

CERTIFICATION OF APPROVAL

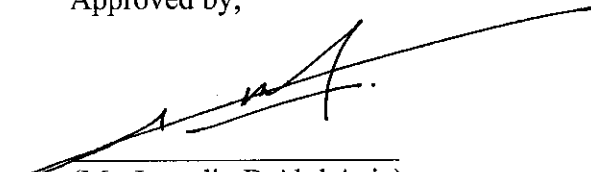
Secure Online Voting System for MPPUTP

by

Hanafi Bin Ab Kadir

A project dissertation submitted to the
Information Technology Programme
Universiti Teknologi PETRONAS
in partial fulfilment of the requirement for the
BACHELOR OF TECHNOLOGY (Hons)
(INFORMATION TECHNOLOGY)

Approved by,



(Mr. Izzatdin B Abd Aziz)

UNIVERSITI TEKNOLOGI PETRONAS
TRONOH, PERAK
June 2006

t
TK
5105.888

.H233

2006

- 1) Internet programming i
- 2) web site development
- 3) IT / IS -- threats

CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.



HANAFI BIN AB KADIR

ABSTRACT

This study is about an online voting system prepared for the Final Year Project. The aims of the project were two: to discover what would be required of an electronic voting system to make it a suitable replacement for the existing paper-ballot system; and to begin the process of designing and implementing a system which could meet the requirements. The first goal was achieved through research into the opinions of experts on electronic voting and computer security. The second was achieved with the use of formal methods.

ACKNOWLEDGEMENT

During the process of completing this project, many obstacles had been gone through but some people had made it easy for me. Many thanks to my supervisor, Mr. Izzatdin bin Abdul Aziz who had helped me a lot by giving a bunch of good ideas. Also to my colleagues especially who are also under the supervision of Mr. Izzatdin. They had been very kind and their generosity in giving friendly advice is very much appreciated. Not to forget my family at home who had always given me moral support as well as financial support. I love you very much.

TABLE OF CONTENTS

CERTIFICATION		i
ABSTRACT		iii
ACKNOWLEDGEMENT		iv
LIST OF FIGURES		vi
ABREVIATION		vii
CHAPTER 1:	INTRODUCTION	1
	1.1 Overview	1
	1.2 Background of Study	1
	1.3 Problem Statement	2
	1.4 Objectives	3
	1.5 Scope of Study	3
	1.6 Timeline	4
CHAPTER 2:	LITERATURE REVIEW	5
CHAPTER 3:	METHODOLOGY	11
	3.1 Requirements	12
	3.2 Designs	14
	3.3 Implementation	15
	3.4 Integration	15
	3.5 Maintenance	16
CHAPTER 4:	RESULTS AND DISCUSSION	17
	4.1 Results	17
	4.2 Discussion	17
	4.2.1 System Flow	17
	4.2.2 Security Elements	19
	4.2.3 Problems faced	20
	4.2.4 Use Case Diagram	21
CHAPTER 5:	CONCLUSION AND RECOMMENDATION	22
	5.1 Conclusion	22
	5.2 Recommendations	23
APPENDICES		24

List of Figures

Figure 3.1: Waterfall Model	10
Figure 3.2: Requirement Types	11
Figure 3.3: System Flow	13

Abbreviation

IIS

Internet Information Services

SDLC

System Development Life Cycle

SSL

Secure Socket Layer

CHAPTER 1

INTRODUCTION

1.1. Overview

This study is about an online voting system prepared for the Final Year Project. It starts with a brief description of election in our nation and the importance of an online voting system based on several problems faced when using the traditional system. Then, it continues with the findings in areas like security and architecture. Next, the methodology used to develop the online voting system is described. Finally, it ends with a conclusion and a few recommendations that could be used as a guide in order to enhance this system.

1.2. Background of Study

Malaysia is one of many nations in the world which practise democracy system. In a democracy system, election method is used. The political leaders are chosen based on the majority of the votes. During an election, many election booths are set up around the country, in rural and urban area. Then, all valid voters who are at least 21 years of age and registered go to their nearest booths to vote. After the invalid votes have been discarded and all votes have been counted, the commissioners will announce the results.

1.3. Problem Statement

Politics in most nations in the world practise democracy system which allows the citizens to vote their leaders. The present election system forces the citizens to make the election at the election sites only and it is manual from throwing the votes to counting them. This traditional system will make room for mistakes which most probably done during the counting process. Further destruction would be the mistakes are not traceable. People will not notice if there are mistakes in the counting process.

The second problem is that the present system takes a lot of time. Citizens are crowded at the selected election sites and this really consumes time. Further problem would be the citizens usually take this opportunity to skip work time especially for the government servants.

The third problem in the present system is that when it forces the citizens to vote at the selected election sites only, the aged citizens would have problem to go. This problem is also faced by people with disabilities. Thinking that it would be dangerous and difficult for them to go to the election sites which are congested with people, they might not go voting.

1.4. Objectives

There are two objectives of this project:

1.3.1. To make a study on online voting system and its security issue.

This study is very important in order to understand how the current voting system works and how it can be enhanced. There will also be a study on the Internet security technology so that the best technology can be applied to the system.

