Automatic Hoax Detection System

By

Adzlan Bin Ishak

Dissertation submitted in partial fulfilment of

the requirements for the

Bachelor of Technology (Hons)

(Business Information System)

AUGUST 2011

Universiti Teknologi PETRONAS
Bandar Seri Iskandar
31750 Tronoh
Perak Darul Ridzuan

Abstract

Hoaxes are non malicious viruses. They live on deceiving human's perception by conveying false claims as truth. Throughout history, hoaxes have actually able to influence a lot of people to the extent of tarnishing the victim's image and credibility. Moreover, wrong and misleading information has always been a distortion to a human's growth. Some hoaxes were created in a way that they can even obtain personal data by convincing the victims that those data were required for official purposes.

Hoaxes are different from spams in a way that they masquerade themselves through the address of those related either directly or indirectly to us. Most of the time, they appear as a forwarded message and sometimes from legit companies such as PayPal. Having known the threat that this non malicious brought, it is important for us to address this problem seriously by providing an automatic hoax detection system as the solution to this matter. Consciousness and Awareness are definitely the first step to be taken for this matter

Acknowledgement

Alhamdulillah I bid for granting me the strength and perseverance necessary to complete this project. The past 8 months have indeed been an amazing experience and journey to embark on. I have learnt so many things from scratch until the completion of my own proposed system which has been fruitful to my learning curve thus far.

I would like to sincerely thank my one and only supervisor, Ms Elaine Chen Yoke Yie for the undying help she has given my colleagues and I throughout the whole 2 semester of Final Year Project. She has been very kind in guiding me to complete my project at this rate of completion thus far. Although there were many complications met throughout the whole journey, yet, she managed to guide us patiently with a strong sense of direction in this project. Ideas and suggestions kept pouring endlessly from her and that they have indeed helped me a lot in getting a lot of problems solved. Thank you, Ms Elaine for your support and guidance. Only God knows how to repay such kindness in equal.

I would also like to take this opportunity to thank my family especially my mother for always being there for me. The warmth of love given has always ensures a high morale confidence and perseverance in completing this project for me. They have been giving me a great deal of encouragement, advices and support which have led me to the completion of my current project today.

Lastly, thanks to my colleagues for the continuous support and guidance, and also to UTP; especially to the CIS Department for their patience and opportunities to allow my project to reach greater heights than it can actually be. Alhamdulillah again and may prayers be with all of you.

Thank you

CERTIFICATION OF APPROVAL

Automatic Hoax Detection System

By

Adzlan Bin Ishak

A project dissertation submitted to the

Business Information System Programme

Universiti Teknologi PETRONAS

in partial fulfilment of the requirement for the

BACHELOR OF TECHNOLOGY (Hons)

(BUSINESS INFORMATION SYSTEM)

Approved by,

(Ms Elaine Chen Yoke Yie)

UNIVERSITI TEKNOLOGI PETRONAS

TRONOH, PERAK

August 2011

CERTIFICATION OF ORIGINALITY

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

ADZLAN BIN ISHAK

Table of Contents

Table of Tables

Table of Figures

Abbreviations and Nomenclatures

Num	Abbreviations & Nomenclatures	Definition
1.	Hoax	Another sort of junk which can lead to misleading of information to the users or the reader of the email himself. Often come from trusted senders
2.	Levenshtein Distance	A method in which it uses a metrics to measure the similarity/differences between 2 strings
3.	Text pre-processing	Method used to process a certain text in which it can be used in the methods presented. For this project, it's mostly to remove the forward clause and prepare the text for hoax detection and insertion to the database via sql
4.	Meme's Theory	Hoax acts like a virus. It spreads easily once something's being perceived as truth although it isn't
5.	n-grams	A method to compare string to string. Complex but fast (not sure how fast though)
6.	Spams	Malicious mails/junks that's being forwarded to your mail. Often being filtered automatically by email service provider. Sender's address are alien to the ones in your contact list

Table 0-1 Abbreviations and Nomenclatures

CHAPTER 1

INTRODUCTION

1. Introduction

1.1 Background

E-mail hoaxes are known not to be malicious towards any created systems. By definition, this hoax is considered to be another sort of junk which can lead to misleading of information to the users or the reader of the email himself. It is important for this paper to distinguish the dissimilarities of both hoaxes and spam clearly in the very first place so as to avoid further confusion to the readers. According to [5], spam is a junk mail which is mostly unwanted or unsolicited which was sent either directly or indirectly by personnel who has no current relationship to the email user. In a simpler version of understanding, spams are mails which were sent from people who are out of the content of our address book. Hoaxes on the other hand, based on [1], are unsolicited and unwanted emails which directly or indirectly sent by personnel who has current relationship with user of the email. In another perspective, hoaxes are the 'smarter' version of spams which masquerade themselves well via the personnel that are present in our contacts.

Usually these hoaxes come in a form of forwarding messages from various sources. It is so tempting and often very convincing that according to [4] is a considerable virus which plagues the minds of the infected. Knowledge plays a vital role in the development of a person's growth. Misleading information within a society will act as a virus which is able to spread its disease as easily as it can [4]. Based on [2] and [4], hoaxes have been understood as to how and why they were being channelled through forwarding emails. At first, similarly like spams, they were introduced with the aim to promote a product. Later on, it was identified that it has the ability to gather information but the worst that could possibly happen in to convince a group of people of certain events which have never seem to occur before. Certain companies like PayPal have paid significant losses due to these hoaxes.

This is where the resolution to create an automatic hoax detection system came to life. Do note that there are not many hoax detection systems created thus far. Based on research, there are only up to 3 research papers thus far with relate to email hoax detection systems; [1], [2] and [3]. Fortunately, this project has its similarities with spams in a way that most of its detection patterns are the same but particularly different by definition [2] and [5] and that text mining tools [13] can be used to assist in this project too. Prior to [1]: the latest research on hoax detection system, the loophole to the previous system has been identified to be on its inability to update its database of hoaxes and to automatically update the database on the latest hoaxes identified using certain algorithms which will be based highly on probabilities. Once the new email has been identified after a certain comparison phase, these hoaxes will be identified and stored as a new hoax in the database. This paper aims to complete a finalized product of an automatic hoax detection system with the ability to update its database on its own.

1.2 Problem Statement

Email hoaxes are not malicious but often disturbing to the society. According to the Meme's theory [4], when a person's mind is being infected by the virus, it will be hard for that infected person to be convinced otherwise of the real truth. Having stated so, it is crucial that hoaxes should be identified primarily to avoid further misunderstandings which can bring significant losses as well as negative impact to a party's growth. Based on history, email hoaxes have given up so many false alarms which have given up a lot of negative impact to certain parties especially to the influential people. For example, PayPal has to undermine and took responsibility after a Nigerian scam (a type of hoax) was forwarded to most of its recipients by an unknown source which masquerade itself by using the email address similar to PayPal in 2003. Another false alarm on the existence of acid rain due to the explosion of nuclear reactor in Japan have also caused chaos to the South East Asians in which this hoax was forwarded to many after a week from the earthquake incident in Japan. Although these hoaxes have proven to be frauds yet, they are still capable in convincing many on the content of the hoax which was of course misleading [1], [2] and [3]. The problem can only be solved once the users were exposed to the real fact of these hoaxes. Even so, it will not be easy to convince a person on the real truth unless the source of the truth presented is valid to that person's view [4].

Another problem which can be identified is the fact that this system has been unable to update the database on its own [1]. This is essential to enable the system to run on its own without regularly updating the program. However, this problem cannot be resolved easily and needs a lot of consideration due to the fact that this system has a very few source of reference. Another problem that compliments to this problem is due to the fact that the latest system [1] is unable to train itself in identifying the characteristics of an email hoax thus, disabling the system to detect new found hoax which is not stored in the database for comparison.

1.3 Objectives

The main objective of this project is to develop an Automatic Hoax Detection System which is able to detect new and old email hoaxes presented in today's world. The main concern is to create higher awareness on the existence of email hoax and enlighten the users on the validity of every email's content that has been forwarded today

1.4 Project Relevancy

This project has a very high relevancy to be completed. This is due to the fact that hoaxes have proven to be notorious and unstoppable over the years. Despite of having anti spams and new virus signatures to abdicate viruses across the internet or systems, hoaxes on the other hand have often been neglected and taken cynically by the society. However, it is amazing as to how can humans tolerate themselves to be wrongly informed on the truth of certain events. Worst case for this matter is when the victims of hoaxes replied certain hoaxes which require them to give out certain personal details which can be used to either cheat the person or as a source of valid marketing data. Another relevancy of this project to the current society is due to the fact that it tore down the images of certain influential parties which have been aimed by these hoaxes. In a summary, hoaxes have proven to be non malicious yet, rather disturbing and often harmful if the victims are not aware of its existence.

Take the acid rain incident for example. While the email hoax presented may not be harmful to anyone and most would say, "Better be safe rather than sorry", this hoax has given a negative impact to BBC as well as to the user itself. Firstly, it gave out a negative impact towards Japanese technology in containing nuclear contaminants. Secondly, it made the readers to believe such science is capable to happen when it couldn't. Finally, the image of BBC was tarnished by bringing out unproven or unsolicited news to the society. The relevancy of this project has proven vital to be executed due to the negative impact it holds towards the society. This project also has a very high relevancy to the academic world due to the fact that this project looks up to the validity of all knowledge as vital to the learning curve of every human. Believing 'facts' from hoaxes will only bring injustice to the growth of that particular victim which will make it harder for them to accept the truth after exposing them to the real truth [4].

1.5 Project Feasibility

The feasibility of this project is depending on certain factors. These factors are as followed:

- Time: Sufficient for implementation (8 months altogether)
- Cost: Most of the hardware and software involved for research are free in particular
- Source of Reference: Although there are limited sources available on this subject but, this subject can refer to certain other topics such as text mining and email spams as point of references
- Settings: the wage of this research does not need any sort of specific setting to be carried out. It can be carried out using our own personal email with full connection to the internet
- Rate of Success: A study in 2009 [1] has proven that this research can be fruitful and that it is feasible for this research to be carried out

Based on the factors above, it is simply feasible for this project to be carried out as it has high relevancy upon its completion. Time, cost and setting factors show positive results and the only thing left is to implement this research paper based on its proposed design.

1.6 Scope of Study, Context and Data Source

Scope of study will aim mainly on email hoaxes. This can be traced using personal email to gather these hoaxes as well as a test subject. Moreover, the context of this project will focus mainly on minor type of hoaxes just as the project scope. The context of this project will aim mainly on hoaxes, spams and text mining concepts. This project will be carried out with aim to prove the ability of the system to detect hoaxes as well as its ability to update and train the database of the new hoaxes presented in the email. Algorithms will be determined based on the compatibility of codes to the program. Source of data can be gained through journals researched papers and archive of hoaxes can be found in hoax-slayerz.com as a point of reference for the database.

1.7 Time Feasibility

Project completion date is estimated to be somewhere in August 2011. Therefore, there is about five months left upon completion starting from the project initiation on the 2nd week of February. Project will be segmented according to the research methodology used. A Gantt char has been attached at the appendices section for easy referencing

CHAPTER 2

LITERATURE REVIEW AND THEORY

2. Literature Review and Theory

2.1 Knowledge about the Area Concerned and Existing System

Automatic Hoax Detection System is a new research finding [1]. It's a system in which was successfully implemented in 2002 [2], 2004 [3] and finally the latest finding in 2009 [1]. This detection system holds the ultimate purpose in creating high awareness among email users on the existence of hoax. This is due to the fact that the legit knowledge is important in every human's growth and misleading information can only distort a human's growth.

Hoax can be considered as a virus [4]. This virus hungers on human's weak perception which can be easily tricked. Human's perceptions have a very easy approach in believing into something as long as it is seen to be true, rational, heard to be true or experienced it to be true [4]. Moreover, hoax has been evolving in a way that most of its components have been modified to ensure that humans can be tricked by its content. This is due to the fact that hoaxes nowadays are often presented with certain percentage of statistics, convincing by experience and additional images to convince the victim of forwarded hoax [12]. Hoax is getting better at it that even anti-spams cannot detect these hoaxes [5]. Although this hoaxes might not seem malicious to the general's point of view, but due to the fact that it is misleading may be costly to the society if it's taken seriously. It can be hard to convince of the truth once a person believes something to be true [4]. It is even worse when the effect of certain hoaxes can jeopardize certain party's credibility and image which may cause losses to that particular party. Worst, hoaxes today have the ability of phishing in which it is able to gather personal data of its victim by convincing them that that email requires their personal data in an instance [12].

