LAUREA
UNIVERSITY OF APPLIED SCIENCES
Together we are stronger

Enhancing security, information privacy on users registration and login

Kimambo, Elisante

2015 Lepävaara

Laurea University of Applied Sciences
Lepävaara

Enhancing security, information privacy on users registration and login

Elisante Kimambo
Degree Programme in BIT
Bachelor's Thesis
May, 2015

Laurea University of Applied Sciences **Abstract**
Leppävaara
Degree Program in Business Information Technology

Kimambo, Elisante

Enhancing security, information privacy on users' registration and login

Year 2015 Pages 39

The thesis is aiming to enhance the registration and login security and raising privacy and terms awareness on client side of the web application. The technological driven world has completely changed the way we interact one another for instance the use of social media has become such important and integral part of everything we do in life.

This is because in today's world, social media has deep influence on how people do things and acting in certain ways. The issue of whether social media has made a world a better place to live will continue to be a controversial topic. However, as English proverb goes, "every coin has two sides" social media has had few negative effects on our society including user's privacy etc. The thesis has examined various ways of protecting user's information on project X social media.

As we all know in recently years, social media has become such a widely used aspect of the web that families, friends, businesses, communities, institutions and organizations depend on it to engage with their audiences, to promote themselves and to communicate better with each other. Users typically join social media by creating a profile and they rely on other users to build up their own network of contacts. In project X social media allows new contacts to be recommended as one way of helping to grow the user's network.

Users are concerned about their information and privacy on social media. The thesis paper had done a contribution to project X and which includes well organized privacy, policy and terms of use, decreasing the number of automated registration robots and examining various security issues awareness facing today's social media. The paper is divided into five different sections, section one gives an introduction about the project:- section two discussed about review of related literature and studies, section three discussed methodology, section four implementations and section five described the outcomes. Implementations and demonstrations were carried out on a local server XAMPP on Windows 7 machine.

Keywords php, mysql, database, oop, html, css, xss, crsf, policy, conditions, terms

Table of contents

1 Introduction

As a company/business grows it increasingly depend on web to reach out more customers and achieving maximum profit, the demand for security also grow rapidly and become more complex to handle. Most companies equip their Web sites with firewalls, Secure Sockets Layer (SSL), and network and host security, but the majority of attacks are on web themselves - and these technologies cannot prevent them (**Beverly, 2008**).

Most of businesses/companies take the need of website for granted and misplace priorities while giving too little awareness about their web security. A recent survey of security executives from Fortune 1000 companies (fishnetsecurity.com/News-Release/Firewalls-Top-Purchase-Priority-In-2010-Survey-Says-) showed that the number one IT security spending priority was network firewalls. By considering the effort these companies are putting on their firewalls, one may guess that these companies are attacked through open ports on their networks. According to security survey conducted by (fishnetsecurity) Fortune 1000 companies and other organizations get attacked was through their web site. We may ask ourselves how often do websites get attacked? Security industry analysts suggests that as much as 70 percent of attacks come through web application (Bryan, 2012).

1.1 Company Background

Project X is a startup medium sized company that its success is driven by users and their commitments. At the time of writing this paper the company provides only online services to its users. Project X is an integrated social network and e-commerce that connects a communities' diaspora around the world and local communities in a certain country. Project X was founded on July 21, 2012 by a student who was trying to unite students from abroad and local students in a certain country to enhance disseminating of knowledge and share information back and forth.

1.2 Challenges/Problem

Project X is facing several challenges and that including users' registration and login system, robot users are doing registrations and logins. They are getting dozens of spams registrations per hour and post some spam links in their users' profiles and after that they never login in again. User registration or login with invalid email address, project X system cannot detect invalid email address, and the email verification is needed to tackle this problem. Project X wants to prevent two users from login at the same time in different locations. Project X is lacking well and organized written social media privacy policy and terms of use. But also the

most important thing is that, project X is lacking awareness about possible security pitfalls available today on web sites.

1.3 Objectives

The thesis aim is to address these challenges and come up with new ideas or mechanisms to prevent automated generated inputs to the registration and login forms. Verification of email addresses to prevent invalid emails, preventing multiple logins with the same email address and writing well organized privacy policy and terms of use, which all users must accept before joining the community. In particular, this paper is intended to address the common significant web page security pitfalls.

1.4 Limitations

The paper focuses at security related issues on user registration forms, the briefly overview on web security is also discussed as its well-known that the web security is a huge topic which in one way or the other the thesis will not be able to cover everything related to web security.

2 Methodology

In this section the paper discussed different kinds of methods used to conduct the research and gaining good understanding of the actual problem project X is facing. The next couple of sections will discuss how the project was done and why it was done the way it's done and what was achieved out of the research. Various methods were used during the research face and they are explained below.

2.1 Interviews

Interviews were the most flexible way of gaining qualitative information about experiences, views and feelings about the project X. Multiple chats and exchanging ideas with a project owner to come up with a possible stronger achievements were done time to time since the inception of the thesis project. Furthermore, several meetings with company's programmers were also achieved to get good understanding of the problem making use of qualitative research by holding a discussions. The first meeting was dedicated for understanding the problem and data gathering and set of milestone assignments. Several other meeting were to demonstrate the thesis proposal and see if there were needs for corrections while focusing on the problem itself.

2.2 Observations

The thesis wanted to know what users or people do under the certain circumstances to gain the experiences and feelings of project X website. Observation was one of the most straightforward way of getting information. Normally, observation can be in form of quantitative or qualitative in this case registering as one of the project X user and see the number of fake users (robots) are registering is quantitative observation and on the other hand seeing number of legitimate users complaining about the fake robots is qualitative observations.

2.3 Benchmarking

Benchmarking is another important method used to measure and compare at the same time trying to figure out how other big companies such as Facebook, LinkedIn etc. they deal with same kind of problems project X facing. Benchmarking gives important insights by helping to understand how companies like project x look at similar problems. In addition benchmarking helps to identify the areas that need some improvements. Other sources of information such as thesis topic related books and online virtual learning Lydanda.com were used as another way of gaining more insights about the problem in general.

2.4 Constructive research approach

Among all methodologies could be possibly used, the thesis has chosen the constructive research approach. Constructive approach is used to solve problems that are already existing in real world by producing innovative constructions. Why constructive research approach? Its constructive research approach because one of the core features of constructive research approach require that the problem is a real world problem and it's relevant to be solved in practice. Also it a constructive because the developed constructions will be implemented depends on the project owner.

3 Review of related literature and studies

In this section concepts and insights related to the thesis project are explained to give more understanding about the challenges and problems most web sites specifically project X is facing.

3.1 Web security

The information technology industry has a big problem, more than sixty billion dollars is spent by global IT industry every year. That's more than the gross domestic product of two-thirds of

the countries in the world. (Bryan, 2012). Too much money is spent on security but still attacks are carried out by hackers, this is probably because the money is spent on wrong things. The vulnerabilities in web applications has been responsible for some of the most damaging, high-profile breaches in recent news. The research done by this paper shows few attacks were carried out in the first half of 2011 includes:

The SQL injection attacks on the Sony Music web sites in May 2011 by the LulzSec organization. While unconfirmed by Sony, it's also believed that SQL injection vulnerabilities were responsible for the attacks against the Sony PlayStation Network and Qriocity that leaked the private data of 77 million users and lead Sony to shut down the services for over a month. The overall cost of this breach to Sony has been estimated to exceed 171 million dollars (US).
A cross-site scripting vulnerability in the Android Market discovered in March 2011 that allowed attackers to remotely install apps onto users' Android devices without their knowledge or consent.

The attack on information security firm HBGary Federal in February 2011 by the hacker group Anonymous. Another simple SQL injection vulnerability in the hbgaryfederal website, combined with a poorly implemented use of cryptographic hash functions, enabled Anonymous to extract the company officers' usernames and passwords, which the enabled them to read the officers' confidential internal e-mails. To mention just a few of them, these attacks convince us that, there is a good reasons to pay more attention on companies' web security specifically in this case social media. The thesis will look at some of these attacks into more details in next sections.

3.1.1 Cross-Site Scripting (XSS) attack

XSS is a hacking technique used by malicious hackers to explore weak websites, XSS reviled the vulnerabilities of the website codes to allow attackers to send malicious contents from users and collect important information from victims. This is done by injecting JavaScript or html code in any vulnerable web page. Any web site that accepts inputs from user and display those inputs back to them is vulnerable to cross-site scripting attack.