1.3.2. To develop a reliable and secure online voting system

Based on the study, a reliable and secure online voting system can be developed. The developed system should be useful to organizations like the Student Representation Council.

1.5. Scope of Study

When we talk about an Internet voting system, area of security and reliability would surely come into picture. But to develop such system would be very costly and time consuming. For the Final Year Project, its scope of study will be narrowed down so that this project will be feasible and could be completed within the time frame. The scope of study includes:

- a research made on the current issues of Internet voting such as the advantages and disadvantages of online voting
- development of the system which covers the aspects of application layer and network layer security, reliability of the system and also part of legality issue.

Some aspects which are out of the scope of study would be:

- development of a robust online voting system which can not be harmed at all.
- covering all of the ethical issues in voting.

1.6. Timeline

Based on the milestone attached in Appendix on the page 32, it is shown that during the first semester, this project focused more on research. The research performed was mainly on security technology and also the process flow of the online system.

During the second semester (refer to page 33), the project was focused on development rather than making study. However, some research was still performed during the second semester due to its need. Most of them were research on programming.

CHAPTER 2

LITERATURE REVIEW

“Voting via the Internet has become a feasible option for political as well as non-political ballots.” [Schryen 1, 2004]

Schryen has stated that Internet voting can benefit the political and non-political elections. This proves that the Internet voting system which yet to be developed can benefit the UTP Student Representation Council.

“As Internet voting is an additional channel for eligible voters the turnout might increase substantially.” [Schryen 2, 2004]

From this sentence, it shows that by using Internet voting, more voters will turn up to vote. This is an advantage for using Internet voting over existent manual voting.

“Cost savings can occur, if less personnel for performing absentee voting and for counting is necessary or if travel activities are reduced. On the other hand building up and operating the poll infrastructure as well as equipping the voters with essential hardware cause cost (see section four). Furthermore, in the foreseeable future of political elections no polling stations will become obsolete.” [Schryen 3, 2004]

To Schryen's view, the overall cost of voting can be decreased if an Internet online system is used. This is due to fewer personnel in charge for performing absentee voting and counting, no poll infrastructure to build and operate and no polling stations will become obsolete.

"Invalid votes can be produced consciously or unconsciously. Consciously producing invalid votes are presumably protest against politics in general, therefore they must be provided in online elections. Unconsciously produced invalid votes could be already identified at "feeding time" with plausibility checks, so that the voting software could point out this mistake." [Schryen 4, 2004]

This shows that invalid votes will not be produced in Internet voting since invalid votes could be identified before the votes being counted. This is another advantage for Internet voting.

"The impact of such vote tampering depends on several factors. Two of the most important are the scale of an attack and the competitiveness of the contest." [Fischer A, 2003]

Based on this statement, it is shown that corruption in voting is influenced by the scale of an attack and the competitiveness of the contest. If the scale of an attack is small, it will not affect the result of the election and if the competitiveness of the contest is low, then vote tampering is not important for the attacker.

"One type of attack might gather information that a candidate could use to increase the chance of winning. For example, if vote totals from particular precincts could secretly be made known to operatives for one candidate before the polls closed, the results could be used to adjust get-out-the-vote efforts, giving that candidate an unfair advantage." [Fischer B, 2003]

From this statement, we know that it is dangerous to leak out information on the current state of the votes because it may be used by the candidate to give an unfair advantage even though the voting process is still running.

“Technical Vulnerabilities. This category includes weaknesses stemming from the computer code itself, connection to other computers, and the degree of auditing transparency of the system.” [Fischer C, 2003]

This statement tells that there are possibilities to have technical errors in the voting system such as computer code, network connection and auditing transparency of the system. So, I should be very vigilant in designing and developing the system.

“Social Vulnerabilities. A significant and increasingly sophisticated kind of attack — dubbed “social engineering” by hackers — involves finding and exploiting weaknesses in how people interact with computer systems. Such social vulnerabilities can include weaknesses relating to policy, procedures, and personnel. Of the 14 specific risks identified in the Maryland study, most were of these types.” [Fischer D, 2003]

However, there are also possibilities to have social vulnerabilities in the Internet voting system. This kind of vulnerabilities is an opportunity to hackers. They can find and exploit weaknesses in how people interact with computer systems.

“There are at least two good reasons to store connection strings outside the application – flexibility and security.” [Esposito, 2006]

Based on the research from Esposito, it is believed that by storing connection strings outside the application, the electronic voting system would be more flexible and secure.

“Password recovery becomes necessary when the user of a system is no longer able to authenticate themselves because they have lost or forgotten their password. Any systems that require authentication will need to have some policy or procedure for password recovery.” [Miller, 2002]

Miller stated that any particular system which has an authentication system, should provide password recovery service to help users who loss their passwords.

“Hash functions are an important tool in the security armory: they are guarantors of the integrity of a piece of information, and as such are used in applications from database security to digital signatures.” [Bursell, 2005]

This statement tells that it is important to use an encryption technique like hash function when storing confidential information like passwords into database.