Current existing research [1] has shown that it is possible to automate a hoax detection system. It is being done using a mining tool called the n-grams to extract the email and make comparison with the database. According to [1], it is believed that n-grams have sped up the speed of processing in which [2] and [3] are unable to achieve. However, the drawback of [1] is due to the fact that it is unable to update its database with the latest hoaxes found. Another drawback of this system is that the system is unable to train the hoax identifier on the latest hoaxes found based on the characteristics of a hoax. According to [12], it is possible for hoaxes to be identified. Hoaxes have certain characteristics in which most of its content can be found bolded, underlined, a lot of exclamation marks and often use words which require instant decisions to be made. Latest architecture used in [1] is as followed:

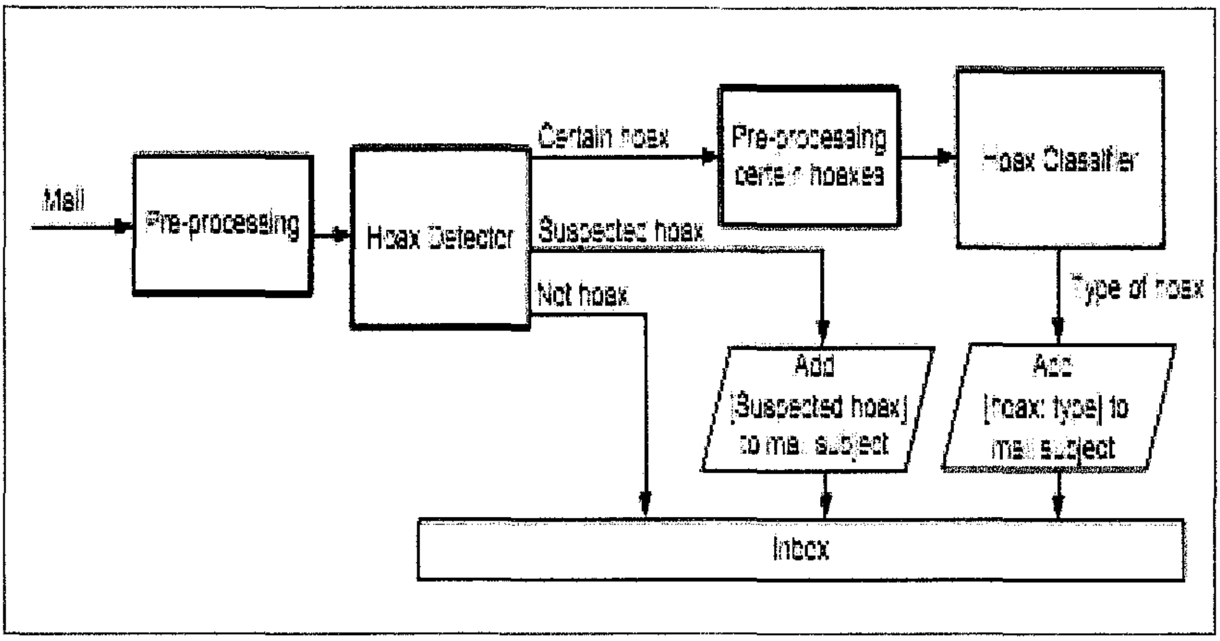

Mail
Pre-processing
Hoax Detector
Certain hoax
Suspected hoax
Not hoax
Pre-processing certain hoaxes
Hoax Classifier
Type of hoax
Add [Suspected hoax] to mail subject
Add [hoax: type] to mail subject
Inbox

Figure 2.1:1 Latest Automatic Hoax Detection System Design

Today's hoax detection system is not commercialized as good as spam detectors [1] and [5]. Spam detectors and spam filters are similar to one another [6]. However, the concept of spam is pretty much simple and easy to understand. If the spam is not traceable, the spam is an effective spam while if the spam is recognizable, the spam filter is definitely an effective spam filter [6]. Looking at the basics of a spam filter, it has a lot of resemblance to a hoax detection system. Perhaps, the only differences in these two subjects are the signatures of the targeted filter subjects, the definition of their own and also the method of extraction. Other than that, the design and architecture as well as the concept can be used to assist this research [1], [2], [3] and [5]. From the figure presented above, the design and architecture of a spam is almost similar to a spam filtering architecture [7] which is presented as followed:

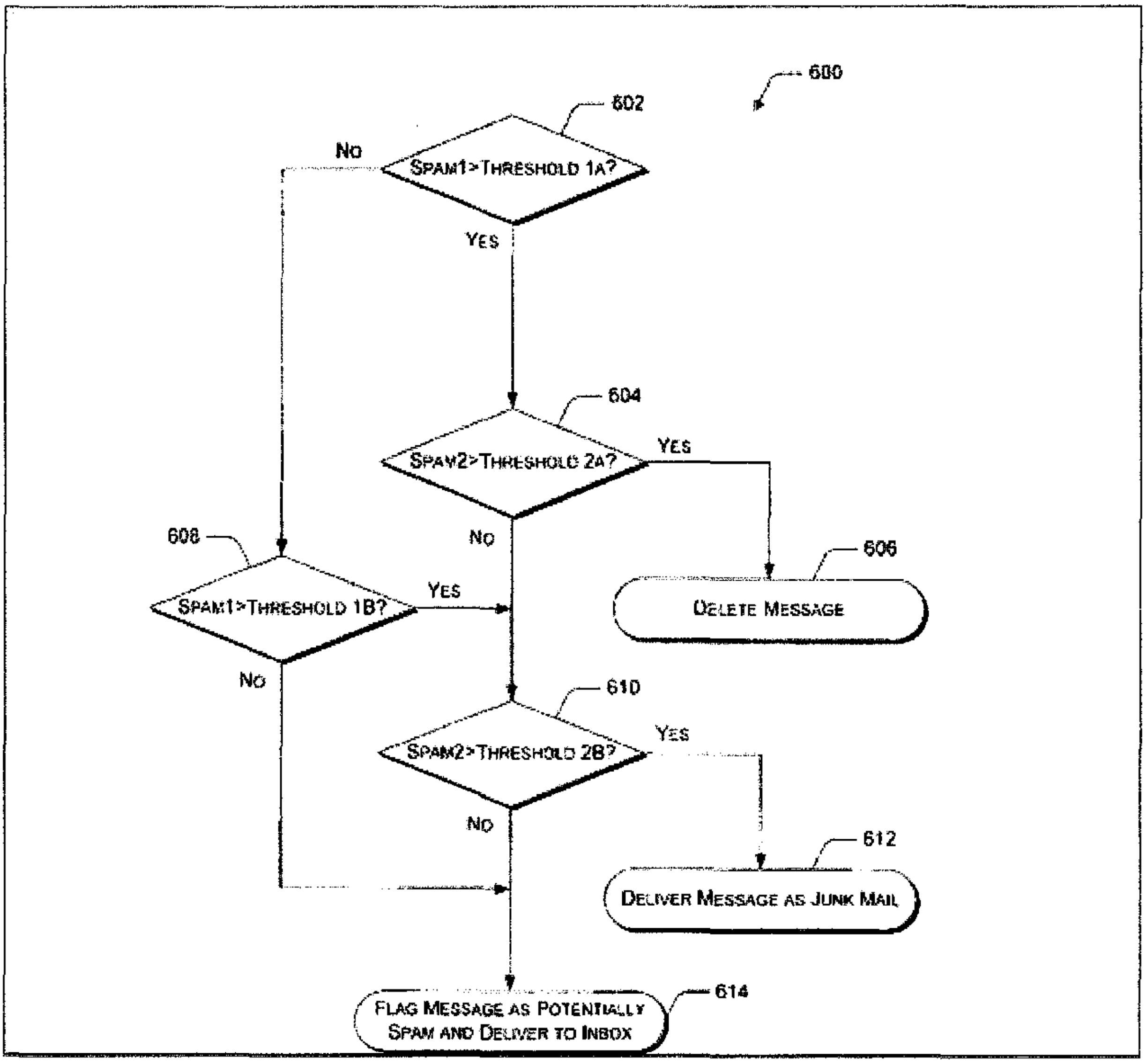


Figure 2.1:2 General Spam Filter Design

Do note that these similarities are almost the same. Another similar pattern as how hoaxes and spam detect the content of that email is via comparison to a database of hoaxes or spams. It is also important to note the process of approving one email either to be a spam or not [1] and [5]:

- Compare email to spam database

- If spam database contains the spam which was detected in the email, flag it as spam and remove them to the junk folder automatically (used by all email service provider today)
- Un-flagged spams can either be flag as neutral (safe) or suspicious depending on the probabilities of the spam and presented to the user to decide on which email should they remove as spam manually later on

Suggestions on improving the latest design have been taken into consideration after the first VIVA. In short, the improvements required are as followed:

- Update database and train the system to learn the characteristics of a hoax
- Ensure that the latest design will ensure that confirmed hoaxes will be shifted to a special hoax folder. Deletion of hoaxes is still up to the user
- Probability of that email of being a hoax should be presented in every email in order to create higher awareness
- Method should be considered. Instead of n-grams, try out other methods to see the impact of the hoax detector to the system

Understanding these facts and the relationship between hoaxes and spams, we are able to identify several methods which can be used as a reference to this project. Some of those methods under surveillance are the Naive Bayes method, Cluster method, Cross-regulation method, and n-grams. Naive Bayes Classifiers method are among the top favourite methods used in today's spam filtering. According to [8], Spam filtering is suitable with a machine learning classification such as the Naive Bayes (NB) and NB has been proven to cope well with this task despite of its simplicity. The main factor as to why this choice of classifier has been the best choice for [8] is due to the fact that the NB forms were showing promising results in handling spams and the best part about this method is that it specifically targets text-based spams in which is very similar to this research's context

On the other hand, the Cluster-based approach is another interesting approach for spam filtering. Understanding the concept of spam thoroughly, the cluster based approach in handling concept drift and skewed class distribution fairly well. Often spams are being distributed more than the ordinary legitimate mails in which it tackles the skewed class distribution concept while the concept drift concept is based on the user's preferences which may change over time or spam topic may vary according to fashionable trends [9]. Below is a sample of cluster-based method design for spam filtering [10]:

