ANDRÉ MIEDE

A CLASSIC THESIS STYLE

LyX PORT

BY

NICK MARIETTE & IVO PLETIKOSIĆ

(refer to **??** for more information)

[July 4, 2014 at 18:46 – Nurul Akbar version 4.1]

A CLASSIC THESIS STYLE

ANDRÉ MIEDE

An Homage to The Elements of Typographic Style

August 2012 – version 4.1

[July 4, 2014 at 18:46 – Nurul Akbar version 4.1]

Ohana means family.
Family means nobody gets left behind, or forgotten.

— Lilo & Stitch

Dedicated to the loving memory of Rudolf Miede.
1939 – 2005

ABSTRACT

Short summary of the contents in English...

ZUSAMMENFASSUNG

Kurze Zusammenfassung des Inhaltes in deutscher Sprache...

PUBLICATIONS

Some ideas and figures have appeared previously in the following publications:

Put your publications from the thesis here. The packages `multibib` or `bibtopic` etc. can be used to handle multiple different bibliographies in your document.

We have seen that computer programming is an art,
because it applies accumulated knowledge to the world,
because it requires skill and ingenuity, and especially
because it produces objects of beauty.

— **?** [**?**]

ACKNOWLEDGMENTS

Put your acknowledgments here.

Many thanks to everybody who already sent me a postcard!

Regarding the typography and other help, many thanks go to Marco Kuhlmann, Philipp Lehman, Lothar Schlesier, Jim Young, Lorenzo Pantieri and Enrico Gregorio[1], Jörg Sommer, Joachim Köstler, Daniel Gottschlag, Denis Aydin, Paride Legovini, Steffen Prochnow, Nicolas Repp, Hinrich Harms, Roland Winkler, Jörg Weber, and the whole LaTeX-community for support, ideas and some great software.

The LyX port was initially done by Nicholas Mariette in March 2009 and continued by Ivo Pletikosić in 2011. Thank you very much for your work and the contributions to the original style.

1 Member of GuIT (Gruppo Italiano Utilizzatori di TeX e LaTeX)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LISTINGS

ACRONYMS

UML Unified Modeling Language

Part I

SOME KIND OF MANUAL

1 INTRODUCTION

This bundle for LaTeX has two goals:

1. Provide students with an easy-to-use template for their Master's or PhD thesis. (Though it might also be used by other types of authors for reports, books, etc.)

2. Provide a classic, high-quality typographic style that is inspired by **?** 's "*The Elements of Typographic Style*" [**?**].

A Classic Thesis Style version 4.1

The bundle is configured to run with a *full* MiKTeX or TeXLive[1] installation right away and, therefore, it uses only freely available fonts. (Minion fans can easily adjust the style to their needs.)

People interested only in the nice style and not the whole bundle can now use the style stand-alone via the file `classicthesis.sty`. This works now also with "plain" LaTeX.

As of version 3.0, classicthesis can also be easily used with LyX[2] thanks to Nicholas Mariette and Ivo Pletikosić. The LyX version of this manual will contain more information on the details.

This should enable anyone with a basic knowledge of LaTeX 2_ε or LyX to produce beautiful documents without too much effort. In the end, this is my overall goal: more beautiful documents, especially theses, as I am tired of seeing so many ugly ones.

The whole template and the used style is released under the GNU General Public License.

If you like the style then I would appreciate a postcard:

André Miede
Detmolder Straße 32
31737 Rinteln
Germany

The postcards I received so far are available at:

`http://postcards.miede.de`.

A well-balanced line width improves the legibility of the text. That's what typography is all about, right?

So far, many theses, some books, and several other publications have been typeset successfully with it. If you are interested in some typographic details behind it, enjoy Robert Bringhurst's wonderful book.

Important Note:

Some things of this style might look unusual at first glance, many people feel so in the beginning. However, all things are intentionally designed to be as they are, especially these:

1 See the file `LISTOFFILES` for needed packages. Furthermore, `classicthesis` works with most other distributions and, thus, with most systems LaTeX is available for.

2 `http://www.lyx.org`

- No bold fonts are used. Italics or spaced small caps do the job quite well.
- The size of the text body is intentionally shaped like it is. It supports both legibility and allows a reasonable amount of information to be on a page. And, no: the lines are not too short.
- The tables intentionally do not use vertical or double rules. See the documentation for the `booktabs` package for a nice discussion of this topic.[3]
- And last but not least, to provide the reader with a way easier access to page numbers in the table of contents, the page numbers are right behind the titles. Yes, they are *not* neatly aligned at the right side and they are *not* connected with dots that help the eye to bridge a distance that is not necessary. If you are still not convinced: is your reader interested in the page number or does she want to sum the numbers up?