“One-way hash functions, or simply hash functions, take a given plaintext and condense it down to a number of a certain size (“hash”), with the stipulation that it must be very difficult to construct another plaintext that, when run through the same function, produces the same hash value (a “collision”). This is often useful when you need to verify that two things are the same without actually storing or transmitting them, such as computer passwords or very large files.” [Argentini, 2004]

From the statement above, we learn that a hash function is really useful for computer passwords. None of the actual passwords are stored in the database. Only hashed passwords are stored. To verify users, passwords entered by them are hashed before being compared to the hashed passwords in the database.

“The *anonymity* of a voter’s ballot must be preserved, both to guarantee the voter’s safety when voting against a malevolent candidate, and to guarantee that voters have no evidence that proves which candidates received their votes. The existence of such evidence would allow votes to be purchased by a candidate.” [Kohn, Stubblefield, Rubin, Wallach, 2003]

This statement shows that there should not be any evidence showing who the voters had voted for. This is to avoid fraud.

“That said, we demonstrate that the insider threat is also quite considerable, showing that not only can an insider, such as a poll worker, modify the votes, but that insiders can also violate voter privacy and match votes with the voters who cast them.” [Kohn, Stubblefield, Rubin, Wallach, 2004]

It is stated that we have to take consideration of what might happen if the insiders perform such misconduct.

“As for uncoercibility, a system which allows the voter to change his vote until the last minute seems like an interesting solution idea (which is actually implemented in several protocols), as this makes coercion not impossible, but still a lot harder. Yet unnoticed but still useful would be the idea to open the polling stations longer than allowing voters to vote electronically. So if »electronic coercion« takes place, the voter can still change his vote by voting physically in a better protected location.” [Karatsiolis, 2006]

Karatsiolis’s idea to reduce the possibility of electronic voting coercion is by allowing the voters to change their votes even if they have voted before. But still, there must be a time limit which is probably just before the vote count.

“He believes that the weakest link in voting systems is probably voter registration, rather than vote collection, which receives all the attention. He highlighted the fact that the voter is not a computer.” [Rivest, 2004]

Dr. Rivest said that the voter registration process is more important than vote collection process. He also said that there might be errors done by voter during registration. So, the system must provide clear instructions to voters.

CHAPTER 3

METHODOLOGY

The methodology that will be used for this project is the waterfall model. Its name describes its behaviour. It has cascading effect from one phase to another just like the illustration below.

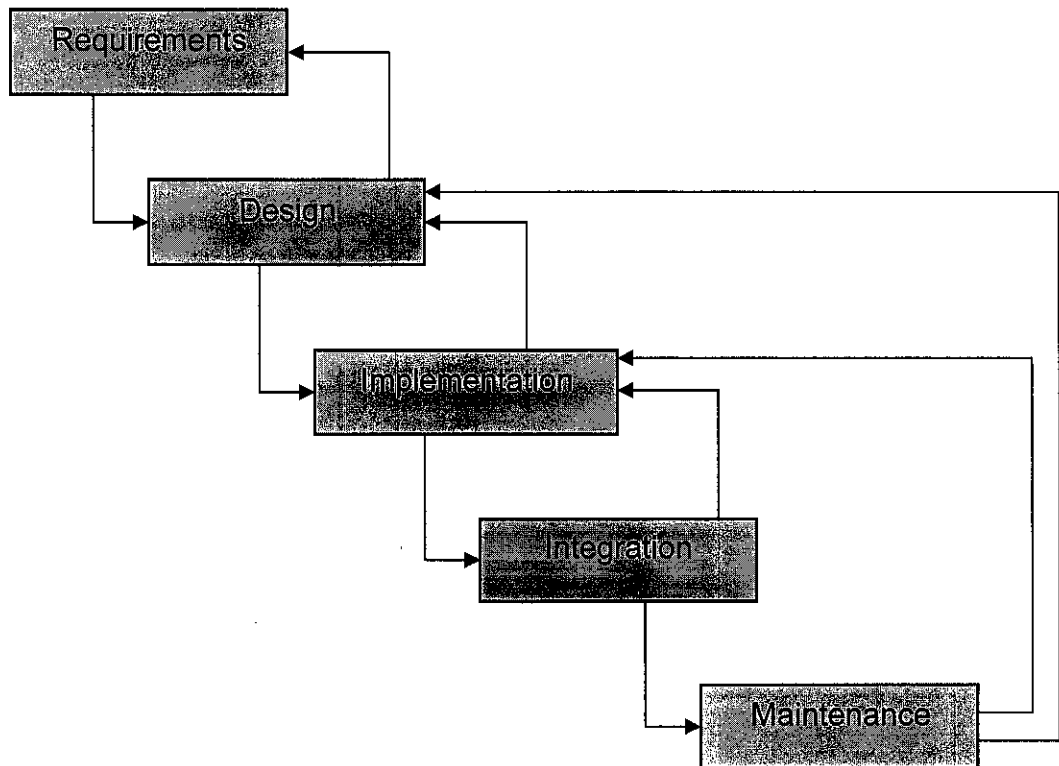


Figure 3.1: Waterfall Model

3.1. Requirements

There are four different kinds of requirement. As illustrated below, they are legal requirement, economical requirement, ergonomic requirement and technological requirement. In the technological requirement, there is a sub requirement which is the security requirement.