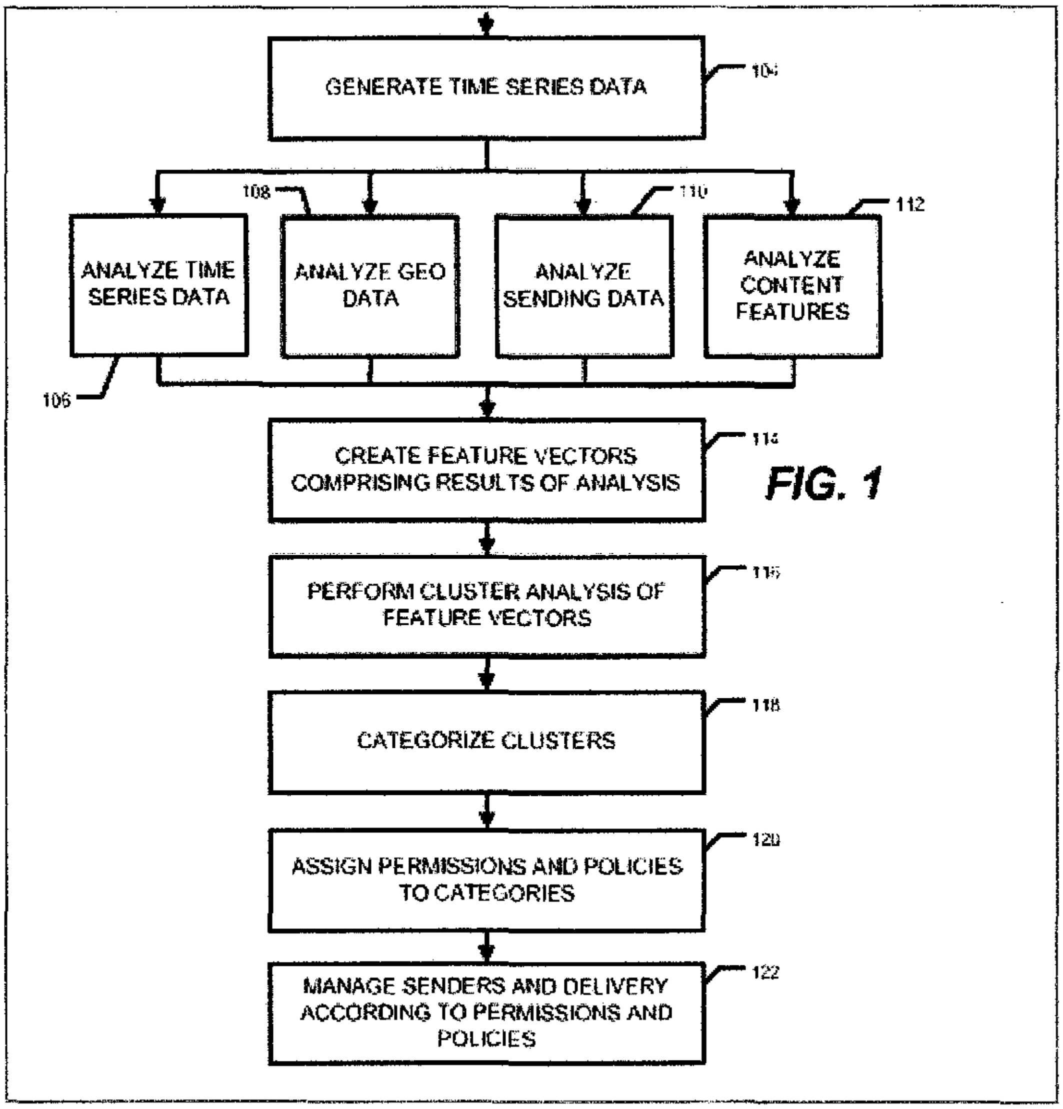

GENERATE TIME SERIES DATA
104
108
110
112
ANALYZE TIME SERIES DATA
ANALYZE GEO DATA
ANALYZE SENDING DATA
ANALYZE CONTENT FEATURES
106
CREATE FEATURE VECTORS COMPRISING RESULTS OF ANALYSIS
114
FIG. 1
PERFORM CLUSTER ANALYSIS OF FEATURE VECTORS
116
CATEGORIZE CLUSTERS
118
ASSIGN PERMISSIONS AND POLICIES TO CATEGORIES
120
MANAGE SENDERS AND DELIVERY ACCORDING TO PERMISSIONS AND POLICIES
122

Figure 2.1:3 General Cluster-Based Method Flow Chart

Based on [11], Cross-regulation method is a new found method for spam filtering. It was introduced with the aim to further detect the spams based on an immune system concept. So far, the research has proven positive result but have very little similarities to hoax detection system based on its architecture. Even so, this thesis has proven not to be best for text comparison requirement of the hoax detection system and indeed will not be part of this research's consideration to apply Cross-regulation method as part of the extraction method.

Positively, after looking through [9], text mining has the ability to tackle key phrases as part of its mined texts. Researches on text mining [13], [14] and [15] show promising effort for comparison between texts. The text mining concept is a data mining tool that is able to extract text based contents on the desired outcome. However, there are several limitations to this concept and the processing method might be a little slower. Regardless of this limitation, text mining is still one of the concepts which we will consider further later in this research.

2.2 Comparison of Works – Similarities and Differences of Existing System

It is very critical to note that there have NOT been many works able commercially for this detection system. Therefore, it is highly recommended that comparisons should be done based on the two research paper which is valid closely to this topic. Below denotes all of the similarities and differences in an existing system.

Criteria	2002 Hoax Detector [2]	Croatian CERT (2004) [3]	2009 Automatic Hoax Detection System [1]
Hoax Detection	Yes. Automated but not really intelligent. Expand to anti-virus software	Yes. Automated and rather intelligent (embed with a database)	Yes. Both automated and intelligent system integrated together
Target	More to Organizational level	Personal Based	Personal Based
Detection	Server based	Server Based	Server based
Form	Text	Text	Text
Approaches / Data mining Concepts highlighted	Not available	Levensthein + Fuzzy Logic + modified nearest neighbour algorithm	n-grams
Result	Able to detect but at minimal effect	Able to detect with probabilities analyzed	Able to detect with probabilities analyzed
Time Require to process	Not Available	Slower with tons of data to process	Light and faster (based on few samplings)
Claims	Time Require to Process made claim and have no specific results to prove on its similarity of comparison thus far		

Gap	It requires an intelligent mechanism to read and compare hoaxes at personal level	Updates require a lot of changes to the database. Unable to self-calculate the characteristics of new hoaxes which is not stored in the database	Updates require a lot of changes to the database. Unable to self-calculate the characteristics of new hoaxes which is not stored in the database

Table 2.2-1 Comparison of Previous Works with Relate to Hoax Detection System

These are the similarities and the dissimilarities that the existing hoax detection system have. However, we will refer to [1], [2] and [3] closely for further analysis.

In general, hoax detector's research began actively in the year 2002 [2]. This model began its research after noticing that hoaxes are indeed a disturbing matter in the society. According to [2], knowledge is essential to human's growth and that misleading information can be costly. Having said so, they began by using the rule-based method as a start for this particular research. In 2002, rule-based method is one of the most leading text extraction method used for anti spams. Rule-based methods are a method in which extraction are made based on certain rules and conditions. This traditional filter method uses simple and straightforward methods which classify spams by matching certain email fields with certain key words. The failure to this method is due to the fact that the complexity of spams today has override the benefits that this method have yet to offer

In 2004 [3], the Croatian CERT took a step further by developing another version of a hoax detection system. This system has proven to be successful in Croatia due to the low internet traffic over in that nation. However, this system has shown some drawbacks in certain area mostly on the idea of using fuzzy logic and Levenhstein method. These two methods have shown promising results in detecting the hoaxes but the speed of processing the hoaxes is deemed to be slow added with heavy traffic of internet and with large volumes of data. Furthermore, the database cannot update on its own (No self-calculate methods) and requires further study which was taken up by the 2009's study [1]

In 2009 [1], the study for an automatic hoax detection system became much more feasible than before. Instead of using rule-based as well as fuzzy logics, this research uses n-grams as text extraction method. N-grams is one of the data mining tool which is a predicting method used to predict the words or the text within the email. N-grams allow the system to predict the subsequent words in the analyzed email in which it can process the words faster (assumed and deemed to be by the inventors) However, given the study for n-grams in spams [16] and [17], the complexity of the probability in executing this concept in computer codes is pretty complicated for hoax detection system. Nevertheless, it has proven successful and will be considered for further development of this research.

2.3 Gap Analysis

Based on the analysis made in Section 2.1 and 2.2, it is essential to say that the point of reference for this subject specifically is very limited. In order for this project to be successful, we have taken the privilege in making other sort of close-related subjects as part of our point of reference in which to be specific are Spams and Text Mining tools. According to [5] and [6], spams filters' patterns are of the similar patter to the hoaxes detection systems created over the years with different definition. Spams can be traced based on the sender's address of that particular email [5] and [6] while hoaxes need to be addressed by its content [1], [2] and [3]. As obvious as this may be, the close relationship of these two subjects can be traced down to its methods. Note that the methods used in [1], [2] and [3] have been used by spam filters before this [7], [8], [9] and [16].

Fortunately, the main gap for this research isn't exactly on its method but rather on the database itself. Thus far, there are no hoax detection systems that are able to self update its database on the existence of new hoaxes. This means that the system is unable to recognize newly created hoaxes. This is the biggest problem encountered in this research. Having this pointed out, the system will enable this feature by identifying the characteristics of hoaxes based on [12].

Another gap that we can see is due to the fact that the methods used in all of the researches related to hoaxes are not processing text that well except for n-grams method. However, after reviewing other text mining tools, it is possible for us to assume that there might be another text mining tool which is capable of extracting texts from the email for comparison with the database. The process of this whole extraction is called text pre-processing [1].

Another problem gap that can be closed down for this research is on how the data is being represented to the users. Hoax detectors previously did not present the probabilities of hoaxes to its users. Again, the objective of this research is to increase awareness of the existence of hoaxes and to avoid unnecessary misleading information and understanding among the society which can lead to other party's losses either unintentionally or intentionally. Having said so, most emails today are presented in a simple form whereby it does not even create the awareness on spams. Therefore, by presenting the data in a form of percentage right beside each of the emails will definitely increase the awareness of the users on the existence of the hoaxes. Furthermore, this can educate the users on what is right and what is not. Therefore, it is essential for this gap to be closed down which can help users to educate themselves on the validity of each and every email they receive.

Finally, the gap which needs to be closed down the most is the need to identify the characteristics of a hoax and how will the hoax detector act in order to classify that email as hoax, suspected as hoax and neutral. Given out these concerns, it is important for the proposed hoax detection system to be able to cater to these gaps and lessen the gaps as best as it could.

CHAPTER 3

METHODOLOGY / PROJECT WORK

3. Methodology / Project Work

3.1 Research methodology

The preferred research methodology as for this project will be the usage of Rapid Application Development method also known as (RAD). This project will require a lot of tries and errors therefore making throwaway prototyping as the best mean of research methodology for this research. This is due to the nature of this methodology for such a short period of time. Below is the research methodology's diagram for further understanding:

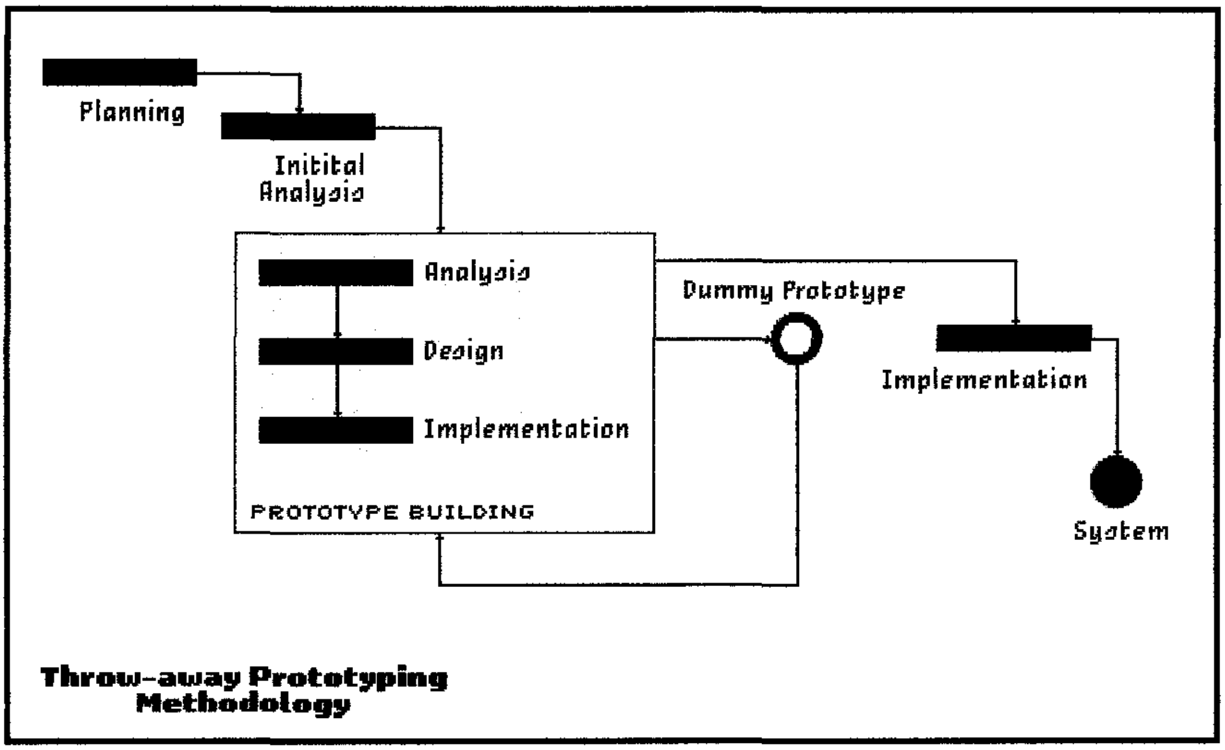


Figure 3.1:1 Research Methodology: Throwaway Prototyping

This prototype was chosen best due to the need for rapid program development. This project has about 31 weeks for total completion with new-found algorithms and theories to be explored. The main projected activities that will be carried out require a lot of trial and errors. Therefore having a throwaway prototype will allow rapid changes to be made to the whole system before it can be finally implemented successfully into a working system.

In the planning stage, this project aims to understand the system in general terms. Most of the analyses involved in planning are in getting the timeline right as well as to compare the current work to the previous work. Planning is how the project will manipulate its limited resources to full use upon completion of the project [18].

