another author-written has some different definitions as Multimedia is the computer's environment. It means the computer environment is a place to combine every element of media. Therefore, this created a great diverse work environment. Another definition is multimedia did not appear before the digital age. Overall, multimedia can be understood as combining two (or more) independent media in a platform environment. The word combination is imperative since interactive media can make a coordinate shape where no one person medium has supremacy (Lachs, 2013). However, with Active learning, people can understand the definition of multimedia which could possibly be combined text, audio, and video or it means computer-based hardware and software packages have developed production on a large scale and owned by individuals for learning (Hilal, Ehab, and Amjad, 2015). With this definition, people can understand that in order for end-users to use multimedia, they need to have personal devices (such as phones, laptop, PC, ...) so that they can get aggregated information from the developers by combining elements (text, images, video, ...) to transfer knowledge to learners.

III. MULTIMEDIA FOR ACTIVE LEARNING

A. TOOLS AND ADVANTAGES

According to Qingsong (2012), the use of multimedia technology can be effective such as the student's interest in learning, giving students direct feedback during the teaching process. Besides, cultivating students' creative thinking is also promoted in teaching multimedia networks. Moreover, the ideal learning environment is built for students to self-study with communication technology. Gunawardhana (2019) conducted that teachers use multimedia applications to present a specific topic which helps encourage students to study particularly effective subjects. These presentations combine a number of different media elements that help students understand the most relevant part of that subject. Multimedia applications also help teachers prepare a variety of teaching materials in less time than the past. In another study, Sauce (2020) found that there are many different tools like using storytelling graphics such as audio. Audio is an effective tool that records what students learn and students can interact directly with their teachers on the course. The display of An Android tablet application that helps blind students learn math through touch and audio, such as video. Moreover, teachers display video applications in the classroom such as animation, documents,

short films, etc to help students more inspired to learn. Teachers create various activities that help students not only interact with teachers but also with students in the classroom. From there, students develop soft skills such as communication skills, teamwork skills, and presentations. From that, teachers know the level of understanding of student papers. Alternatively, teachers insert a few videos that require students to remember and then answer the question at the end of the video which helps students be more interested in the lecture instead of just listening and writing.

B. CHALLENGES

Thanks to advanced tools, student learning is improved. Students can connect, share knowledge through these tools, or they also contribute to the students' self-study ability. However, to access modern tools, students encounter many obstacles. To have the most advanced equipment for the school, the school will have to pay a very expensive fee, not to mention that students have difficulty using them. When it comes to tools such as phones, personal computers, tablets, e-books, ... we will immediately think of their convenience in active learning. However, they are of high value, and not every student is economically viable to have them. Today, people create a lot of software to promote students' ability to self-study. There are very interesting apps, websites have images, colors, sounds ... However, they are almost not free. Students need to pay a sizable fee to use internal functions. Students in traditional classes are often very passive and do not have critical thinking. Therefore, when going to college, it is difficult for them to have an active learning method. They must challenge themselves with self-study and accumulate knowledge. Students who learn the traditional approach will be very timid, and when they switch to the active learning method they will be more likely to be shocked, unable to keep up with their progress. During group activities, discussions, assignments, etc., they will often be ashamed, not giving personal opinions.

C. ADVANTAGES AND CHALLENGES IN HANOI UNIVERSITY

A survey on Hanoi University students' study habits was conducted to see how active learning methods were applied along with the effects of those methods if any..

According to the result of the survey, 55% agree and fully agree that the new method of learning understands the lesson better than the traditional method, however, 40% of students suppose that there is no difference between two methods, and only 5% think traditional one is better. Most students (77,5%) show that the iteration between teacher and student through online software as well as social networks are quietly effective. By contrast, 2.5% of students in total found the interaction ineffective. Besides, information searching skill has improved precisely (70%) which account for the mass of the other skills. The majority of learners with 65% is easy to access to new learning methods. The level of absorbing lessons at 65% is much higher than the traditional one. The result of learning demonstrates agreeing and strongly agreeing accounted for 52,5% which is the largest number of surveys. Meanwhile, 25% of students learn by the social network, interaction channel not effect

Besides the advantages, this survey has some challenges. Most students are sufficiently equipped with learning techniques and tools but they confirmed that all the provided tools have not met the needs (45%). There is a small number of students who report that they have no learning tools (5%). Most of the learning tools are new and modern, so most students have difficulty accessing and understanding new tools (65%). From traditional learning methods to active learning methods, some teachers do not apply appropriate teaching methods (37.5%), so the students' learning results have not improved markedly. Up to 65% of students feel that there is no difference in comprehension between traditional and active learning methods. Skills such as critical, problem-solving, teamwork have not improved significantly (47.5%). With this method of learning, multimedia applications are widely used, so most students are distracted during learning (87.5%) by social networking sites (80%), online games. (35%), the lesson is not attractive and the difficulty in finding information is daunting (27.5%).