A complex websites today deliver information to users based on specific needs, these website suffered from vulnerabilities such as Cross-site Scripting attacks on their information.
The registration form on project X contains text and html that is generated by web server and interpreted by browsers. Browsers interpret static and dynamic web page differently because browser has full control over static webpage while on dynamic webpage browser do not have full control during interpretation of web pages.

When malicious content or html is introduced on dynamic web page neither webpage nor browser can take protection measures. XSS attack normally allows hackers to fill in HTML, JavaScript or Flash etc. to any vulnerable websites to brainwash users so that they could gather valuable information. The webpage that is being exposed by XSS attack may compromise hidden data, change cookies also making request that will probably be mistaken by valid users.

Normally a hacking tool may be used to formulate data into hyperlinks and distribute it via internet by using browsers to find out vulnerability of the dynamic website, in order for this mission to be possible hacker needs to know JavaScript and html languages to be able to create less suspicious URLs and this normally happens to most dynamic website that passes parameters from client to database. In most websites these happens on registrations, logins and forget password forms.

Cross-site scripting is dangerous not just because it can have such high-impact effects, but also because it's the most pervasive web application vulnerability. You're potentially creating cross-site scripting vulnerabilities whenever you accept inputs from a user and display those inputs back to them and this happens all the time (Brayn, 2012).

Below Figure 1: is a diagram that shows how attacker gets users data through XSS attack, what is done by attackers is to put or embed a malicious codes on vulnerable website, when a normal user browse on exposed web page the codes are downloaded to his or her browser and executed. As the picture shows a hacker or attacker [1] visits the page first and implicated malicious codes therefore when a genuine user who is now a victim [2] visits the site executes malicious codes without his/her knowledge and finally something bad is accomplished on his/her behalf by attacker [4].

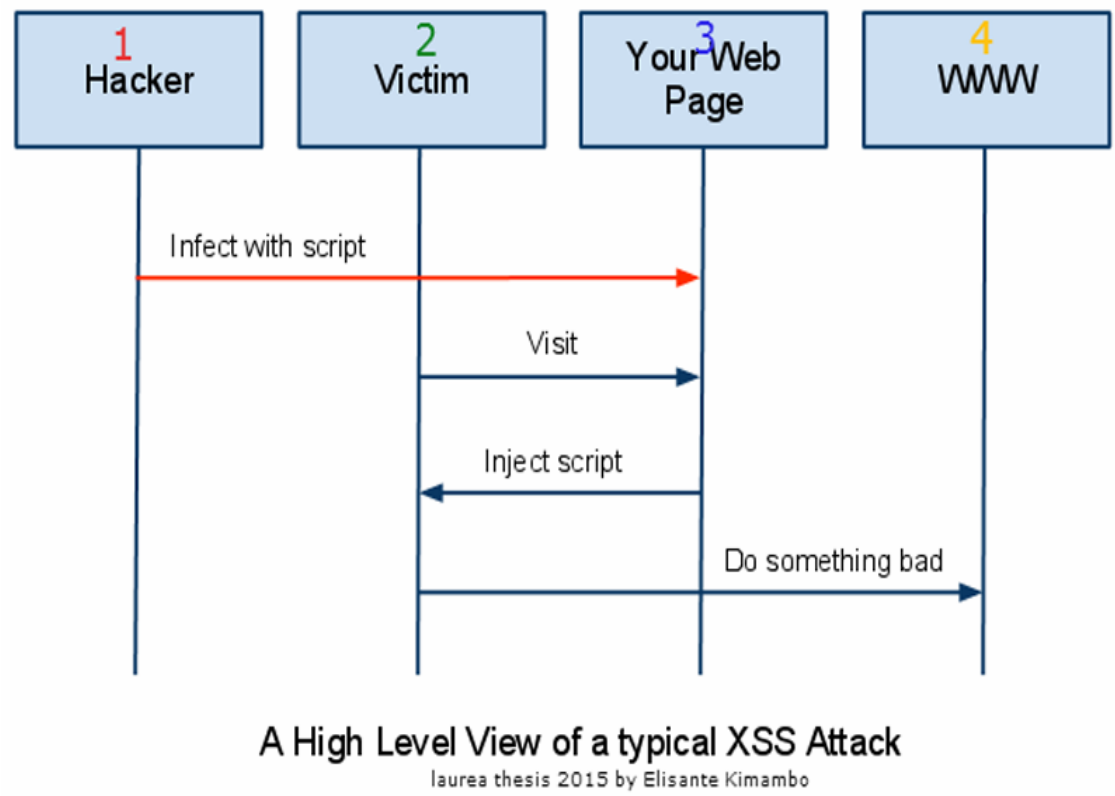

Figure 1: A High Level View of a typical XSS Attack

3.1.2 SQL Injection

The computer language that allows to store, manipulate, retrieve data from the database is called SQL, using SQL language gives the webpages and users the capability to interact with database. The goal of project X is to allow the legitimate visitors to register and login into the site but sometimes not all visitors are legitimate to the website, some have bad intentions. Visitors are allowed to submit and retrieve data to and from a database when they register to the website. Users credentials are stored on a database to be retrieved by legitimate users back and forth whenever is needed.

The same way legitimate visitors can pass data back and forth, attackers may pass something else such as SQL statements expecting that a database will execute them and reveal the data that is stored in a database. This technique is called SQL Injection and is commonly used by attackers to view information on the database and sometimes they many wipe everything out if they want to. SQL injection pose a serious security threat to web application: they allow attackers to obtain a restricted access to the databases underlying the application and to the potential sensitive information these databases contain (William, 2009).

Project X have features like feedback form, contact us form, search from and login form. These forms can be used by both legitimate visitors and attackers to forward data or SQL statements to the database and query it directly; therefore these features are susceptible to attacks because those input fields are meant for legitimate visitors.

The project X has requested this paper to come up with solution to overcome the weakness and disallow any hacker to bypass the login form or any website form barrier and see what is behind it by using SQL Injection attacks.

Based on the research made by this paper intrusion mechanism or firewall and similar mechanisms can do very little about preventing SQL Injection attacks since the website need to be public available for any visitor to forward traffic back and forth through port 80 or through SSL on port 443 and for the purpose of this paper the project X has open access to the database so that the users activities can be updated to the database. There are few SQL commands examples that are commonly used by hackers such as DELETE, DROP, SELECT, INSERT. Delete can be used to delete a particular data from the table within a database, drop can also be used to remove or delete the whole table or a database, select can be used to investigate or lookup for specific data if hackers are interested on what's inside the tables and insert can be used to add more data on the tables.

When a user submit SQL statements on form hoping for problematic SQL to be executed to damage database or reviling of sensitive information, the SQL statements can look like for example: - SELECT * FROM users WHERE email='$email' AND pass=SHA1 ('$pass') so if this was applied on as email input and apply statement like this for example '; DROP TABLE users; as a password if there no precautions to prevent this, the result could look like
SELECT * FROM users WHERE email='whatever@laurea.fi' AND pass=SHA1 ('; DROP TABLE users ;')

There are several ways available of stopping these attacks from being productive. Data should be validated for example making sure the email address has exact correct format and some values should be positive integer. Another possible way to prevent these attacks is by using database special escaping function that validate all strings mysqli_real_escape_string (): $pass = mysqli_real_escape_string ($dbc, $_POST ['pass']); another way is to typecast all values especially the ones that need to be a numeric to be forced to a number for example $id = (int) $_GET ['id']; this time if user submit a password like '; DROP TABLE users; the statement will be type-casted to integer and return zero which by all means it will not harm the database or reviling the sensitive information. Preparing statement could be also an alternative by separating specific values from the statement query and combine them in the database.

3.1.3 Cross-Site Request Forgery (CSRF)

This attack is used by an attacker by executing unauthorized commands from an authorized users or members of the website. The attack is made in such a way that the website trusting a user as if they were previously been authenticated. Previously we discussed about XSS that attacker exploits a trust that user has on website and now we are looking at attacker exploiting a trust browser has on user's website. What a cross-site Request Forgery does is to fool a website to fail to distinguish request on whether they are legitimate or not.

Attacker or hacker will trick a user to access a website or some kind of malicious link that will definitely compromise or steal the user identity and then attacker will behave as legitimate registered user and if a user is an administrator then attacker may shut down the whole application. Attacker may lunch purchase or payment on behalf of legitimate user. The worst part about cross-site request forgery is that every site on the Internet that relies on cookies to identify its users and they are millions of these sites is vulnerable to this attack by default (Brayn, 2012).