Initial analysis will see further as to the items necessary for the implantation of the system in general. Note that currently, there will be a certain text pre-processing models which will be tested out in order to achieve the objective. Hence, in the initial analysis phase, the project will lay down its entire possible completion route and will be documented for further references. Based on this project, the initial analysis phase and planning have both been completed based on the Gantt chart attached in the appendix section. Furthermore, FYP1 only requires for the project to be completed up until this phase. The prototyping phase will be conducted intensely in FYP2 later on in May.

As for the dummy prototyping phase, this project will try out all of the methods listed down in the initial analysis phase and see whether such method can be successfully implemented for this project or not. Trial and error is the best term to describe this phase. Since there are many methods available and analyzed thus far in section 2, it is possible that this phase will take a lot of time before completion. However, the finalized prototype will be the one that will be forwarded to the implementation phase later on.

During the final stage of this methodology, project will modify the system and tune it to its best condition before presenting it as a system to the shareholders later on in FYP2. However, due to the time and financial constraints presented, the quality of the system might not be at its best but certainly possible for completion.

3.2 Projected Activities

There are 4 phases of projected activities based on the methodology chose. Below are the 4 mentioned phases:

- Phase 1: Extensive research for planning stage
- Phase 2: Data comparison of works for initial analysis and possible route of completion
- Phase 3: Design and Implement for dummy prototyping
- Phase 4: Finalized system will be tuned up, implemented and presented to stakeholders

3.2.1 Phase1: Extensive Research for Planning Stage

This is the phase whereby extensive research is required in the planning stage. Research requires a lot of research based on past successful journals. Based on the extensive research done within the first 6 weeks of this project, the automatic hoax detection system is one of the most inactive research topics presented thus far. There are only a few point of references and often relates to text mining. Extensive research is based on general perception of what the system should be at the end of the research period

3.2.2 Phase 2: Data Comparison of Works for Initial Analysis

Second phase is the phase whereby all of the gathered data are thoroughly analysed to see which route of completion this project should foresee. It is the initial analysis of how this project should move forward and understand the gaps that this research should cover. Initial analysis lay down all of the possibilities that this project can take in order to complete a successful final system at the end of the system. The main concerns of the system as of now are:

- To enable the system to self learn the characteristics of a hoax itself
- To enable the system to update the database on its own based on the newly encountered hoaxes
- Present the emails together with the probabilities of them being hoaxes
- To transfer confirmed hoaxes into a special folder (hoax folder)
- To increase the speed of processing the text from the email for comparison with the database

Initial understanding of the whole system should be thorough by the time the project reaches this phase and that it should have at least a general idea of how the design of the proposed system should look like

3.2.3 Phase 3: Design and Implement for Dummy Prototyping

This phase will prepare the project for the creation of dummy prototyping. As describe in section 2, there are a lot of methods in which the text can be pre-processed for comparison. However, we have limited the scope of this research to only try out text mining tool to assist in the completion of this project. Furthermore, the probability calculation will change over time based on the outcome of each and every dummy prototype created. Successful dummy prototype will be used later in the implementation phase

3.2.4 Phase 4: Finalized system will be tuned up, implemented and presented to shareholders

Finalized system will be chosen from the best series of dummy prototyping did during the third phase. The finalized system will be chosen based on its success in closing the gaps presented in section 2. System must be implemented successfully without errors. After noting down the successful system, this system will be tuned-up to its best condition in order for this system to be presented to the stakeholders.

3.3 Key Milestones and Gantt chart

3.3.1 FYP I

FYP 1 will commence in January 11 Semester while FYP 2 will follow in May 11 Semester. By the end of January 11 Semester, these submissions shall be met:

- Proposal (1st Week)
- Extended Proposal (6th Week)
- Proposal Defence & Progress Evaluation (9th Week)
- Interim Report (13th Week)

Further documentations for valid referencing will be attached under the appendix section found in section 7.

3.3.2 FYP II

FYP 2 will commence from May 11 Semester onwards. Below are the submissions that shall be met by the project throughout the timeline given:

- Progress Report (7th Week)
- Pre SEDEX (11th Week)
- Dissertation Submission (12th Week)
- Viva (13th Week)
- Final Dissertation + Technical Report (14th Week)

Further referencing can be done via the Gantt chart attached in the Appendices section later.

3.4 Tools or Equipments Required

Based on the literature reviews made from the papers referenced and read, most of the software required for this project would be:

- Java via Netbeans IDE 7.0
- Glassfish Apache Server
- Hoax database from www.hoax-slayers.com

Note that hoax-slayers.com was chosen due to its credibility of being the only website chosen by PANDORA (National Library of Australia) as their source of reference for hoaxes and internet scams.

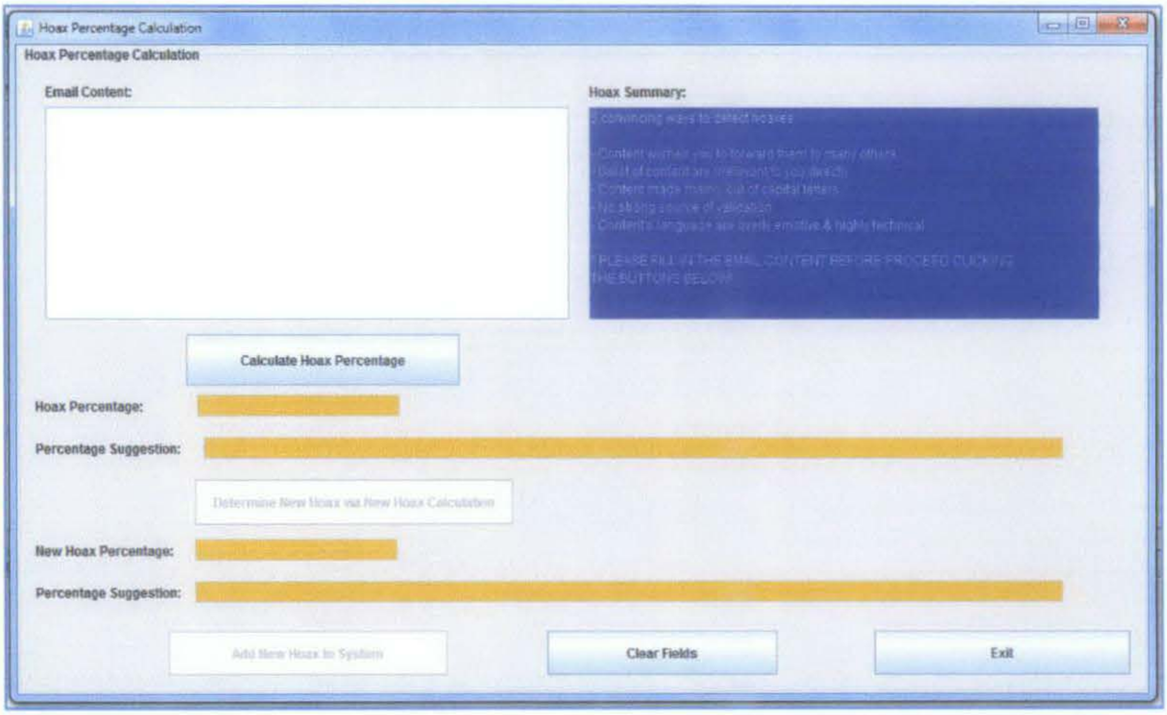
CHAPTER 4

RESULTS AND DISCUSSION

4. Results and Discussion

4.1 Snapshots

Thus far, the system is 85% completed. This system is able to commence at least 2 out of 3 functions and possibly be able to run all methods by the end of the 13th week before the presentation of VIVA. Below are some of the snapshots taken from the 85% completed version of the system:

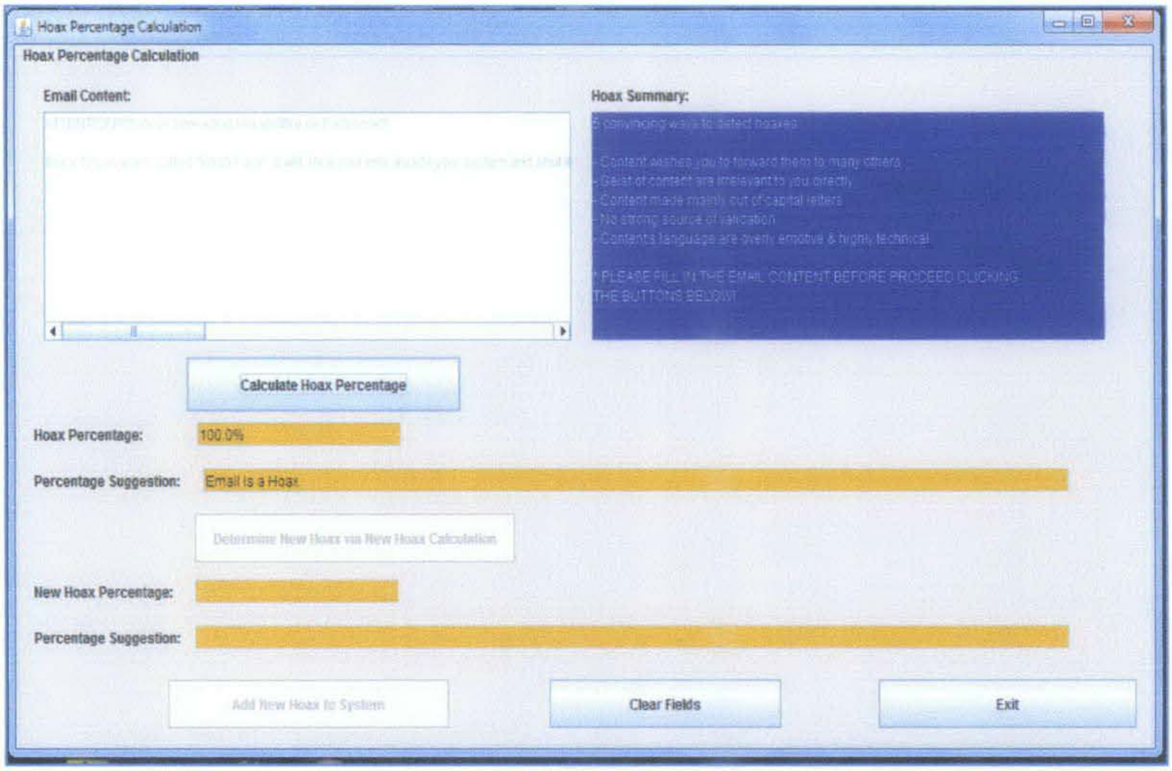


5 convincing ways to detect hoaxes

- Content wants you to forward them to many others
- Bulk of content are irrelevant to you directly
- Content made mainly out of capital letters
- No strong source of validation
- Content's language are overly emotive & highly technical

*PLEASE FILL IN THE EMAIL CONTENT BEFORE PROCEED CLICKING THE BUTTONS BELOW

Figure 4.1:1 GUI Snapshot (Original)



5 convincing ways to detect hoaxes

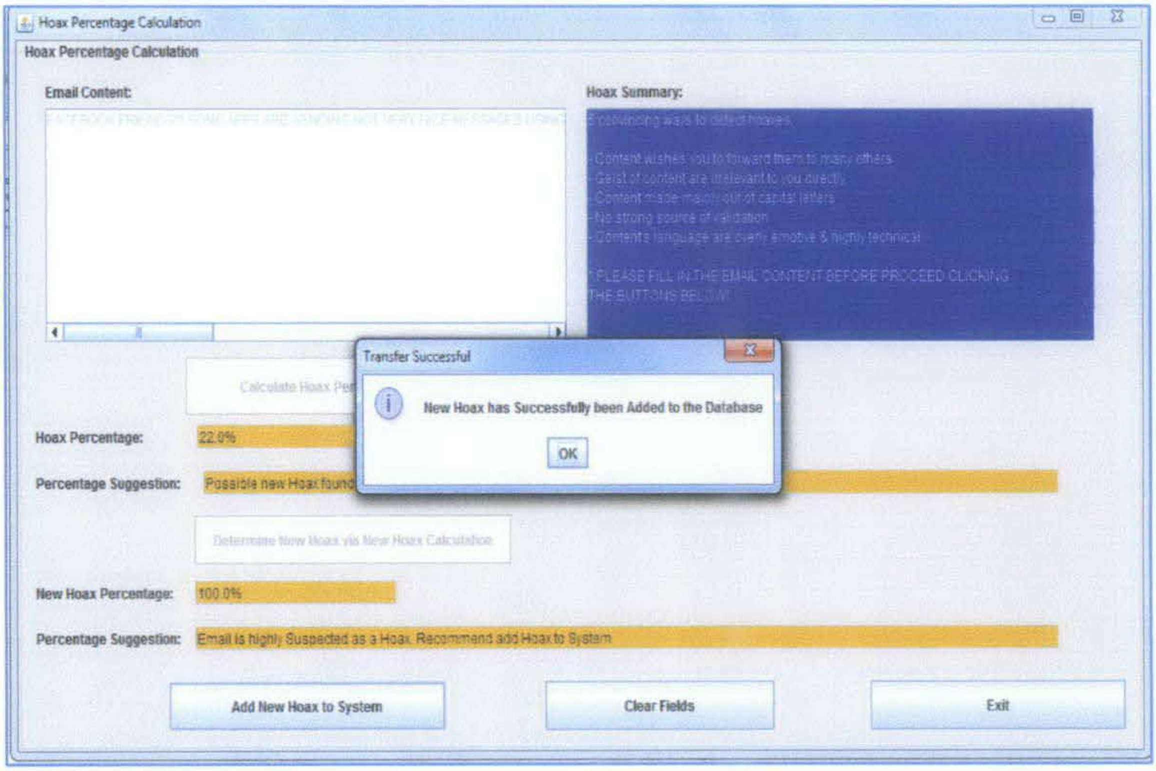
- Content wishes you to forward them to many others
- Geist of content are irrelevant to you directly
- Content made mainly out of capital letters
- No strong source of validation
- Contents language are overly emotive & highly technical

* PLEASE FILL IN THE EMAIL CONTENT BEFORE PROCEED CLICKING THE BUTTONS BELOW!