IV. CONCLUSION

According to the research, giving an overview of active learning by multimedia. The paper shows many advantages and challenges of active learning in Hanoi University. Most of students could approach to this method and bring a great result, however, there are still many difficulties that need to be overcome. Without the trend of increasingly active in learning of Hanoi University students. The methods will further develop its potential and bring about more positive effects. Therefore, teachers and students need to continue improve and investigate to suitable with Hanoi University.

REFERENCES

- [1] Bonwell, C.C. and Eison, J.A. (n.d.). Active Learning. [online] Available at: <https://www.everettcc.edu/files/administration/institutional-effectiveness/institutional-research/outcomeassess-active-learning.pdf>
- [2] Brickner, D. R., & Etter, E. R. (2008). Strategies for promoting active learning in principles of accounting course. Academy of Educational Leadership Journal, 12(2), 87-93. [online] Available at: <http://www.alliedacademies.org/articles/aejvol12no22008.pdf>
- [3] Cambridge-community.org.uk. (2019). Getting started with Active Learning. [online] Available at: <https://www.cambridge-community.org.uk/professional-development/gswal/index.html>
- [4] Del Campo, J., Negro, V. and Núñez, M. (n.d.). TRADITIONAL EDUCATION VS MODERN EDUCATION. WHAT IS THE IMPACT OF TEACHING TECHNIQUES' EVOLUTION ON STUDENTS' LEARNING PROCESS? [online] Available at: http://oa.upm.es/21062/1/INVE_MEM_2012_130820.pdf [Accessed 14 Nov. 2019].
- [5] Farhan, U., Raja, S., Zulfiqar, A., Bhutto and Najmonnisa, S. (2018). Article Journal of Education and Educational Development Comparing Traditional Teaching Method and Experiential Teaching Method using Experimental Research. [online] 5(2), p.276. Available at: <https://files.eric.ed.gov/fulltext/EJ1200262.pdf>.

[6] Felder, R. M., & Brent, R. (2009). Active learning an introduction. ASQ Higher Education Brief, 2(4), 4-9. [online] Available at:

<http://asq.org/edu/2009/08/best-practices/active-learning-an-introduction.%20felder.pdf>

[7]forbes (2019). Active Learning Challenges Old Education Models. [online] Harvard Extension School. Available at: <https://www.extension.harvard.edu/inside-extension/active-learning-challenges-old-education-models>

[8] Gunawardhana, P., 2019. *Possibility Of Using Multimedia Application For Learning*. [ebook] Available at: <https://www.researchgate.net/publication/308579032_Possibility_of_using_Multimedia_Application_for_Learning> [Accessed 12 March 2019].

[9] Hilal, A., Ehab, F., and Amjad, S., 2015. The Effectiveness Of Multimedia Learning Tools In Education. [ebook] p.1. Available at: <https://www.researchgate.net/publication/272433385_Active_Learning_versus_Traditional_Teaching> [Accessed 12 May 2015].

[10] Lachs, V., 2013. *Making Multimedia In The Classroom*. Hoboken: Taylor and Francis.

[11]MULTIMEDIA IN THE CORE MULTIMEDIA ACROSS THE COLLEGE PROGRAMS. [ebook] School of Cinematic Arts University of Southern California, p.4. Available at: <<https://cinema.usc.edu/images/iml/iml mmc studenthandbook.pdf>>.

[12] Powerful Schools. (2018). Why Is Active Learning Better Than Other Old Educational Methods? [online] Available at:

<http://www.powerfulschools.org/why-is-active-learning-better-than-other-old-educational-methods/> [Accessed 12 Apr. 2020].

[13] Qingsong, L. (2012). The Application of Multimedia Technology in Web Education. 1st ed. [pdf] China: Nanyang Normal University. Available at: <https://core.ac.uk/download/pdf/82176089.pdf> [Accessed 25 May 2012].

[14]Sauce, E., 2020. Application Of Multimedia In Education | Education Essays | Essay

S sauce Free Student Essay Examples. [online] ESSAY SAUCE. Available at:

<https://www.essay sauce.com/education-essays/application-of-multimedia-in-education/>

[15]Spaces for Learning. (2017). Common challenges when teaching in an Active Learning Classroom – and how to address them. [online] Available at: https://blogs.shu.ac.uk/learningspaces/enhancing-practice/developing-your-practice/teaching-in-active-learning-classrooms/common-challenges-when-teaching-in-an-active-learning-classroom-and-how-to-address-them/?doing_wp_cron=1587816977.1500399112701416015625

[16] Vaughan, T., 2014. *Multimedia*. 8th ed. New York: McGraw-Hill Education, p.15.