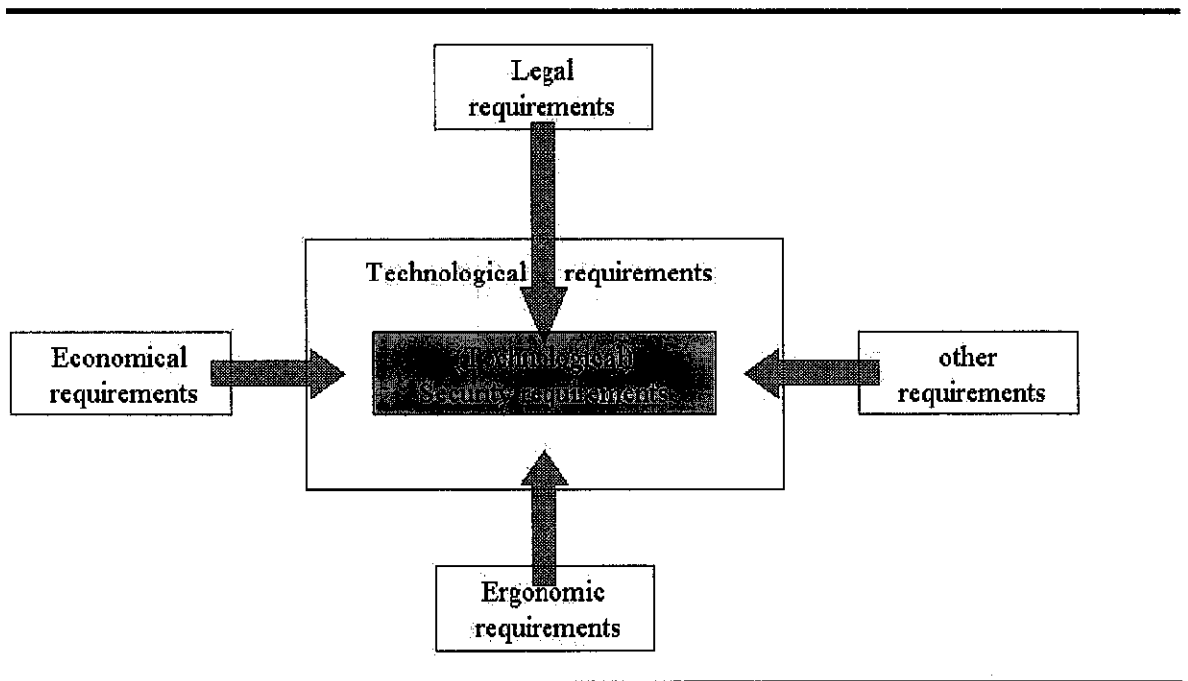


Figure 3.2: Requirement Types

Legal requirement deals with the protocol and how the voting process is going to take place. Different elections have different rules. Elections of the national level have tighter rules if compared to elections of university level. For example, election of the national level may allow vote for only one candidate but election of the university level may allow votes for more than one candidates.

Another issue for legal requirement is who have the authorities to count the votes, view the results and recount the votes. In the Internet voting system, the vote counting is going to be handled by the system itself. But if there is a need to recount, then the authorized person should have an access to the database because the votes are stored in a database. After the vote counting is done, everybody should be allowed to view the result. But if the voting process is still going on, the result of the current state should not be viewed even by the administrator as to avoid any fraudulence.

Economical requirement plays an important role as we know how much money and effort to spend for a certain project and how much of them we have in hand. This is a crucial requirement to follow since many IT projects failed because they were overspending on their projects when the funds were not enough.

Another requirement to meet is the ergonomic requirement. It offers usability and comfort for people to use the system. In order to prepare this requirement, I have to consider the IT knowledge of its users. Users in urban area would surely be more familiar with computers than users from rural area and university students are assumed to have basic knowledge to use computers. So, very basic instructions on how to use the Internet voting system are not necessary for university students.

The last and has most vital requirement for Internet voting system is the technological requirement where it also describes about the security requirement. Level of security for elections differs from each other. Election in the national level surely has more advance security if compared to an election in university level. This is because elections in the national level attract more threats from attackers from all over the nation since they are more competitive and have bigger scale of attack.

3.2. Designs

Based on the chart below, there are five steps involved in the Internet voting process. They are complete voting, encrypt data, record data, send notification to voter and count votes.