Therefore, please do not break the beauty of the style by changing these things unless you really know what you are doing! Please.

1.1 ORGANIZATION

A very important factor for successful thesis writing is the organization of the material. This template suggests a structure as the following:

You can use these margins for summaries of the text body...

- `Chapters/` is where all the "real" content goes in separate files such as `Chapter01.tex` etc.
- `FrontBackMatter/` is where all the stuff goes that surrounds the "real" content, such as the acknowledgments, dedication, etc.
- `gfx/` is where you put all the graphics you use in the thesis. Maybe they should be organized into subfolders depending on the chapter they are used in, if you have a lot of graphics.
- `Bibliography.bib`: the BibTEX database to organize all the references you might want to cite.
- `classicthesis.sty`: the style definition to get this awesome look and feel. Does not only work with this thesis template but also on its own (see folder `Examples`). Bonus: works with both LATEX and PDFLATEX... and LYX.
- `ClassicThesis.tcp` a TEXnicCenter project file. Great tool and it's free!

3 To be found online at http://www.ctan.org/tex-archive/macros/latex/contrib/booktabs/.

- `ClassicThesis.tex`: the main file of your thesis where all gets bundled together.

- `classicthesis-config.tex`: a central place to load all nifty packages that are used. In there, you can also activate backrefs in order to have information in the bibliography about where a source was cited in the text (i. e., the page number).
  Make your changes and adjustments here. This means that you specify here the options you want to load `classicthesis.sty` with. You also adjust the title of your thesis, your name, and all similar information here. Refer to Section 1.3 for more information.
  This had to change as of version 3.0 in order to enable an easy transition from the "basic" style to LyX.

In total, this should get you started in no time.

1.2 STYLE OPTIONS

There are a couple of options for `classicthesis.sty` that allow for a bit of freedom concerning the layout:

. . . or your supervisor might use the margins for some comments of her own while reading.

- General:
  - `drafting`: prints the date and time at the bottom of each page, so you always know which version you are dealing with. Might come in handy not to give your Prof. that old draft.

- Parts and Chapters:
  - `parts`: if you use Part divisions for your document, you should choose this option. (Cannot be used together with `nochapters`.)
  - `nochapters`: allows to use the look-and-feel with classes that do not use chapters, e. g., for articles. Automatically turns off a couple of other options: `eulerchapternumbers`, `linedheaders`, `listsseparated`, and `parts`.
  - `linedheaders`: changes the look of the chapter headings a bit by adding a horizontal line above the chapter title. The chapter number will also be moved to the top of the page, above the chapter title.

- Typography:
  - `eulerchapternumbers`: use figures from Hermann Zapf's Euler math font for the chapter numbers. By default, old style figures from the Palatino font are used.

  - `beramono`: loads Bera Mono as typewriter font. (Default setting is using the standard CM typewriter font.)
  - `eulermath`: loads the awesome Euler fonts for math. (Palatino is used as default font.)
  - `pdfspacing`: makes use of pdftex' letter spacing capabilities via the `microtype` package.[4] This fixes some serious issues regarding math formulæ etc. (e. g., "ß") in headers.
  - `minionprospacing`: uses the internal `textssc` command of the `MinionPro` package for letter spacing. This automatically enables the `minionpro` option and overrides the `pdfspacing` option.

- Table of Contents:
  - `tocaligned`: aligns the whole table of contents on the left side. Some people like that, some don't.
  - `dottedtoc`: sets pagenumbers flushed right in the table of contents.
  - `manychapters`: if you need more than nine chapters for your document, you might not be happy with the spacing between the chapter number and the chapter title in the Table of Contents. This option allows for additional space in this context. However, it does not look as "perfect" if you use `\parts` for structuring your document.

- Floats:
  - `listings`: loads the `listings` package (if not already done) and configures the List of Listings accordingly.
  - `floatperchapter`: activates numbering per chapter for all floats such as figures, tables, and listings (if used).
  - `subfig(ure)`: is passed to the `tocloft` package to enable compatibility with the `subfig(ure)` package. Use this option if you want use classicthesis with the `subfig` package.

The best way to figure these options out is to try the different possibilities and see, what you and your supervisor like best.

In order to make things easier in general, `classicthesis-config.tex` contains some useful commands that might help you.

1.3 CUSTOMIZATION

This section will give you some hints about how to adapt `classicthesis` to your needs.

4 Use `microtype`'s `DVIoutput` option to generate DVI with pdftex.

The file `classicthesis.sty` contains the core functionality of the style and in most cases will be left intact, whereas the file `classicthesis-config.tex` is used for some common user customizations.