Hoax Percentage: 100.0%

Percentage Suggestion: Email is a Hoax

Figure 4.1:2 GUI Snapshot (Calculate Hoax Percentage)



5 convincing ways to detect hoaxes

- Content wishes you to forward them to many others
- Geist of content are irrelevant to you directly
- Content made mainly out of capital letters
- No strong source of validation
- Contents language are overly emotive & highly technical

* PLEASE FILL IN THE EMAIL CONTENT BEFORE PROCEED CLICKING THE BUTTONS BELOW!

New Hoax has Successfully been Added to the Database

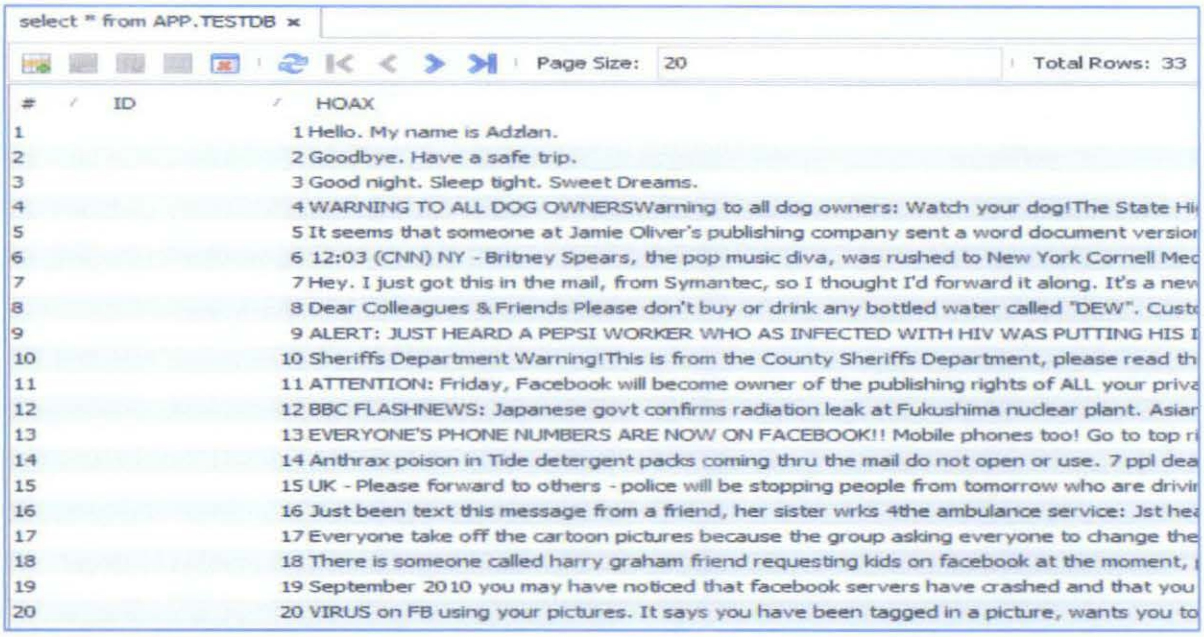
Hoax Percentage: 22.0%

Percentage Suggestion: Possible new Hoax found

New Hoax Percentage: 100.0%

Percentage Suggestion: Email is highly Suspected as a Hoax. Recommend add Hoax to System

Figure 4.1:3 GUI Snapshot (Transfer Successful)



select * from APP.TESTDB ×

Page Size: 20 Total Rows: 33

#	ID	HOAX
1	1	Hello. My name is Adzlan.
2	2	Goodbye. Have a safe trip.
3	3	Good night. Sleep tight. Sweet Dreams.
4	4	WARNING TO ALL DOG OWNERSWarning to all dog owners: Watch your dog!The State Hi
5	5	It seems that someone at Jamie Oliver's publishing company sent a word document versio
6	6	12:03 (CNN) NY - Britney Spears, the pop music diva, was rushed to New York Cornell Me
7	7	Hey. I just got this in the mail, from Symantec, so I thought I'd forward it along. It's a ne
8	8	Dear Colleagues & Friends Please don't buy or drink any bottled water called "DEW". Cust
9	9	ALERT: JUST HEARD A PEPSI WORKER WHO AS INFECTED WITH HIV WAS PUTTING HIS I
10	10	Sheriffs Department Warning!This is from the County Sheriffs Department, please read th
11	11	ATTENTION: Friday, Facebook will become owner of the publishing rights of ALL your priva
12	12	BBC FLASHNEWS: Japanese govt confirms radiation leak at Fukushima nuclear plant. Asia
13	13	EVERYONE'S PHONE NUMBERS ARE NOW ON FACEBOOK!! Mobile phones too! Go to top r
14	14	Anthrax poison in Tide detergent packs coming thru the mail do not open or use. 7 ppl dea
15	15	UK - Please forward to others - police will be stopping people from tomorrow who are drivi
16	16	Just been text this message from a friend, her sister wrks 4the ambulance service: Jst he
17	17	Everyone take off the cartoon pictures because the group asking everyone to change the
18	18	There is someone called harry graham friend requesting kids on facebook at the moment,
19	19	September 2010 you may have noticed that facebook servers have crashed and that you
20	20	VIRUS on FB using your pictures. It says you have been tagged in a picture, wants you to

Figure 4.1:4 Database Snapshot

4.2 Proposed Design

After taking a lot of consideration of how the new Automatic Hoax Detection System should be, this design would be the best point of reference for the following explanations:

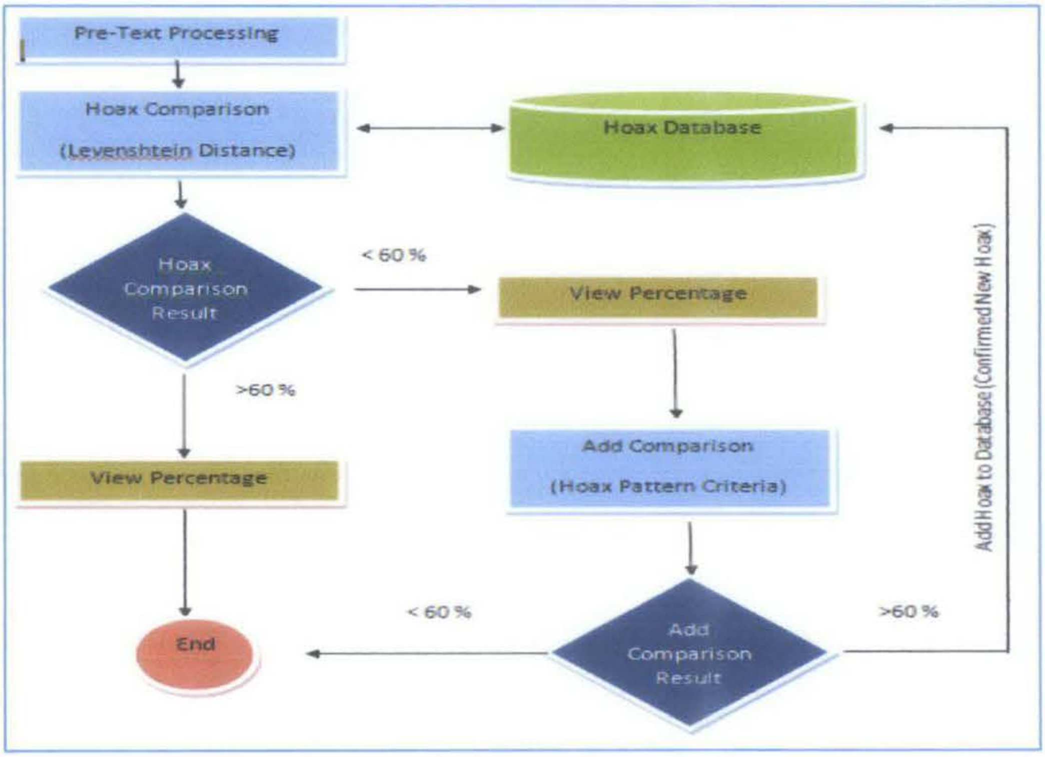


Figure 4.2:1 Proposed Design

This design was created to cater to all of the gaps presented in the gap analysis. Using this design, the database will automatically update the database if a newly found suspected hoax were found during the detection process. Flow will begin with text pre-processing process in which the system will extract the content of the text-based email and process it best for digestion in this proposed system. Text pre-processing as proposed in this system will extract the content of the email in belief that all hoaxes masquerades themselves through forwarded email methods in which spam filters that were placed in email service providers mostly failed to analyze. It is also important to note the fact that SQL language has certain special characters in which cannot be read perfectly by the system. Hence, this system will need to ensure such problem will not exist by performing this text pre-processing method. Thus far, this process has yet subject to completion

Next, the processed text will be forwarded to the hoax detector for comparison with the database. After analysing most of the text mining tools from [1], [13], [14] and [15], and although it is mentioned that the speed via n-grams can improve the lightweight of the system, this proposed system will still refer to Levenshtein Distance's method in belief that it is still capable in handling the variables proposed. Furthermore, previous 2 papers presented have failed to provide specific results to their calculations in terms of speed for best comparison. It is also important to detail out the fact that completed system thus far has been tested with a strong valid sample of 30 data and have believed that it is able to provide results swiftly in less than 5 seconds. Furthermore, Google has it that Levenshtein Distance is indeed one of the best methods used to compare 2 strings by showing these results at the top of their hit lists and also by good discussions via online forums. It is basic yet sufficient to cater the current system's need

4.2.1 Method 1: Levenshtein Distance Calculation (100%)

Levensthein Distance calculation will provide the system with the strings comparison result needed between the detected string as well as the one stored in the database. However, it is also important to note the validity of this system by accepting these results measurement due to the facts that:

- Hoaxes up until today have no specific patterns which is very obvious to calculations

- Hoaxes have no scale of measurement until facts are supported with thorough discoveries and investigations (in which will be done by the hoax detector bodies)
- Only advised patterns can be seen thus far. In which it has only 60% probability to hit such patterns.

It is important to note down the outcomes of the hoax detector. There will be 2 outcomes presented from the above figure. Below are the explanations of the 2 outcomes presented for Levenshtein Distance Calculation method:

- Probability at above 60% = Email is a Hoax
    - This assumption is made believing that scores of 50% holds 50 – 50 chances. Scoring 60% creates higher chances for the tested text to be a hoax in similarity to the ones presented in the database. No strong sense of measurement to admit validity as a strong fact but merely an opinion of a person's thought based on logic and rationale
- Probability below and equivalent to 60% = Email can be suspected as Hoax.
    - Thus far, presented hoaxes or tested hoaxes have shown that the email contents which have scored 60% will not be treated as hoaxes immediately. They will next be subjected to the following calculation = New Hoax Calculation to confirm the authenticity of the hoax to be deemed as one based on the patterns analyzed.