In most websites browsers are made to include credentials related to any website is using them, this credentials can be IP address of the user, basic authentication, sessions and cookies of the user. Therefore if these credentials are still valid attacker can use CSRF attack to intervene legitimate user as browser will not be able to distinguish between a trusted user and attacker. Below figure 2: shows the ability to define parameters the way GET and POST unwillingly user who happen to be a victim will not be using these parameters without their knowledge.

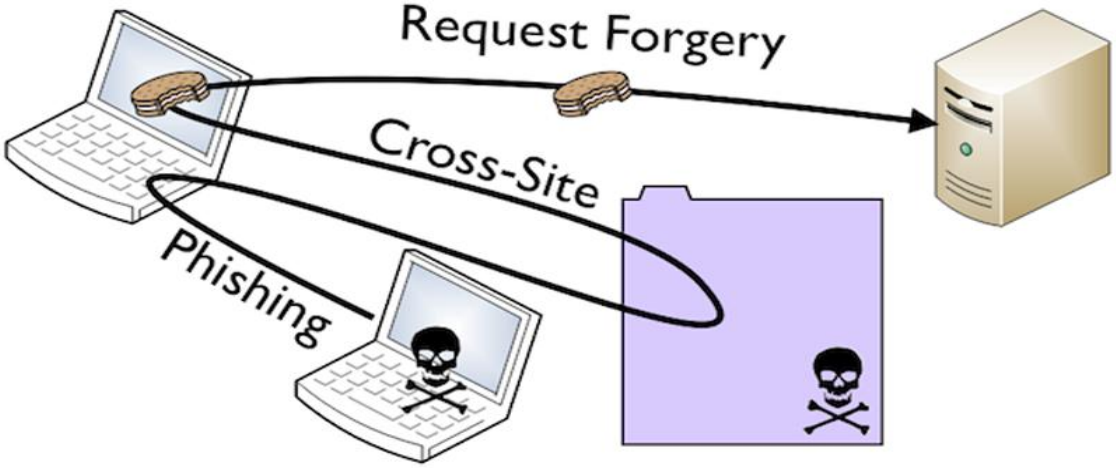

Figure 2: Cross-site request forgery in action

Cross-site request forgery adds a lots of security vulnerabilities that forces end users to run unintentional actions without their knowledge on current application they are authenticated on. When a successful cross-site request forgery (csrf) is applied, it compromises end user's operations and her data. In case if an administrator account is exploited by cross-site-request forgery, the entire application could be compromised. The aim of Cross-site request forgery is to inherit user identity and privileges so that malicious actions can be performed on behalf of user's account.

There several measures can be taken to account in order to overcome CSRF attacks some are not grouped as a complete solution such as use of secret cookies, this method is not considered as perfect way to prevent CSRF attacks because all cookies targeted will be submitted anyway.

We need token class to generate token and check if token is valid, exist and then delete the token, the token will be created for each refresh of page which only that page knows so that another user cannot direct that page because the token would always be checked.
The token need to be generated inside our form on registration page and we need input type hidden, name token and value token class calling generate function. Below Figure 3: is shows how token are generated.

```
    <input type = "password" name = "password" id ="password">
  </div>

  <div class = "field" >
    <label for = "password_again"> Enter your password again </label>
    <input type = "password" name = "password_again" id = "password_again">
  </div>

  <div class = "field">
    <label for = "name">Your name</label>
    <input type = "text" name = "name" value = "<?php echo Input::get('name');?>" id = "name">
    </div>
    <input type = "hidden" name = "token" value = "<?php echo Token::generate();?>">
    <input type = "submit" value = "Register">
</form>
```

Figure 3: A hidden generated token

3.1.4 Sessions

Back in a days, websites only consisted with static content, the HTTP protocol offer authorization which it could be used for authorization and authentication. The authentications were very basic where the user's information's and more specifically username and password where included as the header value. This gave server side application to extract the user's information and making decisions whether to allow the request or not. After gain successful authentication the browser attaches user's information to every subsequent request and since the browser remember these credentials during lifetime of its process to logout is only possible when the browser is closed.

As web grow more complex, web developers wanted to include authentication and application and to authenticate users HTML form where created and user had to enter username and password and by submitting the form the user information were sent to the server where they could be validated. However, since the information is sent in the single request instead of every subsequent request as with authentication header, web needed the way to keep a track on user's authentication state across requests and that was a challenge with stateless HTTP protocol.

The session mechanism tacked the problem, adding session on HTTP protocol was capable of associating multiple requests from the same user with a server side session stated and allow the application to store important session's information

3.1.5 Separate users from robots

Implement a CAPTCHA system (completely automated public Turing test to tell Computers and humans apart)

This is the method of distinguishing whether or not the user is real human or computer during user registration process, there different kinds of CAPTCHA system out there.

Figure 4: Captcha with trick words

No any social media would like to get logged with automated signups that at the end of the day are not going to add up anything to their sites. Captcha is used to tell a computer registered user is a genuine person.

A captcha is the technique that used a series of characters embedded on image and as its known computers are not able to pick text from image automatically, whereas human can do that, and this helps to tell which signup is human and which one is automated computer sign ups.

The CAPTCHA system would be suggested here is going to generate a random phrase and strings then storing these phrases in user's session out of their knowledge, therefore as they move through the registration form to process their registration these images are generated by separate HTTP request and sessions are purposely to maintain session values.

Distorted phrase of strings or numbers will be displayed within an image on registration form and if it's a real user will be able to read the phrase and type it on text box. When user submits the form, the value appropriated session field is compared to those values and if they match registrations proceeds on and if not user is presented with other values to retry again.

Figure 5: ReCAPTCHA with numbers picture

3.1.6 MYSQL Database Design

Simple database is required to save the purpose of the thesis and now we are going to discuss about database and how is being used in our project. First we need to know what data is. In simple words data can be a simple facts related to any object in considerations, for example name, age, weight and height are few examples about data that are related to someone or a person. Things like a file, image, pdf etc. are also considered as a data. Now as we have already looked at what data is it's easier to define database.

Database is a systematic way of collecting data and since the data that is stored in a database is organized it makes data management easy. There are numerous database management systems (DBMS) available on market; the database management system is a collection of programs that work together in order to enable its users to access a database during manipulation of data and representation of data. It also helps to control access to the database to its users.

There good examples on how database is implemented in real life such as on this project. The project needs to store, manipulate and present data related to members, their friends, friend's activities, messages, ads and many more. The usage and expected functionalities of database have been increased immensely and therefore there are different types of database management system (DBMS), hierarchical DBMS, relational DBMS, network DBMS and object oriented DBMS are known types of DBMS.

The relational DBMS is commonly used on market today for the purpose of this project we are going to use relational database. Relational database uses tables to store data; the project is

going to create a database called thesis_project and define three tables; groups table, users table and users_session table.

In our users table this is where our user's data is going to be stored, user uniquely id, username, encrypted password, salted, name, joined date, and group the user is. In our groups table basically we are going to have different groups with different permissions therefore looking at our groups table we have group id, group name and permissions. Lastly we have users_session table which is used as remember me functionality, basically this mean when users' login they can decide whether or not they should be remembered by our login system.

4 Implementations

Previous section the thesis went through few most common web security problems available on today's web applications, the thesis needs to answer the question how to secure the login and registration of project X? The protection is need to be in place to prevent the web security pitfalls;-

CSRF can be used on project X to trick users into performing actions they didn't intend to do, therefore the strategies defined above should be implemented.
CAPTCHA on signup, the thesis recommended CAPTCHA on signup form to reduce automated signups. This is very important because robots signups makes web site slow and annoys real users.

Secure login, the login needs to happen over HTTPS to reduce the risks of the user's information being captured by man in the middle attack.

Email confirmation is required to make sure that all user's email are verified as part of signup process, they shouldn't be allowed to login until their emails are confirmed. This is very important because there times valid email is needed to reset the password.

Secure password reset, a secure method of resetting password is very important to disallow hackers from resetting users' password without their knowledge. The thesis focus about security on signups and logins and the next section shows the html form.

4.1 Client side validation

To save unnecessary server process that may be caused by average users it's highly recommended to validate user's inputs on client-side validation, if they happen to enter invalid email address or password instead of forwarding the information to the server side validation,

the client side validation takes care of that. The next section will describe the simple signup and login system that was used to demonstrate and test the web security.