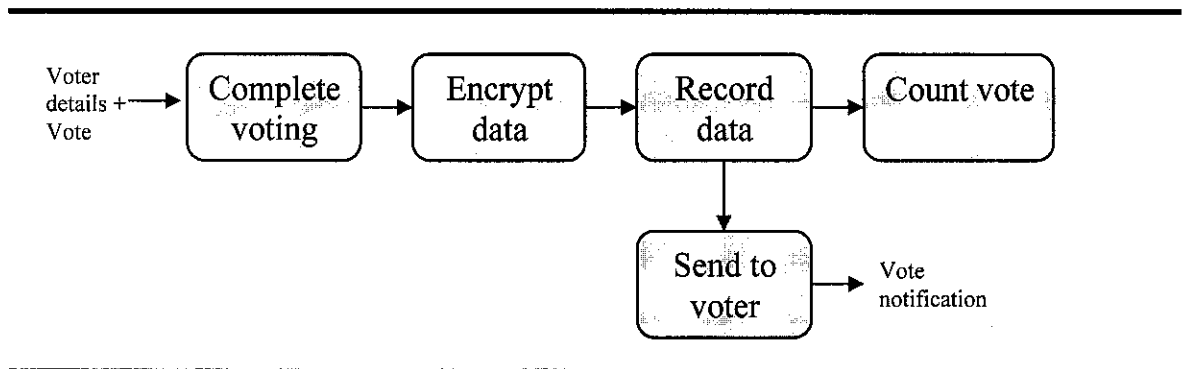


Figure 3.3: System Flow

Firstly, a voter has to fill a certain particular such as name, I/C number and ID number as well as the vote. Then, the voter's details and vote will be encrypted separately. After all information has been received, system will send a notification to the voter informing that his or her vote has been successfully casted. Next, this information will be decrypted and vote will be counted. Then, the vote and its voter's particular will be recorded in two different databases. This separation is needed in order to avoid any attempt to dig out which candidate did the voter voted for.

3.3. Implementation

The language that will be used for this project is the ASP.Net using VB.Net as the code behind. ASP.Net is a server-side scripting and it is easy to understand. Plus, it has several built-in cryptographies that I can use to enhance the security of data travel. Besides that, by using VB.Net as its code behind, I can write important codes which could not be viewed by the Internet users since the VB.Net file is excluded from the ASP.Net and it is hidden.

One of the tools which support ASP.Net is the Visual Studio.Net. Internet Information Services (IIS) is going to be running as the server as it is easy to use and comes along with the Windows package. As for the database, I prefer MS SQL than any other software. MS SQL which is manufactured by Microsoft has been well known as one of the most reliable database equipped with the latest security.

3.4. Integration

Two methods will be used to test the integration between all the modules in the system. They are top-bottom testing and black-box testing. Top-bottom testing tests the results made by a combination of several modules. Nothing much different from top-bottom testing, black-box testing also tests several modules as a one unit but it focuses more on the behaviour.

3.5. Maintenance

For maintenance, consistent checks have to be performed so as to detect any misbehaviour done by the system. Any defect has to be recorded and correction action has to be made. If the problem persists, then I have to go back to the previous stages of the System Development Life Cycle (SDLC) which are design and implementation to solve the problem.

CHAPTER 4

RESULTS AND DISCUSSION

4.1. Results

The product of this project which is the electronic voting system has been tested and results have been recorded:

- Number of turnouts when use paper system = 1300
- Number of turnouts when use electronic system = 1980
- Percentage of growth = 52%
- Time taken to count vote when use paper system = 3 hours
- Time taken to count vote when use electronic system = Almost instantly

Based on the above result, it can be said that the objectives of electronic voting system has been achieved although the growth of turnouts is not really high. This happened because the system was tested during two weeks before final examination when UTP students have got a lot of projects to complete. Furthermore, this test was not made formal and made known to all students.

4.2. Discussion

4.2.1. System Flow

This test started from voter registration process until showing the result of the election to public. Firstly, users have to register using the online registration system in order to get a password. The password is auto-generated and sent to the users through emails so that only they know their password. The

passwords will not be directly stored into database. The passwords need to be hashed first before being stored. The users then have to use the given password to sign in into the voting system before they could cast their votes.

Not all can access the voting area. Only authenticated users can access it. To verify the users, passwords entered by users will be hashed. Then, they will be compared to the hashed password in the database. After the users signed in, they will be directed automatically to either student's area or administrator's area.

In a student's area, users have options to cast vote, recast vote and also see results. When they cast votes, they will not be given a receipt stating which candidates they have voted for. This is to maintain voters' privacy by eliminating any evidence of votes. If the votes are not kept secret, it might invite fraudulence. How? The evidence will become a proof that the voter has voted according to what a vote buyer has asked for.

After the voters have performed the casting, they will be marked as voted users. It means that they can not cast another vote. This limitation is to ensure that the system has a good reusability procedure. If there is no limitation to users, they will vote again and again, making the election unfair.

Another choice in a student's area would be to recast vote. This option is given to students who want to make a change to their votes. Of course, the previously casted votes will be discarded and will not be shown. This option is an option until the time comes to count all the votes. Users have this option of recasting their votes so that the robustness of the system is upheld. For example, if a voter is forced to vote a candidate, he or she can still change the vote by selecting the recast vote option.

In the other area which is the administrator area, users can manage the candidates database. They have the authority to add, edit and delete any candidate they want. However, they do not have the option to cast any vote. Lastly, they can stop the voting system from accepting anymore vote when they choose to yield the result of the election. Only now, everyone can view the results including the users in the student's area.

4.2.2. Security Elements

Now, let us discuss on the security area. The first security element used in this system is hashed password. A hash is a numerical value of fixed length which unequivocally identifies files of arbitrary lengths. An example of a hashing algorithm is SHA1. People might now say that saving the password as a hash would be sufficient, but why is this wrong?