4.2.2 New Hoax Calculation (200%)

New Hoax Calculations (Add Comparison) are a calculation method in which it will take the text and compare them with the user-defined calculations based on the perceived understanding of hoaxes. Bear in mind that hoaxes will undergo two main tests in this method. They are:

- Uppercase letter comparison (100%)
    - Calculation for uppercase characters above the total number of characters in the email.
- Frequently used and found words comparison (100%)
    - Hoaxes are claimed to have similar words being used in the content of their emails. For instance: "PLEASE FORWARD THIS MESSAGE /

SEND THIS MESSAGE!!!! / URGENT!!!" Hence, by analysing these words and add them in a new database for comparison, this can create another percentage to assist the users in determining whether the hoax can be confirmed as a new hoax and can be added as a hoax entirely. Some of the samples identified frequently used by hoax are as followed:

Num	Words	Num	Words
1	URGENT	11	VIRUS
2	ATTENTION	12	FOLLOW
3	PLEASE	13	INSTRUCTIONS
4	FORWARD	14	IMMEDIATELY
5	SEND	15	MESSAGE
6	INFECTED	16	NOT
7	WORST	17	ADD
8	CLOSED	18	TELL
9	EVERYONE	19	EVERYBODY
10	LOVED	20	SPREAD

Table 4.2.2-1 Bag of Words Used in Frequent Words Formula

New Hoax Calculations (Add Comparison) have certain calculation results and the maximum calculation result up until now is at 200%. The rationale behind 200% is due to the fact that the combinations of these 2 factors will never reach 200% in all situations. Therefore, the system needs to amplify the output to ensure that the new hoax can be detected as best as it could based on the factors presented and identified .Below are the assumed results:

- Above 60% = Confirmed Hoax based on patterns.
  - These hoaxes can now be added to the database for further recognition
- Below and similar to 60% = Email is not strong enough to be deemed as hoax
  - These emails can be ignored but however, with such calculations, it will create awareness among the users on the existence of hoax automatically.

Probabilities will be displayed in the text fields presented which will be available easy for readings and understanding.

4.3 System Architecture

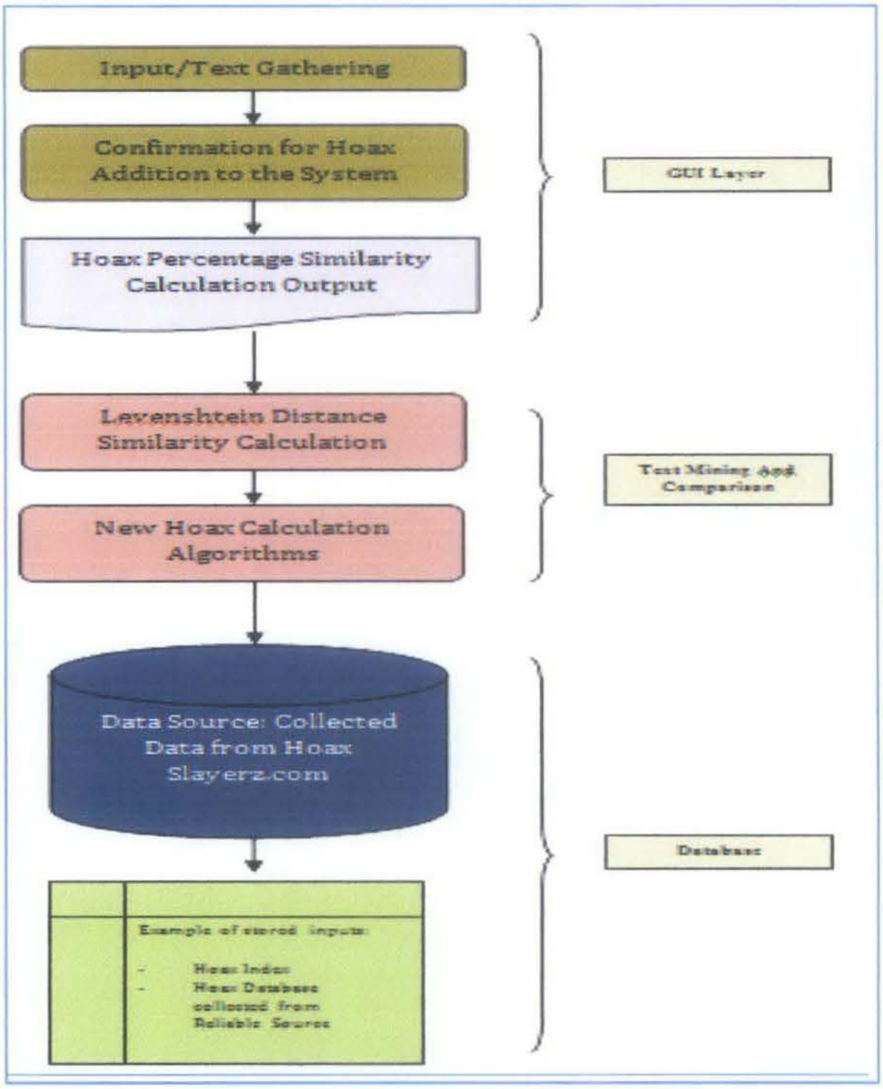

Input/Text Gathering
Confirmation for Hoax Addition to the System
Hoax Percentage Similarity Calculation Output
GUI Layer
Levenshtein Distance Similarity Calculation
New Hoax Calculation Algorithms
Text Mining And Comparison
Data Source: Collected Data from Hoax Slayerz.com
Database
Example of stored inputs:
- Hoax Index
- Hoax Database collected from Reliable Source

Figure 4.3:1 System Architecture

Above is the pictured system architecture of the proposed system. There are several layers involved and they are:

- Database Layer
  - o This Layer has 2 databases proposed. However, only one has been successfully completed thus far. The 2 databases which are supposed to be completed in this layer are the Hoax Database and Frequent Words database.
  - o Hoax Database is the database whereby the system stores the hoaxes for the Levenshtein Distance comparison. It is also the database whereby the Add Comparison calculation may add the hoax once it is

found confirmed to be one. It comprises of 2 data: Hoax and ID (primary key)

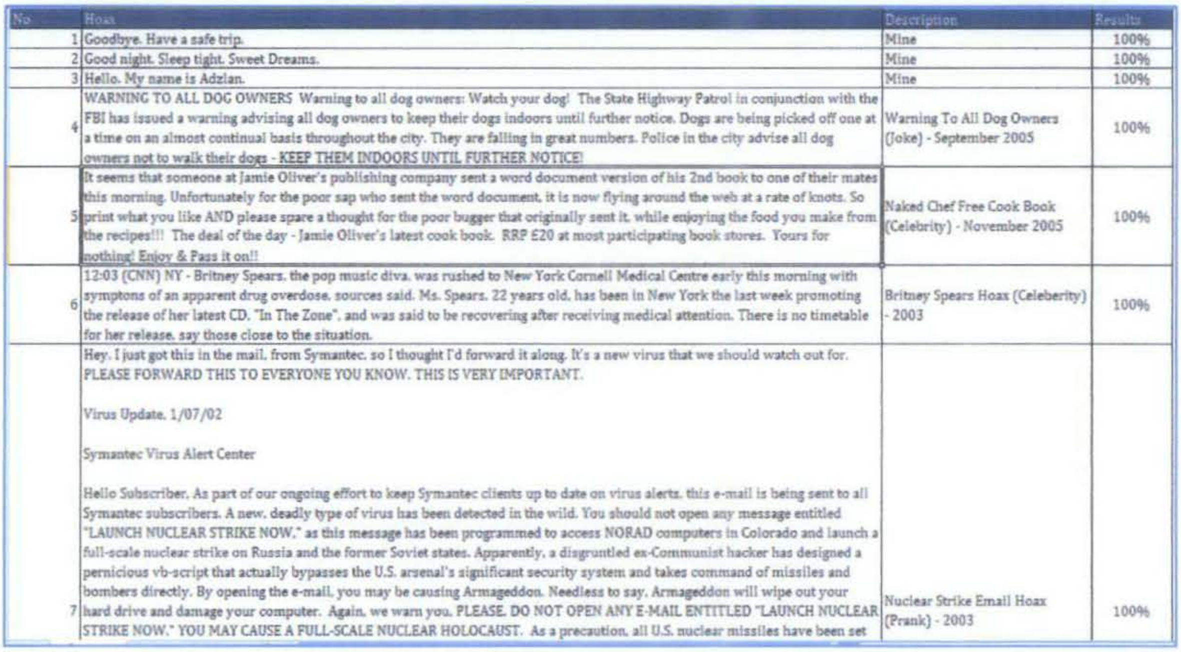
  - Frequent Word Database is the database whereby the system is able to store the frequent words found in hoaxes that can be used in the New Hoax Calculation method. This comparison is important to nominate the tested text as a hoax or not. It comprises of 2 data: Frequent Words and ID (primary key).

- Information Extraction and Comparison Layer
  - Currently, there are 2 methods in which have been successfully completed. However, there are actually 3 methods used in this system but the third method is still under construction. Therefore, only 2 have been added in the latest system architecture thus far.
  - The first method used is the Levenshtein Distance string comparison methods in which it detects the differences in distances between two strings. In other words, it detects the similarity of texts presented. The result of this calculation can be proven valid in the Results and Discussion section later on
  - The second method used is the New Hoax Calculation (Add Comparison). This method's being used once the text's being identified as suspected possible hoax. Therefore, this calculation will commence in accordance to the perceived pattern of hoaxes found from experts and logic. The result of this calculation can be proven valid later in the Results and Discussion section later on.
  - Third method proposed is the text pre-processing method. This method removes all forwarded clauses in emails and ensures that the content of the email can be preserved and used best with SQL statements to avoid such problems. These methods shall be used at both the beginning and at the end of the system to ensure that content can be preserved and be saved into the database.
- GUI Layer
  - This layer was designed in accordance to the HCI principles.
  - There is only one interface and this interface serves the purpose well

4.4 Results Discussion

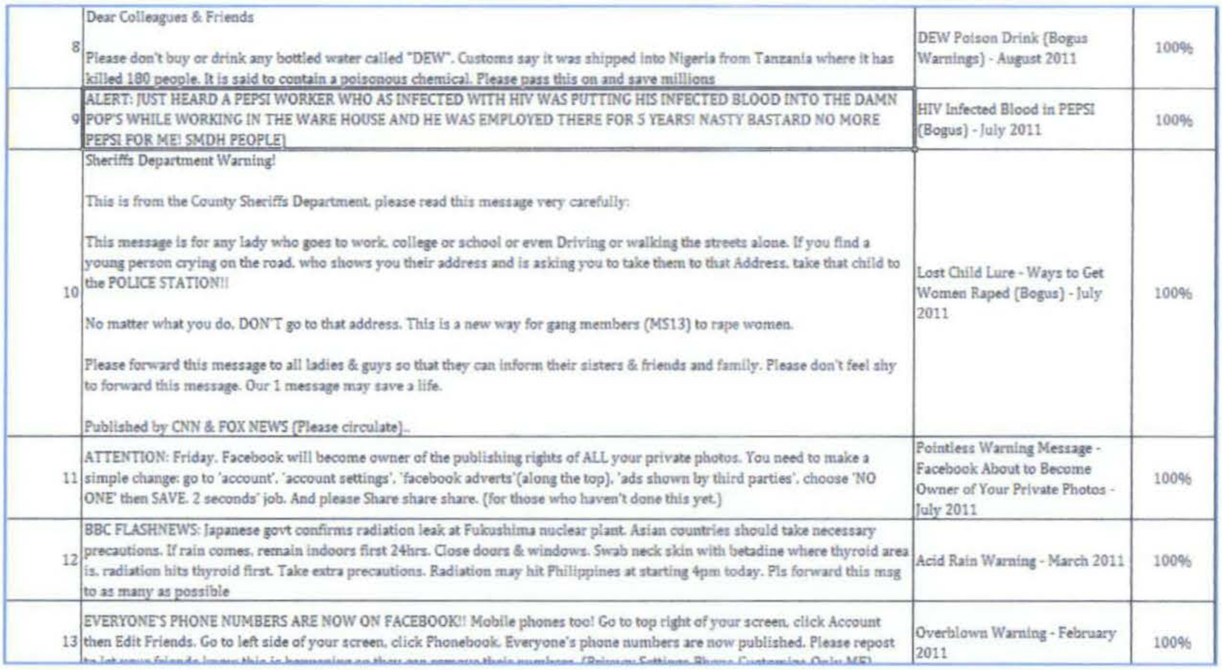
4.4.1 Sampling Results and Outcome

This system has completed 30 successful samplings thus far with additional of 3 personal sampling (33 samplings). According to the results of each of the samplings made through the Levensthein Distance, the samplings have shown positive results in approval of this system in being successful. Below are some of the snippets of the results made through the Calculations in this system:



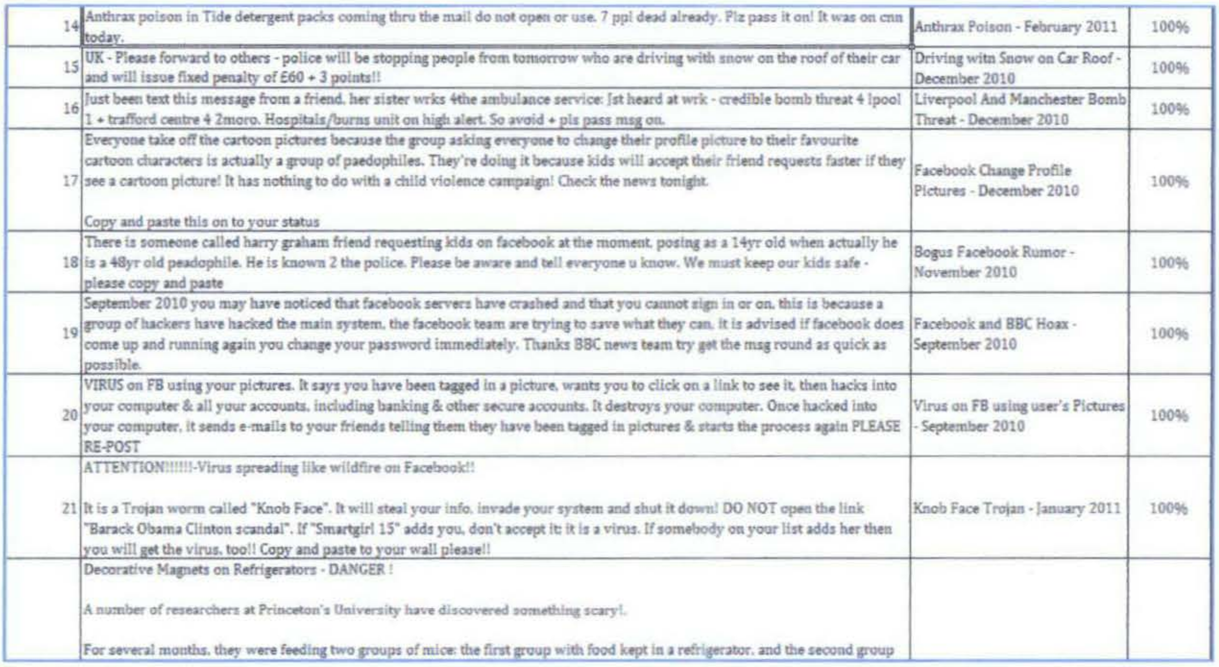
No	Hoax	Description	Results
1	Goodbye. Have a safe trip.	Mine	100%
2	Good night. Sleep tight. Sweet Dreams.	Mine	100%
3	Hello. My name is Adzlan.	Mine	100%
4	WARNING TO ALL DOG OWNERS Warning to all dog owners: Watch your dog! The State Highway Patrol in conjunction with the FBI has issued a warning advising all dog owners to keep their dogs indoors until further notice. Dogs are being picked off one at a time on an almost continual basis throughout the city. They are falling in great numbers. Police in the city advise all dog owners not to walk their dogs - KEEP THEM INDOORS UNTIL FURTHER NOTICE!	Warning To All Dog Owners (Joke) - September 2005	100%
5	It seems that someone at Jamie Oliver's publishing company sent a word document version of his 2nd book to one of their mates this morning. Unfortunately for the poor sap who sent the word document, it is now flying around the web at a rate of knots. So print what you like AND please spare a thought for the poor bugger that originally sent it, while enjoying the food you make from the recipes!!! The deal of the day - Jamie Oliver's latest cook book. RRP £20 at most participating book stores. Yours for nothing! Enjoy & Pass it on!!	Naked Chef Free Cook Book (Celebrity) - November 2005	100%
6	12:03 (CNN) NY - Britney Spears, the pop music diva, was rushed to New York Cornell Medical Centre early this morning with symptons of an apparent drug overdose, sources said. Ms. Spears, 22 years old, has been in New York the last week promoting the release of her latest CD, "In The Zone", and was said to be recovering after receiving medical attention. There is no timetable for her release, say those close to the situation.	Britney Spears Hoax (Celeberity) - 2003	100%
7	Hey, I just got this in the mail, from Symantec, so I thought I'd forward it along. It's a new virus that we should watch out for. PLEASE FORWARD THIS TO EVERYONE YOU KNOW. THIS IS VERY IMPORTANT. Virus Update, 1/07/02 Symantec Virus Alert Center Hello Subscriber, As part of our ongoing effort to keep Symantec clients up to date on virus alerts, this e-mail is being sent to all Symantec subscribers. A new, deadly type of virus has been detected in the wild. You should not open any message entitled "LAUNCH NUCLEAR STRIKE NOW," as this message has been programmed to access NORAD computers in Colorado and launch a full-scale nuclear strike on Russia and the former Soviet states. Apparently, a disgruntled ex-Communist hacker has designed a pernicious vb-script that actually bypasses the U.S. arsenal's significant security system and takes command of missiles and bombers directly. By opening the e-mail, you may be causing Armageddon. Needless to say, Armageddon will wipe out your hard drive and damage your computer. Again, we warn you. PLEASE. DO NOT OPEN ANY E-MAIL ENTITLED "LAUNCH NUCLEAR STRIKE NOW." YOU MAY CAUSE A FULL-SCALE NUCLEAR HOLOCAUST. As a precaution, all U.S. nuclear missiles have been set	Nuclear Strike Email Hoax (Prank) - 2003	100%

Table 4.4.1-1 Sampling Results 001



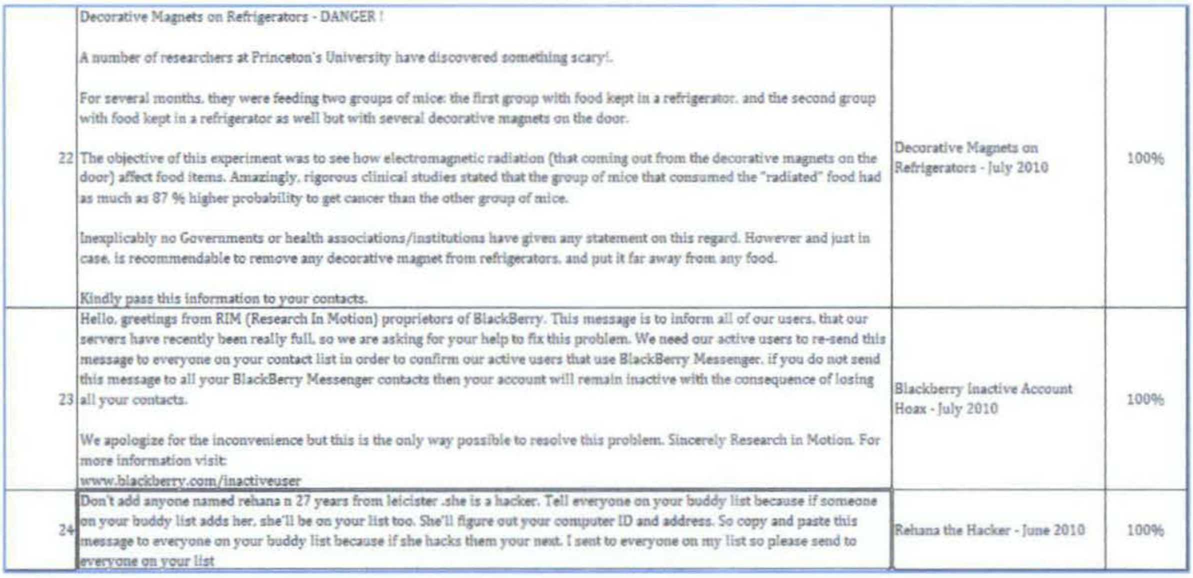
8	Dear Colleagues & Friends Please don't buy or drink any bottled water called "DEW". Customs say it was shipped into Nigeria from Tanzania where it has killed 180 people. It is said to contain a poisonous chemical. Please pass this on and save millions	DEW Poison Drink (Bogus Warnings) - August 2011	100%
9	ALERT: JUST HEARD A PEPSI WORKER WHO AS INFECTED WITH HIV WAS PUTTING HIS INFECTED BLOOD INTO THE DAMN POP'S WHILE WORKING IN THE WARE HOUSE AND HE WAS EMPLOYED THERE FOR 5 YEARS! NASTY BASTARD NO MORE PEPSI FOR ME! SMDH PEOPLE]	HIV Infected Blood in PEPSI (Bogus) - July 2011	100%
10	Sheriffs Department Warning! This is from the County Sheriffs Department. please read this message very carefully: This message is for any lady who goes to work. college or school or even Driving or walking the streets alone. If you find a young person crying on the road. who shows you their address and is asking you to take them to that Address. take that child to the POLICE STATION!! No matter what you do. DON'T go to that address. This is a new way for gang members (MS13) to rape women. Please forward this message to all ladies & guys so that they can inform their sisters & friends and family. Please don't feel shy to forward this message. Our 1 message may save a life. Published by CNN & FOX NEWS (Please circulate)..	Lost Child Lure - Ways to Get Women Raped (Bogus) - July 2011	100%
11	ATTENTION: Friday. Facebook will become owner of the publishing rights of ALL your private photos. You need to make a simple change: go to 'account'. 'account settings'. 'facebook adverts'(along the top). 'ads shown by third parties'. choose 'NO ONE' then SAVE. 2 seconds' job. And please Share share share. (for those who haven't done this yet.)	Pointless Warning Message - Facebook About to Become Owner of Your Private Photos - July 2011	100%
12	BBC FLASHNEWS: Japanese govt confirms radiation leak at Fukushima nuclear plant. Asian countries should take necessary precautions. If rain comes. remain indoors first 24hrs. Close doors & windows. Swab neck skin with betadine where thyroid area is. radiation hits thyroid first. Take extra precautions. Radiation may hit Philippines at starting 4pm today. Pls forward this msg to as many as possible	Acid Rain Warning - March 2011	100%
13	EVERYONE'S PHONE NUMBERS ARE NOW ON FACEBOOK!! Mobile phones too! Go to top right of your screen. click Account then Edit Friends. Go to left side of your screen. click Phonebook. Everyone's phone numbers are now published. Please repost	Overblown Warning - February 2011	100%

Table 4.4.1-2 Sampling Results 002



14	Anthrax poison in Tide detergent packs coming thru the mail do not open or use. 7 ppl dead already. Plz pass it on! It was on cnn today.	Anthrax Poison - February 2011	100%
15	UK - Please forward to others - police will be stopping people from tomorrow who are driving with snow on the roof of their car and will issue fixed penalty of £60 + 3 points!!	Driving witn Snow on Car Roof - December 2010	100%
16	Just been text this message from a friend. her sister wrks 4the ambulance service: Jst heard at wrk - credible bomb threat 4 lpool 1 + trafford centre 4 2moro. Hospitals/burns unit on high alert. So avoid + pls pass msg on.	Liverpool And Manchester Bomb Threat - December 2010	100%
17	Everyone take off the cartoon pictures because the group asking everyone to change their profile picture to their favourite cartoon characters is actually a group of paedophiles. They're doing it because kids will accept their friend requests faster if they see a cartoon picture! It has nothing to do with a child violence campaign! Check the news tonight. Copy and paste this on to your status	Facebook Change Profile Pictures - December 2010	100%
18	There is someone called harry graham friend requesting kids on facebook at the moment. posing as a 14yr old when actually he is a 48yr old peadophile. He is known 2 the police. Please be aware and tell everyone u know. We must keep our kids safe - please copy and paste	Bogus Facebook Rumor - November 2010	100%
19	September 2010 you may have noticed that facebook servers have crashed and that you cannot sign in or on. this is because a group of hackers have hacked the main system. the facebook team are trying to save what they can. it is advised if facebook does come up and running again you change your password immediately. Thanks BBC news team try get the msg round as quick as possible.	Facebook and BBC Hoax - September 2010	100%
20	VIRUS on FB using your pictures. It says you have been tagged in a picture. wants you to click on a link to see it. then hacks into your computer & all your accounts. including banking & other secure accounts. It destroys your computer. Once hacked into your computer. it sends e-mails to your friends telling them they have been tagged in pictures & starts the process again PLEASE RE-POST	Virus on FB using user's Pictures - September 2010	100%
21	ATTENTION!!!!!!-Virus spreading like wildfire on Facebook!! It is a Trojan worm called "Knob Face". It will steal your info. invade your system and shut it down! DO NOT open the link "Barack Obama Clinton scandal". If "Smartgirl 15" adds you. don't accept it: it is a virus. If somebody on your list adds her then you will get the virus. too!! Copy and paste to your wall please!!	Knob Face Trojan - January 2011	100%
	Decorative Magnets on Refrigerators - DANGER ! A number of researchers at Princeton's University have discovered something scary!. For several months. they were feeding two groups of mice: the first group with food kept in a refrigerator. and the second group		