4.1.1 HTM Design

To save the purpose of the thesis html design will be demonstrated on registration and login forms. The goal at this point is to get users to register by going through authentications process that will be implemented on this paper. The figure 3: shows readymade html design with CSS and JavaScript to validate user's data before sent to server side validation process.

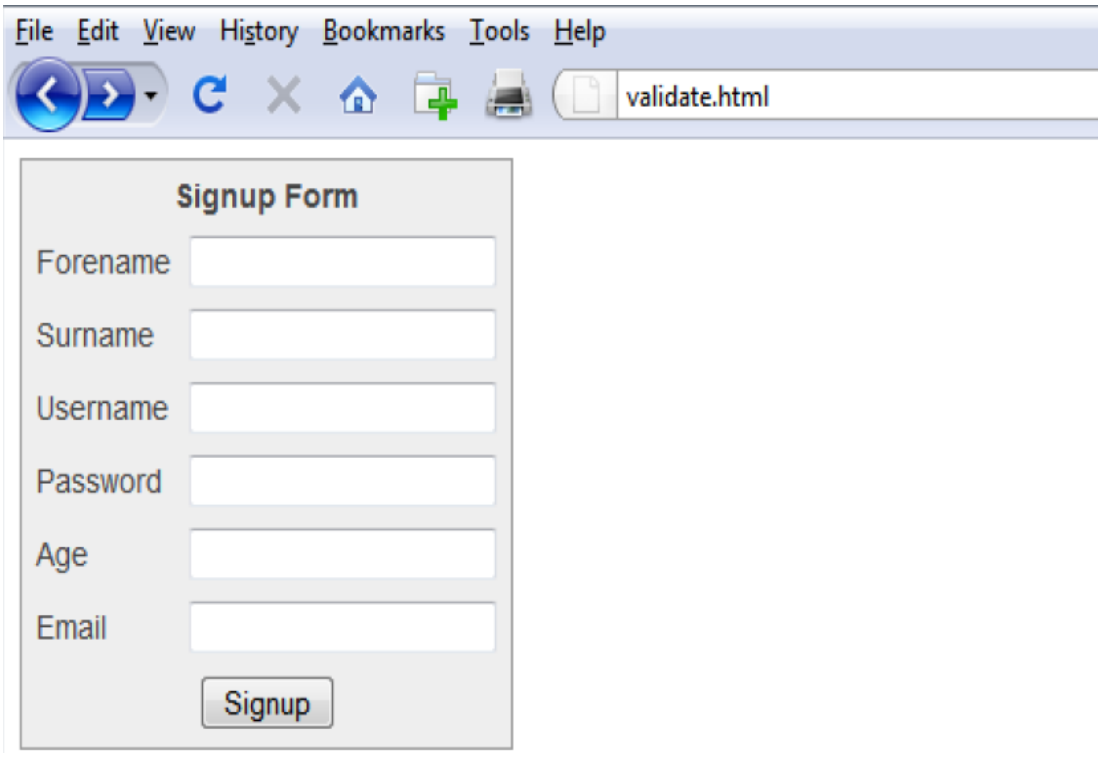

Figure 3: HTML Form Design

Users are presented with the form above during registration process, what users can type into those field have to be validated based on the application requirements. Validations can be done both on the browsers and server side. This paper has suggested both ways of users' data validation.

4.2 Server side validation

Malicious users may bypass JavaScript's client side validation and submit dangerous inputs to the server, so it's very important to revalidate the users' input again with server side validation. Server side validation is all very important because it's possible that not all users have JavaScript enabled on their browsers.

4.3 Protocols and Languages

The protocols and languages used on this project or on web programming are, the main ones are HTTP/HTTPs protocols, the Hypertext Transfer Protocol used on the browser through port 80 to communicate between client-side and server-side. The client-side sends HTTP request and server-side sends back HTTP respond. The HTTPs is for security which is done on the port 443 and that is basic we need to get certificates from third part provider in order to encrypt the transmission between back and forth so that they cannot be sniffed out by web sniffers that sniffers packets. The packets are communicated back and forth between client and server via TCP/IP protocol.

Markup language are used to display the communication that is taking place between client and server, HTML- Hypertext Markup Language it is a main web language used by browsers to have pages displayed. XHTML-Extensible HTML is just an extension of HTML library, and XML-Extensible Markup Language which often used to do things like display data or transmit data. Php and mysql……here…

4.4 System requirements

A simple system to demonstrate the thesis recommendations is going to be developed on local server. XAMPP is used as a local server to host our system. XAMPP is cross-platforms free open source software that includes apache server, SMTP server, ftp server, php and Mysql database. It is very easy to install and with XAMPP you don't need to install all these services individually therefore no configurations are required, install and run.

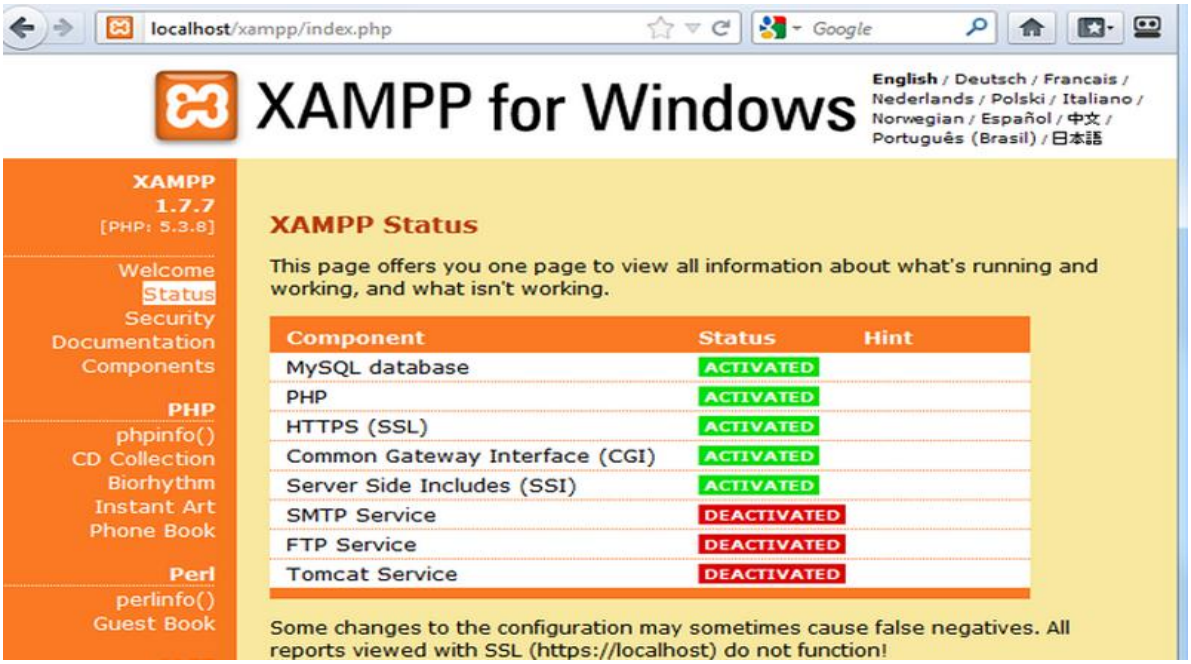

Figure 2: XAMPP server

4.5 Database Structure

4.5.1 Users Table

On our database user table has seven columns, the first column is going to be an Id data type integer (int) and this also going to be auto incremental column and it will be set as primary key so every time user registers or any data inserted in our users table, this will increment. Username column data type is going to be varchar and maximum length of 32, password column is varchar and maximum length of 64 characters and hash, and this means the password is not going to exceed 64 characters length. Salt column is varchar as data type length is 32, name is going to 40 characters, joined column is going to be daytime data type and group is going to be an integer.