The reason for this is that usually so called 'Dictionary Attacks' are run against hashed passwords - a good example being the MD5 hashed passwords of NT4. This is a Brute Force attack: all entries in a dictionary were hashed using MD5 and those hash values then are compared against the password database. Have a guess how quickly some passwords are found this way.

The intention behind a Salted Hash is to have this type of attack fail by attaching a random value - the so called salt - to each password and only then compute the hash over password and salt. For comparison of the password the salt has to be stored alongside the salted hash, but the only vector of attack is to re-code the dictionary for each individually stored password with the salt - and this takes quite a long time. Another way of doing salted hash is by combining the password with the salt, and only then we hash them. This way, a hacker would never know if there is any salt in the password. This system uses the second technique where salt is not stored separately from password.

One thing to remember now is that the hashed password could not be used to retrieve the original password. So, for password recovery system, this voting system will provide new passwords for users who lost their passwords through emails, the same mean as they get their first passwords. The password recovery service is very important as it can sustain the users' accounts. If there is no password recovery service in this system, users have to create new account once they loss their passwords. If this happens, it will create such a messy database and users can vote as many times as they want as new users.

The second security element in this system is encrypted connection string. A connection string is where you store information in order to connect to the database. The information in a connection string would include database name and password to access it. That is why it is so important to encrypt the connection string to avoid people from see it with bare eyes. If the connection string is not encrypted and somebody has an access to the server, he or she can always find the database and access it with the information in the connection string. If it happens, the integrity of the database is doubtful.

The third security element implemented in this system would be the method of giving away passwords to users. This system only gives away auto-generated passwords though emails. This way, only the intended users can know the passwords, not even the database administrators know them.

4.2.3. Problems faced

In completing this project, there were a few problems occurred. One of them happened during the mail server setup. Incorrect setup of mail server would cause the system of sending passwords through emails to fail. This was worsening as there was no error message or what so ever to notify what went wrong to the mail server. In term of the system itself, it would be easy to detect any coding error and corrective action could be taken immediately.

Second problem was during the implementation of Secure Socket Layer (SSL). To implement an SSL, I have to subscribe to the SSL provider. In my case, I used SSL provided by Startcom because it was free of charge and they were trusted by many people. They also have a good record in providing over thousands of SSL to their subscribers. But the problem was that I did not get the SSL certification that I should get from them. When I issued another proposal, they said I had already given an application. So, I had to leave the SSL aside from this system due to time constraint.

4.2.3. Use Case Diagram

There are two actors involved in the system (refer to the use case in Appendix). They are student and administrator. Student has the option to register him/her self in case he or she has not registered yet. Then, after receiving a password through email, he or she can cast the vote. When the election process has finished and the administrator allow student to see the election result, then only the student can view the result.

Administrator has a duty to update the candidate database which includes the process of adding, editing and deleting. Next, after finalizing the candidate database, the administrator can allow all registered students to cast their votes. When the time is up for the students to cast vote, administrator can disallow students to cast any votes. Lastly, administrator can allow all students to see the election results.

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.1. Conclusion

The present voting system is not very efficient in terms of time and work force management. Counting votes process take much more time since it is done manually. Man power has to be allocated to count the vote and to prepare the voting booths. This is very costly and if recounting is necessary, then it will cost even more. Furthermore, counting votes manually may allow mistakes.

Voting using Internet voting system is a solution for the above problems. It will count the votes as the voting process is still running. This will save up much time even if the system needs to recount all the votes. Vote counting is going to be done by the system itself thus eliminate any human error in the counting process. Money can also be saved since there is no need to pay for counting the votes. Another advantage of using this system is that it can be operated 24 hours everyday provided that the server is up.

5.2. Recommendation

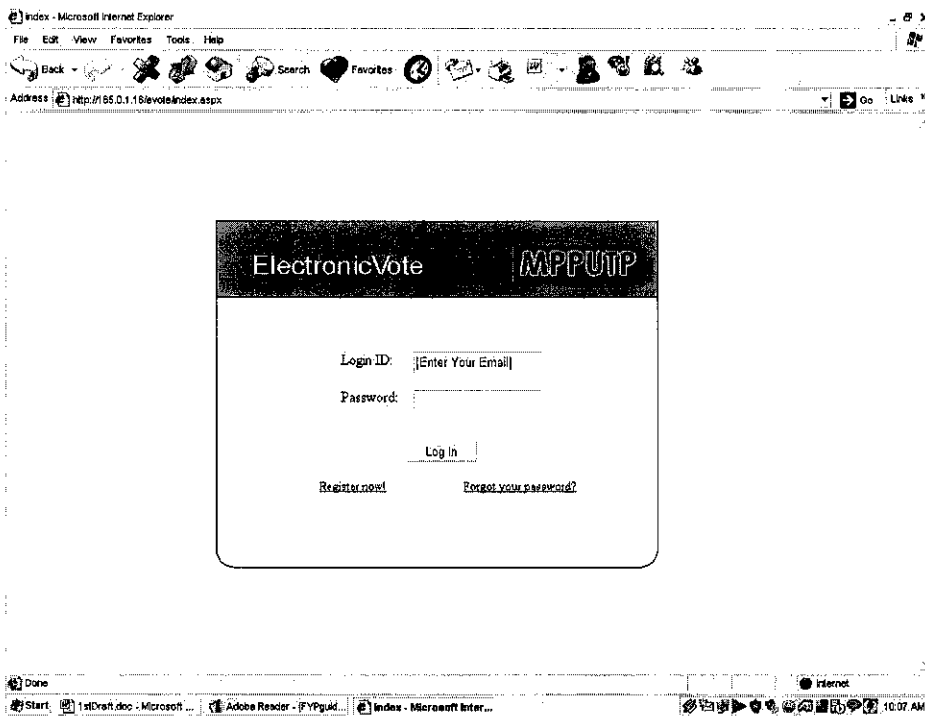
Even though this system sounds good enough but we have to remember that there is no technology which is good enough for tomorrow. There are a lot of rooms to improve the system. But due to limited time and money, I could only develop the system which is already described in the previous chapter. Here, I would like to suggest a few things in order to make this system even better.