The first customization you are about to make is to alter the document title, author name, and other thesis details. In order to do this, replace the data in the following lines of `classicthesis-config.tex`:

Modifications in `classicthesis-config.tex`

```
% ********************************************************
% 2. Personal data and user ad-hoc commands
% ********************************************************
\newcommand{\myTitle}{A Classic Thesis Style\xspace}
\newcommand{\mySubtitle}{An Homage to...\xspace}
```

Further customization can be made in `classicthesis-config.tex` by choosing the options to `classicthesis.sty` (see section 1.2) in a line that looks like this:

```
\PassOptionsToPackage{eulerchapternumbers,listings,drafting,
    pdfspacing,subfig,beramono,eulermath,parts}{classicthesis}
```

If you want to use backreferences from your citations to the pages they were cited on, change the following line from:

```
\setboolean{enable-backrefs}{false} % true false
```

to

```
\setboolean{enable-backrefs}{true} % true false
```

Many other customisations in `classicthesis-config.tex` are possible, but you should be careful making changes there, since some changes could cause errors.

Finally, changes can be made in the file `classicthesis.sty`, although this is mostly not designed for user customisation. The main change that might be made here is the text-block size, for example, to get longer lines of text.

Modifications in `classicthesis.sty`

1.4 ISSUES

This section will list some information about problems using `classicthesis` in general or using it with other packages.

Beta versions of classicthesis can be found at the following Google code repository:

http://code.google.com/p/classicthesis/

There, you can also post serious bugs and problems you encountered.

Compatibility with the `glossaries` Package

If you want to use the `glossaries` package, take care of loading it with the following options:

```
\usepackage[style=long,nolist]{glossaries}
```

Thanks to Sven Staehs for this information.

Compatibility with the (Spanish) `babel` Package

Spanish languages need an extra option in order to work with this template:

```
\usepackage[spanish,es-lcroman]{babel}
```

Thanks to an unknown person for this information (via Google Code issue reporting).

Compatibility with the `pdfsync` Package

Using the `pdfsync` package leads to linebreaking problems with the `marginpar/graffito` command. Thanks to Henrik Schumacher for this information.

1.5 FUTURE WORK

So far, this is a quite stable version that served a couple of people well during their thesis time. However, some things are still not as they should be. Proper documentation in the standard format is still missing. In the long run, the style should probably be published separately, with the template bundle being only an application of the style. Alas, there is no time for that at the moment... it could be a nice task for a small group of LaTeXnicians.

Please do not send me email with questions concerning LaTeX or the template, as I do not have time for an answer. But if you have comments, suggestions, or improvements for the style or the template in general, do not hesitate to write them on that postcard of yours.

1.6 BEYOND A THESIS

It is easy to use the layout of `classicthesis.sty` without the framework of this bundle. To make it even easier, this section offers some plug-and-play-examples.

The LaTeX-sources of these examples can be found in the folder with the name `Examples`. They have been tested with `latex` and `pdflatex` and are easy to compile. To assure you even a bit more, PDFs built from the sources can also be found the folder.

Listing 1: An Article

```
% article example for classicthesis.sty
\documentclass[10pt,a4paper]{article} % KOMA-Script article
    scrartcl
```

```
\usepackage{lipsum}
\usepackage{url}
\usepackage[nochapters]{../classicthesis} % nochapters

\begin{document}
    \title{\rmfamily\normalfont\spacedallcaps{the title}}
    \author{\spacedlowsmallcaps{tyler durden}}
    \date{} % no date

    \maketitle

    \begin{abstract}
        \noindent\lipsum[1] Just a test.\footnote{This is a
            footnote.}
    \end{abstract}

    \tableofcontents

    \section{A Section}
    \finalVersionString \lipsum[1]
    \subsection{A Subsection}
    \lipsum[1]
    \subsection{A Subsection}

    \section{A Section}
    \lipsum[1]

    % bib stuff
    \nocite{*}
    \addtocontents{toc}{\protect\vspace{\beforebibskip}}
    \addcontentsline{toc}{section}{\refname}
    \bibliographystyle{plain}
    \bibliography{../Bibliography}
\end{document}
```