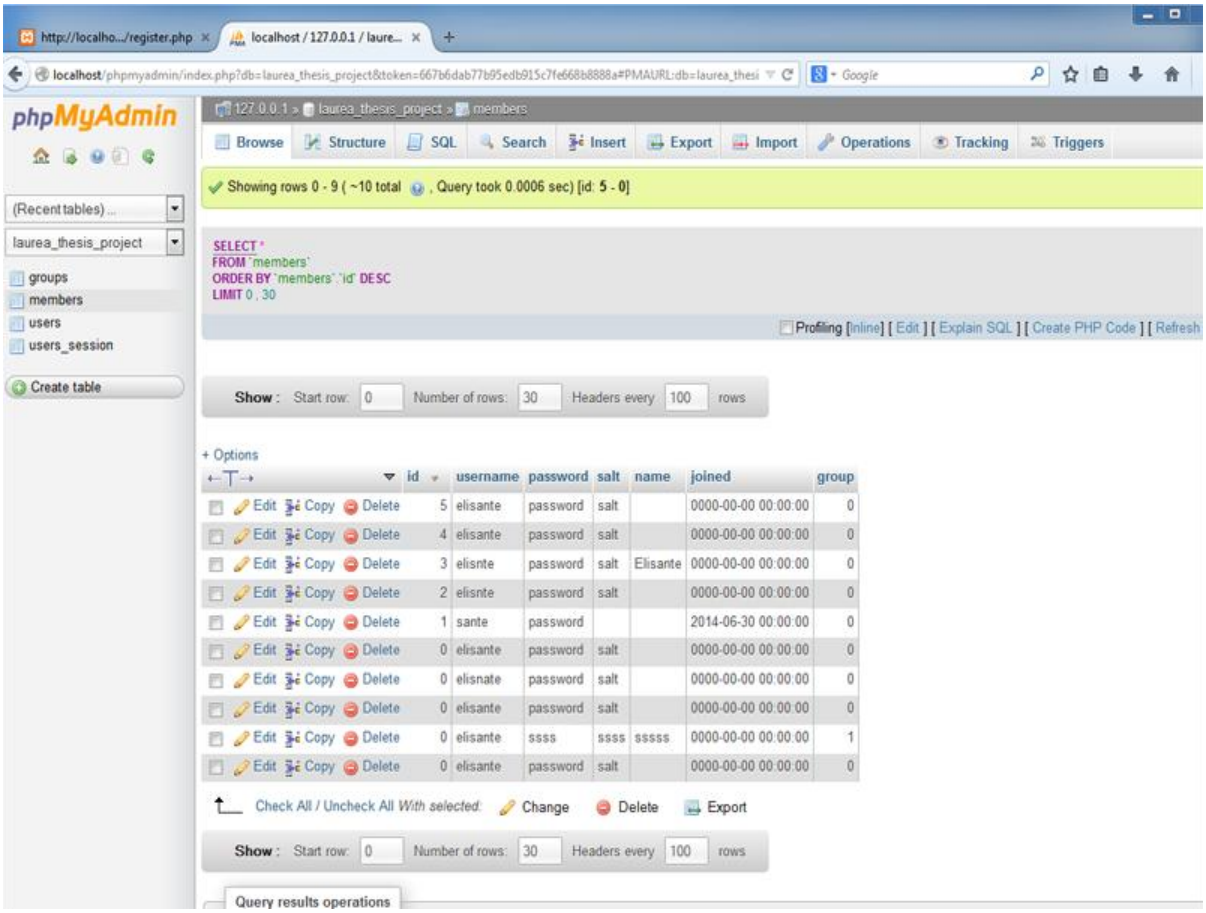

Figure 3: Users' table

4.5.2 Group table

The groups table is going to have 3 columns Id, name and permissions. The Id is a primary key and also auto incremental, the name is going to hold varchar data type and length of 20 characters, the last column permissions is going to hold text varchar data type and this column will hold the different permissions for example administrator will hold 1 and normal users 0.

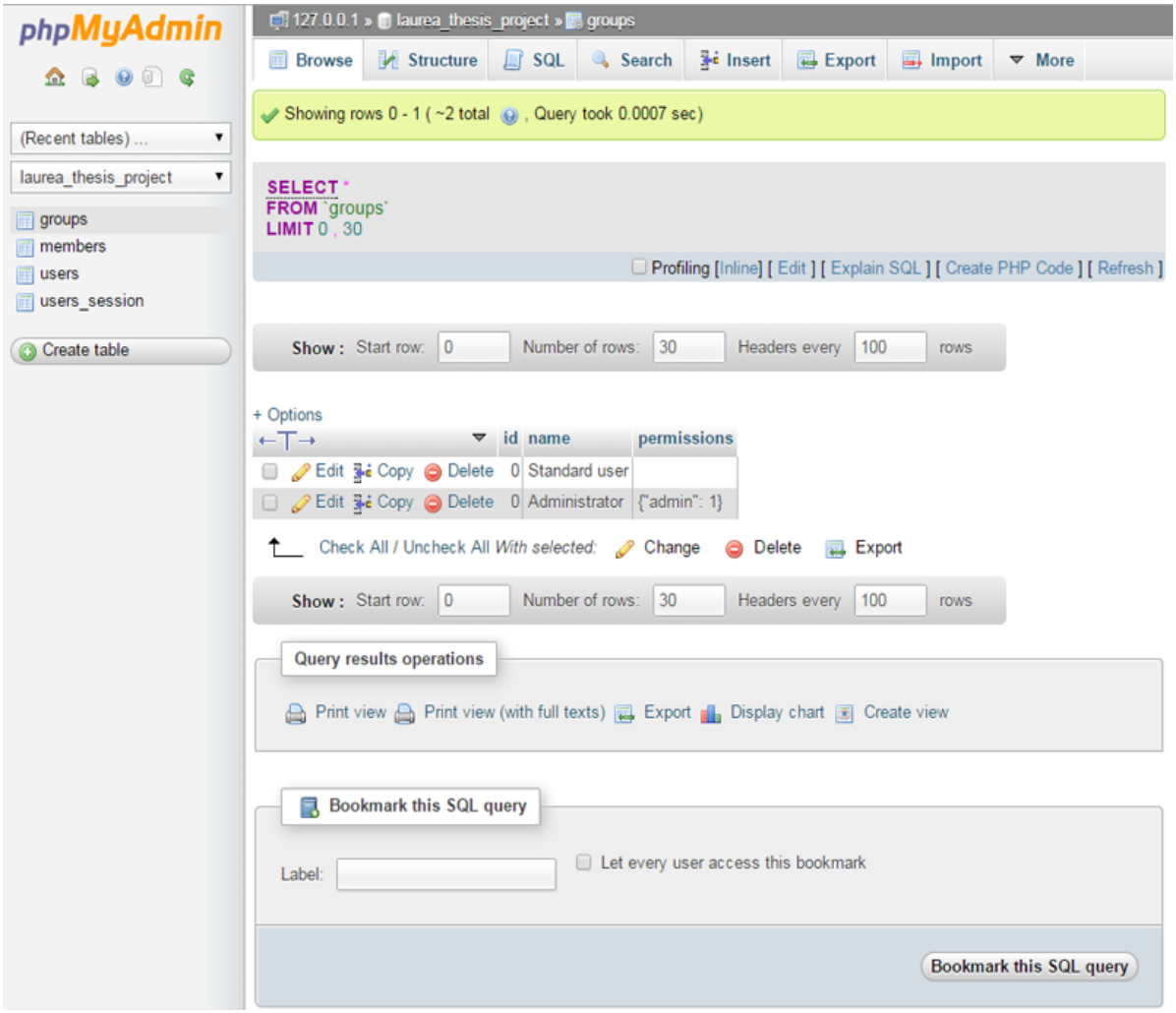

Figure 4: Groups table

4.5.3 Users session table

This table will have three columns Id, user ID, and hash. ID is an integer and is going to be auto incremental, User ID is going to identify users when sessions need to be compared, hash will store the hash.