Table 4.4.1-3 Sampling Results 003



22	Decorative Magnets on Refrigerators - DANGER ! A number of researchers at Princeton's University have discovered something scary!. For several months, they were feeding two groups of mice: the first group with food kept in a refrigerator, and the second group with food kept in a refrigerator as well but with several decorative magnets on the door. The objective of this experiment was to see how electromagnetic radiation (that coming out from the decorative magnets on the door) affect food items. Amazingly, rigorous clinical studies stated that the group of mice that consumed the "radiated" food had as much as 87 % higher probability to get cancer than the other group of mice. Inexplicably no Governments or health associations/institutions have given any statement on this regard. However and just in case, is recommendable to remove any decorative magnet from refrigerators, and put it far away from any food. Kindly pass this information to your contacts.	Decorative Magnets on Refrigerators - July 2010	100%
23	Hello, greetings from RIM (Research In Motion) proprietors of BlackBerry. This message is to inform all of our users, that our servers have recently been really full, so we are asking for your help to fix this problem. We need our active users to re-send this message to everyone on your contact list in order to confirm our active users that use BlackBerry Messenger, if you do not send this message to all your BlackBerry Messenger contacts then your account will remain inactive with the consequence of losing all your contacts. We apologize for the inconvenience but this is the only way possible to resolve this problem. Sincerely Research in Motion. For more information visit: www.blackberry.com/inactiveuser	Blackberry Inactive Account Hoax - July 2010	100%
24	Don't add anyone named rehana n 27 years from leicister .she is a hacker. Tell everyone on your buddy list because if someone on your buddy list adds her, she'll be on your list too. She'll figure out your computer ID and address. So copy and paste this message to everyone on your buddy list because if she hacks them your next. I sent to everyone on my list so please send to everyone on your list	Rehana the Hacker - June 2010	100%

Table 4.4.1-4 Sampling Results 004

Based on the above snippets, the system is able to compare strings and match these hoaxes to the database at a rate of 100%. Calculated Standard Deviation is equivalent to 0 while the average or mean is at 100%. This has proven the system's accuracy in detecting hoaxes via its database at 100% rate with no failures.

4.4.2 Levenshtein Distance Formula

$$O\left(M \cdot N \cdot \max(M, N)\right)$$

Figure 4.4.2:1 Simplified Levenshtein Distance Formula

Levensthein Distance formula is the core foundation of this system. It builds up to almost 60% of the whole system. Levenshtein Distance (LD) is a metric for measuring the number of differences between 2 strings. Technically, it's very lengthy to explain but with the example provided below, this formula can easily be understood.

Basic Example 1:

- If the original string (os) is "bold" and compared string (cs) is also "bold", then LD(os,cs) = 0, because no transformations are needed. The strings are already identical to each other.
- If original string (os) is "bold" and compared string (cs) is "bonk", then LD(os,cs) = 2, because two substitution (changes of "l" to "n" and "d to "k") is sufficient to transform the original string to the compared string.

However, this formula has been slightly modified for this system. It uses the optimal string alignment concept instead of the simplified Damerau-Levenshtein Distance concept in which it ensures the organization of the text is similar to the original text. Further example will explain this concept better.

Basic Example 2:

- Simple Damerau-Levenshtein Distance:
  - If LD(CA,ABC) = 2
- Optimal String Alignment Concept (Used Levenshtein concept)
  - If LD(CA, ABC) = 3

Hence, getting to understand the capability of Levenshtein Distance, the system uses this method to compare string to string; between the original email text as well as the hoax text found in the database.

4.4.3 New Hoax Calculation (Add Comparison) Formula (200%)

As mentioned in the above section, the new hoax calculation comprises of 2 main focuses. They are with regards to the number of upper case characters found in a string as well as number of frequently used words found in the frequent word database suggested above (almost completed).

- Upper Case Character (100%)
  - This upper case character formula is derived by calculating the number of upper case characters to the total number of characters in the string.
  - Total percentage (%) = Total Upper Case Characeter / Total Characters
- Frequent word (100%)
  - The frequent word formula is derived by calculating the number of frequent words to the total number of words found in the string.
  - Total percentage (%) = Total Frequent Words / Total Words

Total of both calculations should reach 200%. The rational reason behind this is due to the fact that these patterns are still weak in recognition. Not all hoaxes uses upper case and not all of the frequent words are being used in total of all of the strings used. Hence, if we average the calculation, result will not show such logic calculation. It will be much convincing to the users if hoaxes are identified if it is 60% and above. To reach this objective, it is best to remain these two factors' addition at 200% maximum.

After all, even the experts do not have specific measurement as to how to identify hoaxes at first glance. Patterns recognition is essential and by all means, this pattern recognition has shown substantial results. So far, at least 6 out of 10 tries shows that this pattern recognition is valid and samples can be provided to prove this theory capable to handle new found hoaxes on its own. In other words, not only does the system able to compare string to string with its database but also detect possible new found hoaxes on its own term.

4.4.4 Interesting Argument

By accident, some of the samplings made were on trial and error basis. There are some samples which have been run even before the samples were being added into the hoax database. Having said so, with the accumulation of hoaxes previously done, the comparison shows that the result of 60% remains valid as most of the sampling hoaxes mentioned in particular to this section scores 60% and above even though the data has not been added to the database. Indeed, there are patterns to hoaxes and this interesting argument could be the hidden third factor in which the system may have and help the users in the near future to understand hoaxes better.

4.4.5 Time Consumption

Note that most of the time needed to run all these samples for results only takes less than 3 seconds in order to obtain the results.

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5. Conclusion and Recommendation

5.1 Relevancy to the Objective

In conclusion, this project is able to serve its objective presented in the earlier section of this paper. In the end, it is expected that the system should:

- Be able to self add new hoaxes to the database based on the identified characteristics of hoaxes presented in the referenced website
- Be able to present the probabilities in the interface for the user to see in order to create greater awareness for the users on the existence of hoax
- Able to compare string to string and provide accurate results
- Able to provide results in a short period of time (less than 5 seconds)

This project's results and outcomes have proven to be both relevant and beneficial to the society. Being aware of the existence of hoaxes create greater awareness for finding the right knowledge which can be used to further develop a human's perception as they grow. Misleading information can only cost losses to certain group of people including tarnishing their image in the eye of the public.

5.2 Future Recommendation for Work Expansion and Continuation

- Focus strongly on the hoax patterns and investigate more ways to identify hoaxes
- Suggest to create standards for hoaxes and conform to it for easier and valid comparison
- Note that the current system is made solely out of text based emails.

- Assume that the forwarded emails are based from Gmail service providers.
- Perhaps there are better methods to outrun current method's speed
- GUI interface should be improved using email layouts with these email hoaxes probabilities being listed beside the title of the email in inbox
- Ensures that there are no more special characters in SQL that will bother the insertion statement.
- System can be used more than just for hoax detection!

6. References

[1] Vukovic, M., Pizpuzic, K., & Belani, H. (September 30, 2009) An Intelligent Automatic Hoax Detection System. Zagreb, Croatia

[2] Hernandez, J.C., Hernandez, C.J, Sierra, J.M, Ribagorda, A: A First Step towards Automatic Hoax Detection. In: Proceedings of the International 36th Annual Carnahan Conference on Security Technology, pp. 102-114. IEEE, Piscataway

[3] Petkovic, T., Kostanjcar, Z., & Pale, P.: Email System for Automatic Hoax Recognition. In: XXVII. International Convenion MIPRO 2005 Bd. CTS & CIS, Opatija, Croatia, pp. 117-121 (2005) ISBN 953-233-012-7

[4] Brodie, R. (1995) Virus of the Mind: The New Sicence of the Meme, 1st edn. Integral Press, USA

[5] Blanzieri, E., Bryl, A. (July 10, 2009) A Survey of Learning-Based Tecniques of Email Spam Filtering. University of Trento, Trento, Italy

[6] Cormack, G.V. (April 2007) Email Spam Filtering: A Systemic Review. Hanover, MA, USA

[7] Buckingham, J.T., Hulten, G.J., Goodman, J.T., Rounthwaite, R.L. (December 31, 2004) Using Message Features and Sender Identity for Email Spam Filtering. WA, USA.

[8] Paliouras, G., Karkaletsis, V., Spyropoulos, C.D., Stamatopoulos, P. (2007) Learning to Filter Spam E-Mail: A Comparison of a Naive Bayesian and a Memory-Based Approach Ion Androutsopoulos. Athens, Greece

[9] Hsiao, W.F, Chang, T.M. (April, 2008) An Incremental Cluster-Based Approach to Spam Filtering. National Pingtung Instititue of Commerce, Taiwan: National Sun Yat-sen, Taiwan.

[10] Stanley, W. Et al (September 29, 2008) Classfication and Cluster Analysis Spam Detection and Reduction. Palo Alto, CA, USA.

[11] Abi-Haidar, A., Rocha, L.M. (2008) Adaptive Spam Detection Inspired by a Cross-Regulation Model of Immune Dynamics: A Study of Concept Drift.

Indiana University, Bloomington, IN, USA: Instituto Gulbenkian de Ciencia, Oeiras, Portugal

[12] Christensen, B. (2003) About Hoax Slayers from Hoax Slayers website

< /http://www.hoax-slayer.com/about.html/>

[13] Isa, D., Lam, H.L, Kallimani, V.P., Rajikumar, R (September, 2008) Text Document Pre Processing with Bayes Formula for Classification using the Support Vector Machine. Univesity of Nottingham, Malaysia

[14] Viola, P., Narasimhan, M. (2005) Learning to Extract Information from Semi-Structured Text Using a Discriminative Context Free Grammar. Redmond, WA, USA: University of Washington, Seattle, WA, USA

[15] Rozilawati, D., Aono, M. (2011) Ontology Based Approach for Classifying Biomedical Text Abstracts. Toyohashi University of Technology Hibarigaoka, Aichi, Japan

[16] Kanaris, I., Kanaris, K., Houvardas, I., Stamatatos, E. (2007) Word Versus Character N-Grams for Anti-Spam Filtering. University of Aegean, Karlovassi, Samos, Greece

[17] Ifrim, G., Bakir, G., Weikum, G. (2008) Fast Logistics Regression for Text Categorization with Variable-Length N-Grams. Saarbrucken, Germany: Zurich, Switzerland: Saarbrucken, Germany.

[18] Dennis, A., Wixom, B.H., Tegarden, D. (2005) System Analysis and Design, 2nd edn. Massachusetts: John Wiley & Sons.

7. Appendices

APPENDIX 2-1

Timelines for FYP 1

No.	Detail/ Week	1	2	3	4	5	6	7		8	9	10	11	12	13	14
1	Selection of Project Topic	■	■						Mid-semester break							
2	Preliminary Research Work		■	■	■	■										
3	Submission of Extended Proposal Defence						●									
4	Proposal Defence									■	■					
5	Project work continues											■	■	■		
6	Submission of Interim Draft Report														●	
10	Submission of Interim Report															●

● Suggested milestone

■ Process

APPENDIX 2-2

Timelines for FYP 2

No.	Detail/ Week	1	2	3	4	5	6	7		8	9	10	11	12	13	14	15
1	Project Work Continues	■	■	■	■	■	■	■	Mid-Semester Break								
2	Submission of Progress Report									●							
3	Project Work Continues									■	■	■	■	■			
4	Pre-EDX												●				
5	Submission of Draft Report													●			
6	Submission of Dissertation (soft bound)														●		
7	Submission of Technical Paper														●		
8	Oral Presentation															●	
9	Submission of Project Dissertation (Hard Bound)																●

● Suggested milestone

■ Process