- Security in network layer and transport layer should be developed since this system only caters cryptography for the application layer.
- A firewall should also be designed to prevent any attempt to destroy the system.
- Lastly, if this system has a sufficient level of security, it should be prepared for the national election. It can be prepared in term of legitimacy and availability for the whole nation.

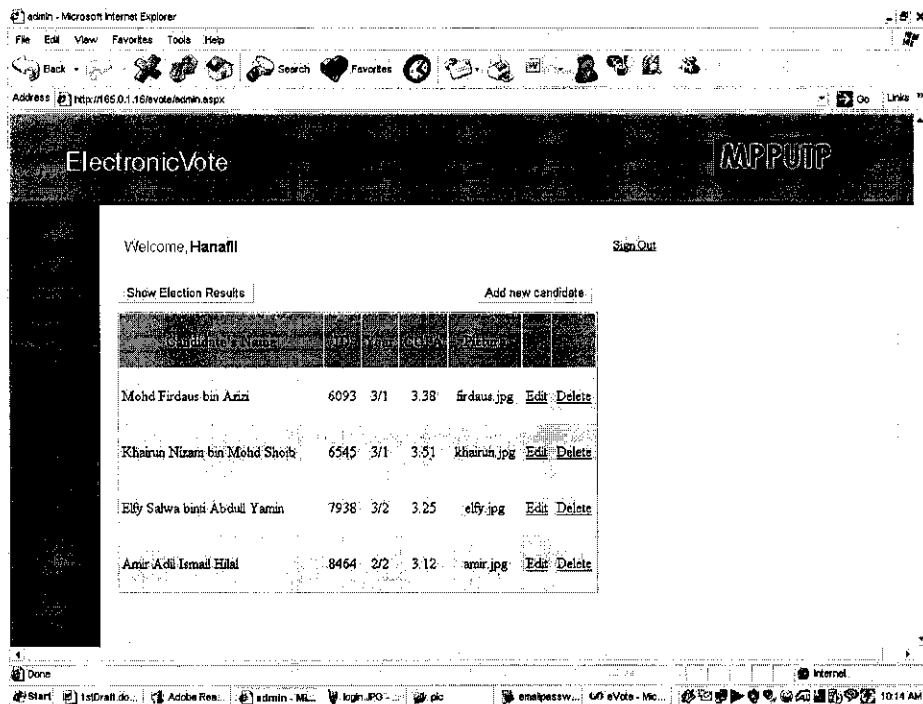
APPENDIX

Screen Shots

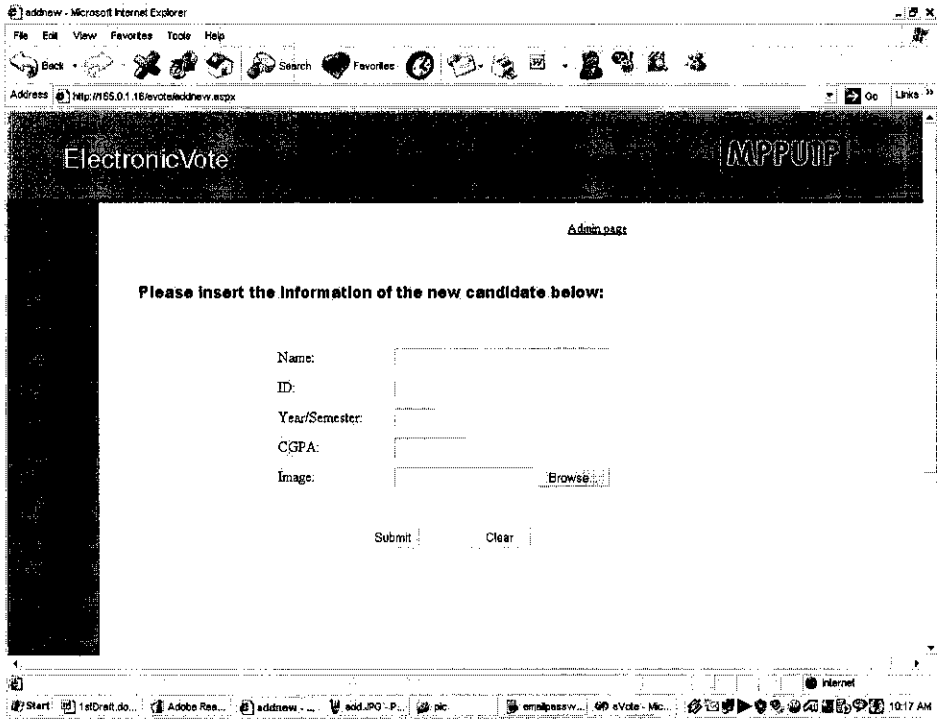
Login page:



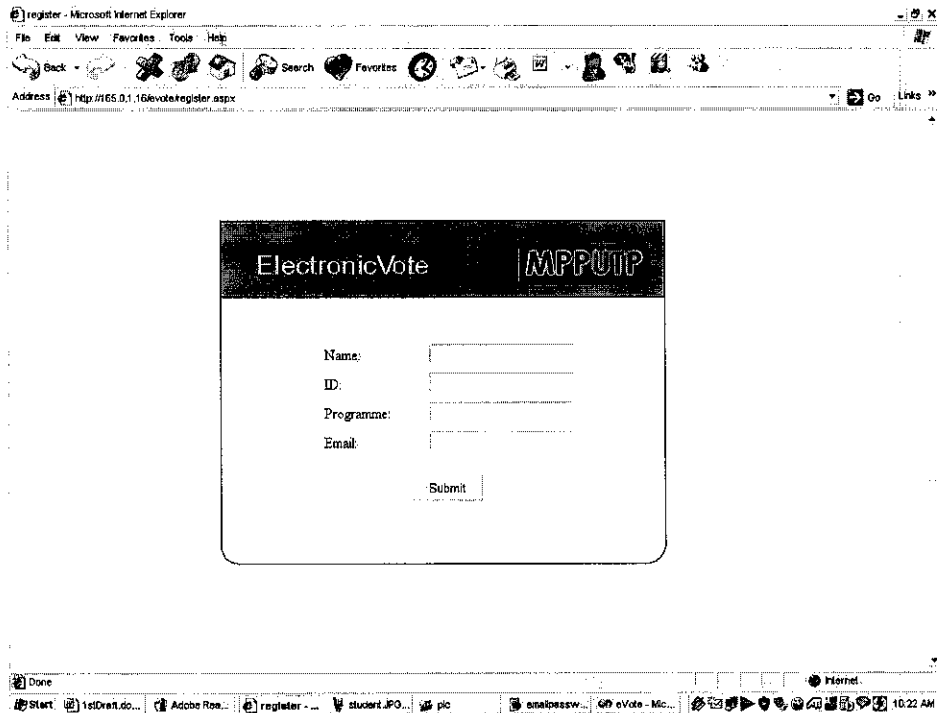
Administrator page:



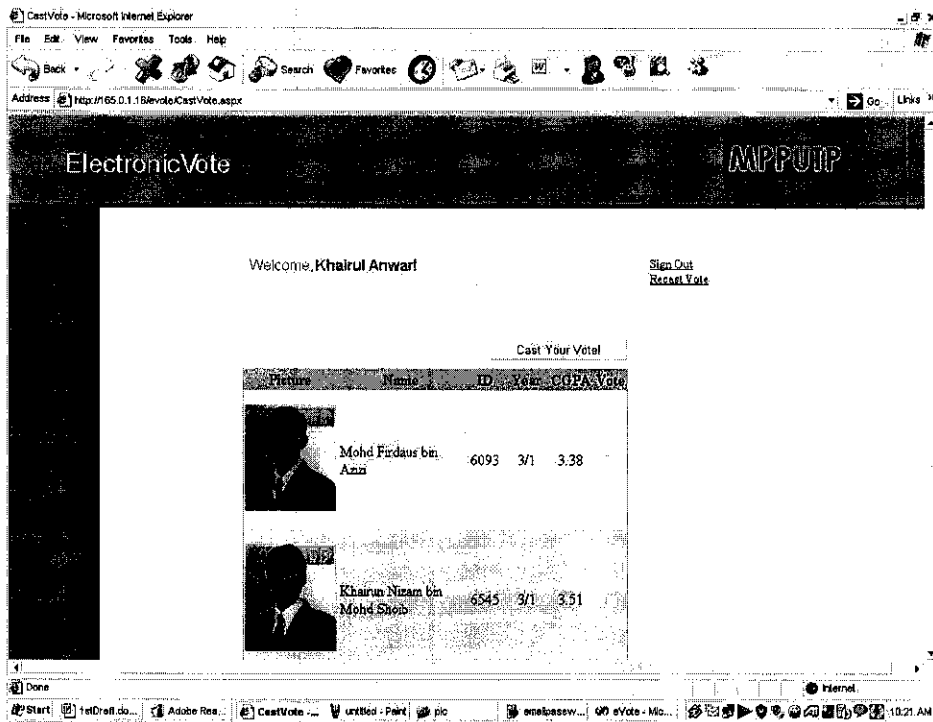
Adding new candidates page:



User registration page:



Students' page:



Milestone

Use Case Diagram