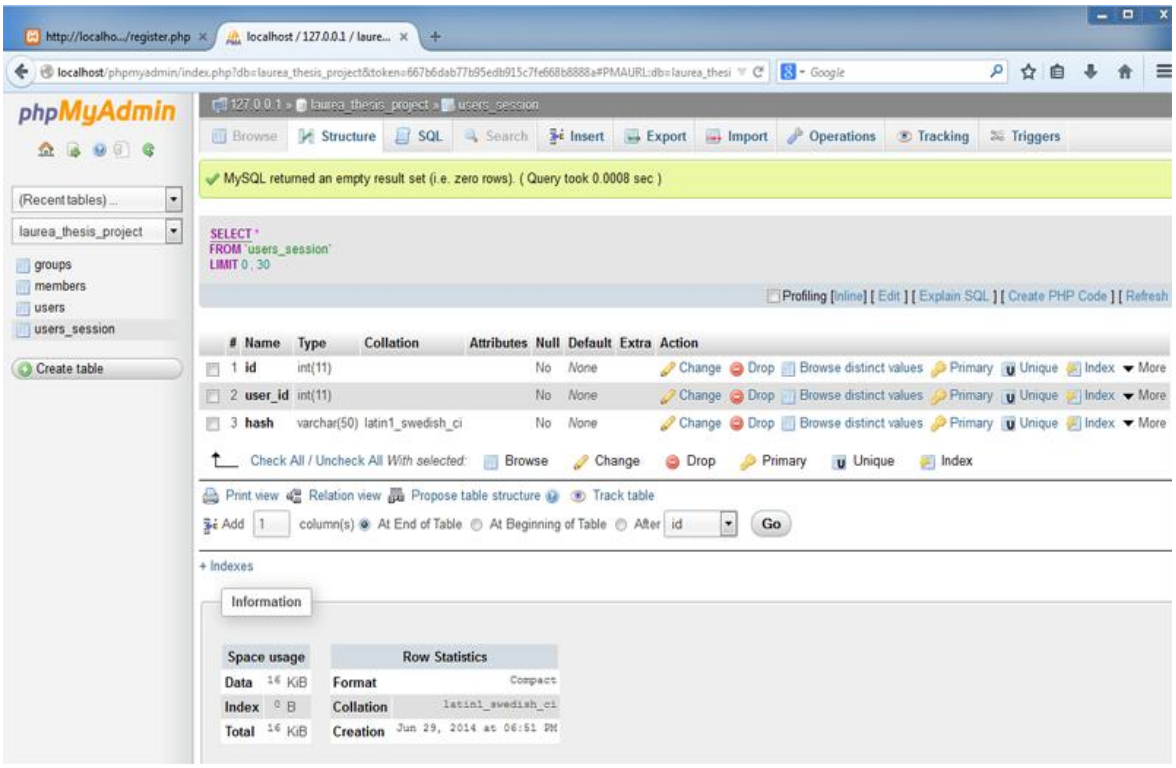

Figure 5: Session table

4.6 Directory Structure and Files

The root directory is going to be called LaureaThesisProject and four folders will be added in it and each folder will hold files, the first folder is going to be classes' folder because we're going to be working with classes, all of our classes will be called from one place which is from this folder called classes.

Another folder is going to be called core folder, and this is going to hold our initialization file such as auto load classes, another folder is going hold our function and actually its only one function call sanitize function and finally we are going to have an includes folder, the folder is going to store things like errors for example if user profile is not found then the error will load up 404 error and this will be automatically displayed to the user.

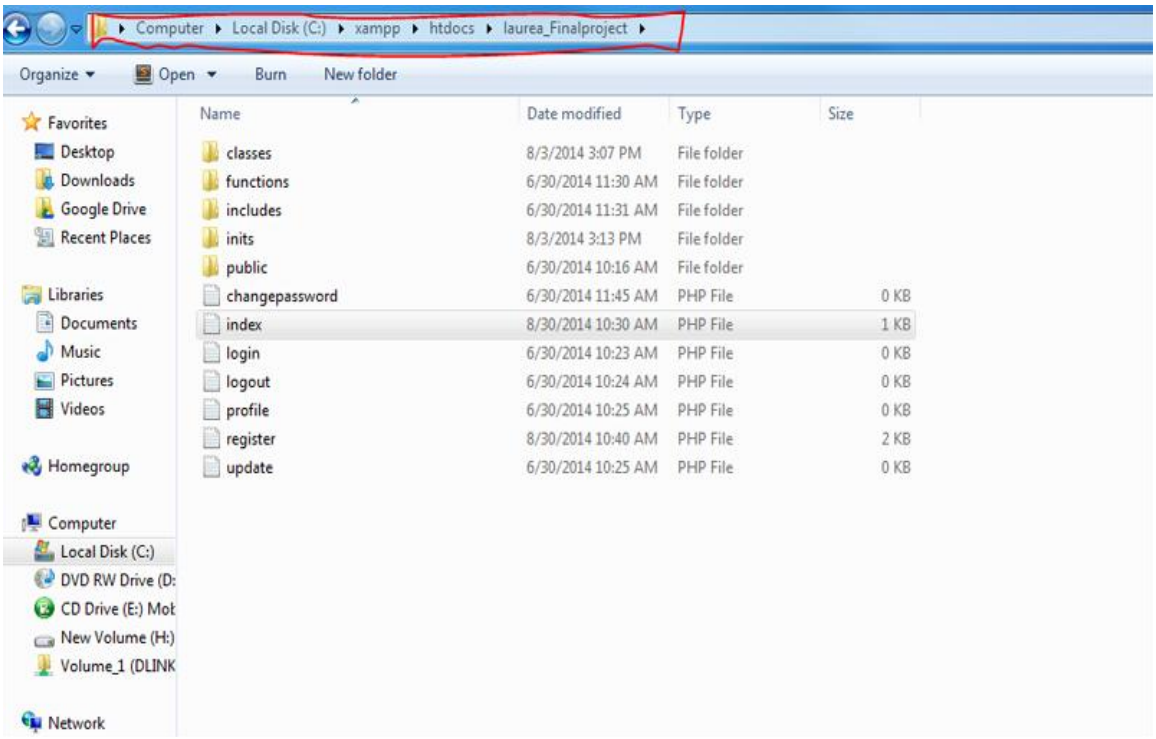

Figure 6: Directory structure

These four folders are going to be stored in our root directory folder also we need to add some file within our root directory folder. Index file will be added because this is the file where user are going land anytime they want to login. The login file is used to process users login and is going to be stored in our root directory folder, logout and view profile files are also going to be stored in root directory folder.

Register, update profile and change password files are needed and will be stored on root directory folder.

4.6.1 Classes Folder

Into this folder we are going to have things like input helpers to help us to work with inputs real easy, validation classes to quickly help us to validate users.

4.6.2 Database Classes

This class will provide basic level of abstractions when the database is accessed, this will also simplify other tasks such updating the database, editing tables, inserting data into database. The class is going to be able to connect more than one database, run some queries and return some results set from executed query. Return id of some records that was last executed.

In order to connect to more than one database connection, maintaining of different connections is needed and by using array it is possible to store different connections. When the query is executed, it will be compared to active connections in our array.

To connect to our database we need three variables; server host, username and password as well as the name of database we may wish to connect to. The connection ID is returned as our results are stored in array. If we want to see the result of query we need to define a variable called $final to store the results and later use the method to access it. When results from query is needed, the method called MysQLifetch_array () is called with results stored on $final variable.

Once we are connected to the database, there are few things we may want to achieve and that includes inserting data to the database, updating the existing data on database and delete data from our database. These operations are often times repeated; however we could abstract them to the low level so that we don't need to repeat ourselves. For example deleting data from our database we can use table, condition and limit to do that. $delete = "DELETE FROM {$table} WHERE {$condition} {$limit}";

Updating that can also be abstracted by passing the table name and arrays of field name and value pairs. $update = "UPDATE". $table. "SET "; same apply to insert $insert = "INSERT INTO $table ({$fields}) VALUES ({$values})";

4.6.3 Data Sanitize

Escaping characters that may be prone to our database it's a great security risk to overcome, sanitize function is used, this make our life easier and, provide single place for changes to be made.

4.6.4 MySQLi Functions

The class has few functions that are helping for example fetching data that are executed by queries, getting number of rows returned by query and getting number of rows that are affected by our query.

4.6.5 Close Connection

The final thing we would like to do in our database class is to close the connection when object is no longer needed. To be able to know what connection should be closed as all are saved on array, we need to use foreach to loop through all connections by using deconstructor.

4.7 Programming

The primary focus of this paper is not to write codes but to enhance security of users' registration and login codes. That said, the programming language used on this paper is PHP which pretty much act like clue to put everything together such as users, browsers, email and database. Since the project deals with users' data the code should not be only functional but also secure, extensible and reusable.

To achieve these goals the code must be well organized and documented each and every step, following this manners the codes will be easier understood by whoever happens to read them. The habit of writing secure codes is crucial when one is dealing with users' data or sensitive information.

4.7.1 Object Oriented or Procedural

In PHP one has a choice to use OOP or Procedural programming approach, based on research made by this paper there is no evidence that tells one approach is better than the other. Different source believe that OOP is more extensible and secure than procedural but the truth is poorly written code in either approach is security flaw. It's not about which approach is used, rather it's all about how secure the codes are.