APPENDIX

Khảo sát thói quen học tập của sinh viên Hanu

1. Mô tả

● “Active learning”-”học chủ động” là một phương pháp học tập mới, các lớp học sẽ trở nên năng động sáng tạo hơn nhờ áp dụng những phương pháp và công cụ giúp ích cho quá trình giảng dạy của giáo viên. Đồng thời học sinh trong các lớp học này sẽ trở nên sáng tạo hơn, chủ động tìm hiểu và tiếp thu kiến thức và ghi nhớ lâu. Bằng những công cụ ví dụ: máy tính, điện thoại thông minh, trang web(google meeting), quizlet,

● Đa phương tiện có thể hiểu theo nhiều định nghĩa khác nhau. Trong bài nghiên cứu, đa phương tiện là kết hợp thông tin nhiều dạng khác nhau, có nội dung mà sử dụng kết hợp giữa video hay chữ, tranh ảnh, âm thanh

● Với phương pháp truyền thống sẽ luôn là “cô đọc trò chép” sự tương tác giữa giáo viên và học sinh ở mức thấp, trái ngược hoàn toàn với phương pháp mới “học chủ động”. Học sinh được tương tác nhiều với thầy cô và bạn cùng lớp thông qua những hoạt động trong lớp học. Từ đó hình thành những kỹ năng mới như: thuyết trình, kỹ năng giải quyết vấn đề, tư duy phản biện,vv..

2. Câu hỏi được dựa trên những nghiên cứu

a. Những khó khăn học sinh gặp phải như áp dụng phương pháp mới

1. Bạn có được trang bị đầy đủ công cụ ở trường học cũng như ở nhà để đáp ứng quá trình học tập phương pháp mới? (đầy đủ/có nhưng chưa đáp ứng nhu cầu/không có)

2. Trong lớp học áp dụng phương pháp mới,bạn có thấy hiểu bài hơn so với phương pháp truyền thống? (hoàn toàn đồng ý/đồng ý/trung lập/không đồng ý/hoàn toàn không đồng ý)

3. Giáo viên có hỗ trợ học tập qua các kênh tương tác để đảm bảo việc học tập thông qua gmail, facebook, zalo,..?(hoàn toàn đồng ý/đồng ý/trung lập/không đồng ý/hoàn toàn không đồng ý)

4. Các kỹ năng bạn rèn luyện được khi áp dụng phương pháp học tập mới đó là?(kỹ năng giải quyết vấn đề, kỹ năng thuyết trình, kỹ năng làm việc nhóm, kỹ năng tìm kiếm thông tin, khác)

b. Trở ngại về phương pháp sử dụng công nghệ mới

1. Bạn có gặp phải khó khăn khi tiếp cận với cách thức học mới không?

A. Có

B. Không

2. Điều gì khiến bạn gặp khó khăn khi tiếp cận với phương pháp học chủ động?(nhiều hơn 1 đáp án)

- A. Trang thiết bị đồ dùng học tập không đủ điều kiện
- B. Cách thức giảng dạy chưa thực sự phù hợp, khó tiếp thu
- C. Bản thân chưa có ý thức học tập và tìm hiểu thông tin
- D. Gặp khó khăn trong việc tiếp cận và hiểu rõ những công cụ mới
- E. Khác

3. Trong khi bạn sử dụng đa phương tiện để tự học bạn có thường hay bị sao nhãng không? Đó là gì?

- A. Có
- B. Không

4. Điều gì gây ra sao nhãng cho bạn

- A. Game
- B. Các trang mạng xã hội
- C. Bài học không hấp dẫn
- D. Khó khăn khi tìm thông tin, gây nản chí
- E. Khác

c. Các công cụ

1. Bạn thường sử dụng những gì để tự học?(nhiều hơn 1 đáp án)

- A. Video
- B. Audio

C. E-book

D. Website

E. Khác

2. Các phần mềm có dễ dàng sử dụng? (Dễ sử dụng/ Trung lập/ Khó sử dụng/ Khác)

d. Ảnh hưởng tích cực với việc học

1. Với phương pháp học này thì mức độ tiếp thu bài có dễ dàng cho việc tự học không? (Dễ/ Trung lập/Khó/ Khác)
2. Các kỹ năng: giải quyết vấn đề, làm việc nhóm, phản biện, có được cải thiện đáng kể không? (Có/ Trung lập/ Không/ Khác)
3. Việc áp dụng phương pháp mới có tạo sự hứng thú trong quá trình tự rèn luyện ở nhà cũng như trên lớp không? (Hoàn toàn đồng ý/Đồng ý/Trung lập/Không đồng ý/Hoàn toàn không đồng ý)
4. Kết quả của phương pháp học tập mới có đạt được hiệu quả như mong đợi không? (Hoàn toàn đồng ý/Đồng ý/Trung lập/Không đồng ý/Hoàn toàn không đồng ý)