On this paper the author was much interested of enhancing the skills and knowledge of OOP therefore the thesis paper has use OOP approach to achieve the purpose of the thesis.

The (Hypertext Preprocessor) PHP language was used to send and retrieve data and display the data on the browser, the next section will give more information about how PHP is used to achieve server side validations.

5 Results

On this project we've created a fully functional improved object oriented registration system for project X social media. We created helper classes and some main classes to make it work; all these functionalities can be used elsewhere such as replacement of current login and registration of project X social media.

Throughout the project there have been a learning curve of object oriented programming and how it used in real world projects. The system is very basic in terms of design while trying to focus on how to solve the problem itself.

Users are most important aspect of any site, without users there is no a social media therefore, project X needs to allow user to sing up and login as well as getting information about currently logged in users. Project X needs to manage the permissions of their users to make sure they can do what they are permitted to do and prohibit them from wrong doings.

5.1 Users Data

User’s data are pieces of information that represent users on the site; each user is uniquely identified by user ID, users can uniquely choose their easily membered name such as their username or even their email address, password is also used to prove users they are who they claimed they are. For the registrations and authentication purpose these are few data that users need to provide.

5.2 Users Account

In order for users to be able to login they need to be registered first, the user's account is created by users themselves. Users are given a chance to register then login and logout also whenever they forget their passwords or want to change the password they should be given a chance to do so. This paper has taken into consideration at security issues related to user account during registration and login and figuring out how to enhance them.

5.2.1 Protect Passwords

The issue of how secure the password is depends on so many factors including how and where the password is stored. There are different ways of storing passwords, storing on plan text which is not secure at all, encrypted format which can also be decrypted easily and in hashed format which cannot be decrypted.

5.2.2 Encrypted Format

Storing password in encrypted format may seems secure because after all is not seen as plan text passwords but if someone get to where passwords are stored in this case on the server he can be able to see all password by decrypting them.

5.2.3 Hash Format

Letting someone to discover users’ password is an enormous security violations because some people are lazy and the same password and email address may have been used in several

websites. Hashing the passwords is more secure compared to plan text and decryption mechanism MD5 has been there for years as one of the most used hashing mechanism.

Hashing the word “password” will be represented in MD5 hash as 5f4dcc3b5aa765d61d8327deb882cf99 there is no different words have the same hash values. Storing passwords in hash is believed to be very strong because hacker or attacker will not be able to decrypt it, instead attacker will create the hash of common words hopping to find a match by using attack such as dictionary attack or brute force attack.

5.2.4 Hash Algorithm

MD5 is common used hashing mechanism, but at the same time is not very secure hashing algorithm. SHA or SHA1 are considered more secure and fine for some applications. Hashing password can take place either on the database or by using PHP codes, this paper recommends hashing passwords on PHP codes because PHP seems to have a very sophisticated hashing function out there such as password_hash() compared to MYSQL database function.

5.3 Invalid email address

The login system should be able to detect an invalid email address by confirming whether email ever existed. There is native PHP method to validate email address therefore, there is no need to re-invent the wheel.

```
function validateEmail ($email){
  return filter_var($email, FILTER_VALIDATE_EMAIL);
}

$email = $_POST['email'];
if (validateEmail($email)){
  echo 'Valid email!';
}else{
  echo 'Email NOT valid!';
}
```

5.4 Verify user’s email address

Although CAPTCHA is implemented and we are sure users are humans and not automated robots, however there other number of reasons why we should verify user’s email addresses. By verifying user’s email we have avoid multiple sign up and we are sure records are up to date just for incase users lost a passwords or their email address.

If for some reasons users are misbehaving we have ability to prevent them from repeating sign up unless if they have multiple email addresses. There times when user may receive notifications from other users or from site administrator, therefore if their emails are invalid they won't be able to achieve that.

Sending emails to users is crucial as it might be required to inform them about their details, informing them about users who are trying to connect with them and updates in general for that reason email verification is needed.

5.5 Process users' sign up and login forms

The figure 10: below will accept input data and PHP will process the data entered on inputs form to the database therefore, we need a script that will do the job. The paper has created the script to demonstrate how user's data is validated before sent to the database. The script name is called register.php

```
<form action="register.php" method="post" accept-charset="utf-8">
<?php
create_form_input('first_name', 'text', 'First Name', $reg_errors);
create_form_input('last_name', 'text', 'Last Name', $reg_errors);
create_form_input('username', 'text', 'Desired Username', $reg_erro
echo '<span class="help-block">Only letters and numbers are allowed
create_form_input('email', 'email', 'Email Address', $reg_errors);
create_form_input('pass1', 'password', 'Password', $reg_errors);
echo '<span class="help-block">Must be at least 6 characters long,
one uppercase letter, and one number.</span>';
create_form_input('pass2', 'password', 'Confirm Password', $reg_err
?>
<input type="submit" name="submit_button" value="Next &rarr;" id="s
/>
</form>
```

Figure 7: Registration Form

We need regular expressions to control what users can fill in the forms inputs, starting with a first name the codes below insists that submitted values must be between 2 and 45 characters long and only contains combination of characters letters case insensitive, space, period and apostrophe and the same technique will be used on last name.

```
if (preg_match ('/^[A-Z \'.-]{2,45}$/i', $_POST['first_name'])) {
   $fn = escape_data($_POST['first_name'], $dbc);
} else {
   $reg_errors['first_name'] = 'Please enter your first name!';
}
```

Figure 8: Checking for first names

Checking on username, we may want restrict only letters and numbers by using regular expressions as follow.

```
if (preg_match ('/^[A-Z0-9]{2,45}$/i', $_POST['username'])) {
   $u = escape_data($_POST['username'], $dbc);
} else {
   $reg_errors['username'] = 'Please enter a desired name using only letters and numbers!';
}
```

Figure 9: Checking for usernames

Checking for an email address if is valid format of an email, regular expression will do that for us and this will increase the time for automated robots sign up figuring out the correct format though it will not completely prevent them from sign up.

Finally we check for the password and confirm password by using regular expressions

```
if (preg_match('/^(\w*(?=\w*\d)(?=\w*[a-z])(?=\w*[A-Z])\w*){6,}$/', $_POST['pass1']) ) {
   if ($_POST['pass1'] === $_POST['pass2']) {
      $p = $_POST['pass1'];
   } else {
      $reg_errors['pass2'] = 'Your password did not match the confirmed password!';
   }
} else {
   $reg_errors['pass1'] = 'Please enter a valid password!';
}
```

Figure 10: Checking for passwords

5.6 The users' registration and login overview

Obviously the user will be able to do registrations and when the registration is successful the user should be able to login to our system. While they are logged in user can look at his or her profile based on their username, users should be able change their password and update their profile and finally they should be able to logout.

6 Information Privacy

Project X is free online service that gives users a freedom of speech on top of that it gives them the option to choose from on how they should be been seen publicly online in whatever or whenever format they may like.

6.1.1 Review

The information age has caused some of the rapidest and most alarming to personal privacy than ever before witnessed in the history of mankind, organizations and governments of all kinds are asking for and collecting more private or sensitive information about people than ever before. And on the other hand most people are now sharing or are willing to share more private details about themselves, their lives and their preferences than ever before.

In fact sharing private information has become such a common activity that many people now freely share highly share sensitive information about themselves without ever considering the consequences. The gathering and sharing a private information has become so widespread that the information age despite all its benefits must be looked upon as having been absolute disastrous from perspective of protecting one's personal privacy.

In order to understand the rapid collapse of information privacy since the dawn of the information age we must become familiar with the three major facets of information privacy.

6.1.2 Three major facets of information privacy

Disclosure control refers to the extent which a person can exercise control over the disclosure or sharing up her own private information. The second facet of information privacy is data sensitivity which refers to extent which information is sensitive within a particular context, note that the extent in which information can be considered sensitive is often a matter of perspective.

The third major facet of information privacy is the affected part which refers to the part or parties that would be negatively impacted if private information were to be disclosed and note the part in this sense might be a person, a group, an organization, a government or any combination. The widespread adaption of information and communication technologies or ICT has created several critical and serious threats to information privacy.

6.1.3 Critical and serious threats to information privacy

Data collection, information and communication technologies allow massive amount of potentially private data to be collected share and analyzed, the second critical threat is lack of informed consent potentially privacy information about the action or preferences of specific individuals is routinely collected without the explicit consent of the individual themselves. Example of the type of information that are commonly collected without explicit informed consent include web searches, product preferences, IP addresses, location information etc.

The third major threat related to widespread adoption of ICT is loss of control in information age potentially private or sensitive information is commonly sold or shared between business partners, individual typically have little knowledge of control over such exchanges. Many companies for example commonly share or sell individuals mailing addresses, product preferences, email address and so forth without informing those individuals or obtain their permission.

The final critical threat to information privacy relate to the ownership or private or sensitive data, for example if a company is sharing or selling your private data should you get a share of profits? Further consider that after you have shared your private data the person or an organization with whom you have shared your data with, may not protect it sufficiently well or may divulge your personal data to other parties without your knowledge or consent.

As a means of constraining the abuse of individual's private data several principle should be used to guide the collection of a potentially sensitive information and the first to these is the principle of limited collection, information should only be obtained in a fair and lawful manner.

Second is the principle of data quality, information should be of high quality and relevant to the context for which it is being gathered.

Third is the principle of purpose specification, the purpose for which information will be used should be clearly specified and the information should be permanently destroyed after that purpose has been achieved.

Fourth is the principle of unlimited use, information should be used only for a specified purpose unless the individual explicitly consents to it being otherwise or the law require it to be used otherwise.
Fifth is the principle of safeguarding security, procedures must be established to protect sensitive information from being lost, damaged or misused.

Sixth is the principle of openness, information about the acquisition storage and use of personal information should be easily obtainable.

Seventh is the principle of individual participation, individual should have the right to examine challenge and correct information about themselves their lives or their preferences.

The eighth and final principal is the principal of accountability organization and governments must monitor the extent which they hear to all of these principles and must held accountable for violating these principles. Clearly many of the information collection principle described are commonly overlooked or entirely ignored by modern organizations and governments alike.

6.1.4 Privacy and law

One of the greatest challenges associated with information privacy in the information age relates to the law put simply privacy law vary from country to country. The situation is problematical because in the information age private data can easily cross geographic or political boundaries.

For some countries the privacy art regulates how data are collected by their governments. The privacy art ostensibly applies to all personal data about an individual held anywhere in the government. Additional laws regulate the collection and maintenance of personal data by other organization but unfortunately these laws are inconsistent and typically apply to only one target domain example of such laws are the fair credit reporting act, the health insurance portability and accountability art, the federal education rights and privacy art and so forth with respect to privacy laws.

The EU's European privacy directive requires that both governments and organizations operation within the European Union maintain privacy rights when collection and using data about individuals. This directive also largely requires the governments and organizations adhere to the 8 information collection principle mentioned earlier.

Other societies such as Canada, Australia and Japan have also instituted laws that regulate the collection and protection a personal data relating to their citizen.
Many people do not realize countries such a United State government web sites are required by law to address five specific factors relating to privacy namely choice, notice, security, access and enforcement. With respect with choice individual using government web site must be given the option to choose whether to provide personal information, and if so must be able to choose how those data will be used. With respect to notice the data collection practices used by a government web site must be disclosed to individuals before the website collects

their personal information. With respect to security: private information that is collected from individuals via a government website must be secured against unauthorized

6.1.5 Collected personal data

Previous section on privacy security overview gave an insight or knowledge about how information privacy should be handled. Project X collects data from users such as email address, first name, last name, username or nicknames and password and these are few data users are required to fill into the inputs form.

6.1.6 Content and Information

Project X believes that contents and information shared on the site is owned by user themselves, therefore user are responsible for what they share as well as shared contents and information on the site can be controlled by using privacy and account settings. Users are able to customize on how other users interact with them.

The thesis paper suggests that the language used to write privacy policy and term of use should be easily understood by readers, and the collected information whether it is anonymous or identified should state so.

Throughout this section we shall explore in great details about the advantages and disadvantages of using social media and what information might be safe to share or post on social networking site, also how to protect the information in order to make sure that what is posted can only be accessed by intended users.

Individual chooses information he/she is willing to share within the site, most profiles shows user's town, home address, e-mail address and sometimes even phone number. Later on it might be opportunity for users to share things like where they attended their schools, personal interest, place of work and much more.

To save the purpose of the thesis, the network allows users to create their profile that gives them capability to connect with other users and finally create friendship

When a user registers to any website, he or she is forced to agree on to the terms, conditions and the privacy policy of using the site. Terms and conditions are concern about the liability of using the site at the same time privacy and policy are concerning about how user’s data is used within the site. It’s very important both sides to be honest to each other, often times

some users don't bother to read terms and conditions of using the site and on the other hand user's information is misused without user's knowledge.

Being honest is a good thing, on how users' data is used and giving them assurance about their privacy and security when they are using the site. The issue of user's privacy is not taken seriously as it should, big sites like Facebook has had a lot of criticisms about the way they deal with users data.

The aim of thesis project is to improve and enhance project X user's privacy security as well raise concern about terms and conditions of using the site, by being clear and honest when defining policy is very important, users get bored when the have to read numerous pages of text before they have to login and use the service.

Users should be informed straight forward who can get to access their data that they are adding on the site, users should be also informed what is going to private and what is going to public or are all profiles public or private? This is very important to let them know in advance that whatever posted is for their own risks.

In case if none of the above is true, then users should be informed how much information is available and to who can access it and how information can be restricted. It's very important users should be allowed to change their privacy setting to have the flexibility of what profile information they may want to share to the public, and whether it's public to their friends or to everyone else, things such as date of birth and contact details.

User's data should be held securely to insure that no one can access it unless users themselves give permissions to do so. Make sure the users data is kept up to date and is relevant, users fully request of their profiles removal should be given to them whenever they want to leave the site.

7 Conclusions and recommendations

I would like to express once again my great appreciation and enormous thanks to all involve in thesis project. This is being a great learning process and an opportunity to apply skills and knowledge to a real world project. The paper solve the challenges project X were facing but rooms are still available for more improvements. The results were tested on local server and showed to the client. The client will decide whether to apply the outcomes to a real web site on not. The paper explained some of hot security vulnerabilities available on today's website

and how to overcome them, including SQL injection, cross site request forgery and cross-site scripting (XSS). Among other things the paper discussed were means to prevent automated registration robots, filtering email format and validations of users' data before stored on the database.

The paper has rewrite and improve the project X privacy terms and condition of use that is accepted before their users get registered. Writing privacy terms and condition can be a thesis by itself therefore, what this paper suggested is not all written here but few guidelines were used to formulate a well organize privacy terms and condition are listed above.

8 References

Bryan, V. (2012). Web application security. Chicago: Mc Graw Hill.

Schwartz, P. M. (2010-2011). Information Privacy. New York: Wolters Kluwer.

Shaw, T. J. (2012). Information Security and Privacy. Chicago: Adventure Works Press.

Solove, D. J. (2015). Information Privacy Law. New York: Wolters Kluwer.

Duckett, J. (2011). HTML & CSS Design and Build Websites. Indianapolis, Indiana: John Wiley & Sons, Inc.

Chesters, J. (2015) JavaScript Frameworks in the Real World. Info Queue, 2 March. Accessed 24th May 2015. http://www.infoq.com/research/javascript-frameworks-2015

Jr., T. C. (December 2, 2013). Simple Terms And Conditions And Privacy Policy For All Businesses. Califonia.

June , J. P., & Oja, D. (2014). New Perspectives on Computer Concepts. Boston, MA 02210: Course Technology.

Lavin, P. (2006). Object-oriented PHP concepts, techniques, and code. San Francisco: William Pollock.

Steve , S., Tim , C., & Joyce , P. (2009). PHP 6 and MySQL 6 Bible. Indianapolis, IN 46256: Wiley Publishing, Inc.

Web application security 2012 ChicagoMc Graw Hill

Cross-Site Scripting (XSS) Tutorial :(2014, October 13). Retrieved from http://www.veracode.com/security/xss

SQL Injection Tutorial :(2014, October 13). Retrieved from http://www.veracode.com/security/sql-injection

Cross-Site Request Forgery Guide: (2014, October 13). Retrieved from http://www.veracode.com/security/csrf

Google Terms of Service: (2015, January 1). Retrieved from http://www.google.com/intl/en/policies/terms/

Telling Humans and Computers Apart Automatically: (2014, October 12). Retrieved from

http://www.captcha.net/

http://www.cc.gatech.edu/~orso/papers/halfond.viegas.orso.ISSSE06.pdf

9 Figures