Listing 2: A Book

```
% book example for classicthesis.sty
\documentclass[11pt,a5paper,footinclude=true,headinclude=true,
    english]{scrbook} % KOMA-Script book
\usepackage[T1]{fontenc}
\usepackage{lipsum}
\usepackage[linedheaders,parts]{../classicthesis} % ,manychapters
%\usepackage[osf]{libertine}
\hypersetup{linktocpage=true,bookmarksnumbered=true,pageanchor=
    true,hypertexnames=false,naturalnames=true,plainpages=false}

\begin{document}
%       \pagestyle{scrheadings}
%       \manualmark
%       \markboth{\spacedlowsmallcaps{\contentsname}}{\
    spacedlowsmallcaps{\contentsname}}
```

```
        \tableofcontents

%       \automark[section]{chapter}
%       \renewcommand{\chaptermark}[1]{\markboth{\
    spacedlowsmallcaps{#1}}{\spacedlowsmallcaps{#1}}}
%       \renewcommand{\sectionmark}[1]{\markright{\thesection\
    enspace\spacedlowsmallcaps{#1}}}

    % use \cleardoublepage here to avoid problems with
        pdfbookmark
    \cleardoublepage\part{Test Part}
    \chapter{Test Chapter}
    \lipsum[1]

    \section{A Section}
    \lipsum[1]

    \chapter{Test Chapter}
    \lipsum[1]

    \section{A Section}
    \lipsum[1]

%       \include{multiToC}

    \appendix
    \cleardoublepage\part{Appendix}
    \chapter{Appendix Chapter}
    \lipsum[1]

    \section{A Section}
    \lipsum[1]

\end{document}
```

Listing 3: A Curriculum Vitæ

```
% cv example for classicthesis.sty
\documentclass[10pt,a4paper]{scrartcl}
\usepackage[LabelsAligned]{currvita} % nice cv style
\usepackage{url}
\usepackage[ngerman]{babel}
\usepackage[nochapters]{../classicthesis}
% Some font experiments
%\usepackage[osf]{libertine}
%\usepackage{hfoldsty}
%\usepackage[math]{iwona} %[light,condensed,math]
%\renewcommand{\sfdefault}{iwona}
%\usepackage{lmodern} % <-- no osf support :-(
%\usepackage[urw-garamond]{mathdesign} %<-- no osf support :-(
```

```
\renewcommand*{\cvheadingfont}{\LARGE\color{Maroon}}
\renewcommand*{\cvlistheadingfont}{\large}
\renewcommand*{\cvlabelfont}{\qquad}

\begin{document}
    \begin{cv}{\spacedallcaps{Curriculum Vit\ae}}
        %\pdfbookmark[1]{Pers\"onliche Daten}{PersDat}
        \begin{cvlist}{\spacedlowsmallcaps{Pers\"onliche Daten}}\
            label{PersDat}
            \item   Dr.-Ing.~Andr\'e Miede
            \item   Geboren am \dots\ \texttt{(-;} \\
                    Europ\"aer, Deutsche Staatsb\"urgerschaft
            \item   \url{http://www.miede.de} \\
                    \url{https://www.xing.com/profile/Andre_Miede
                        }
        \end{cvlist}

        %\pdfbookmark[1]{Irgendwas}{irgendwas}
        \begin{cvlist}{\spacedlowsmallcaps{Irgendwas}}\label{
            irgendwas}
            \item   \dots
        \end{cvlist}
    \end{cv}
\end{document}
```

1.7 LICENSE

Part II

THE SHOWCASE

2 INTRODUCTION OF PHISHING

While the Internet has brought convenience to many people for exchanging information, it also provides opportunities to carry malicious behavior such as online fraud on massive scale with a little cost to the attackers. The attackers can manipulate the Internet users instead of computer system (hardware or software) that significantly increase the barriers of technological crime impact. Such human centered attack could be done by social engineering. Phishing is a form of social engineering that aim to retrieve credential from online users by mimicking trustworthy and legitimate institutions [7]. These fraudulent attacks are most frequently done by electronic communication such as emails to direct users to fake websites and prompt for sensitive information.

This chapter provides an overview to phishing modus operandi and its current countermeasures such as detection and prevention techniques along with a brief exploration of direct and indirect cost of phishing attacks.

2.1 WHAT IS PHISHING?

In the real world, phishing has a similar basic principle as 'fishing'. Instead of fish, online users are lured by authentic looking communication and hooked by authentic looking websites. Phishing is a malicious technique to manipulate online users into sharing their sensitive information by masquerading as legitimate and trustworthy institutions. Phishing attack also is a form of online criminal activity by using social engineering technique [7]. An individual or a group who uses this technique is called *Phishers*. After successfully gain sensitive information from the victim, phishers use this information to access victim's financial accounts or committing credit card frauds. The technique of phishing may vary, but the most common technique of phishing attacks done by using fraudulent emails and websites [8]. A fraudulent website is designed in such a way that it would be very identical to its legitimate target.

2.2 HISTORY OF PHISHING

The first time the term "phishing" was published by the American Online (AOL) UseNet Newsgroup on January 2, 1996 and began to expand in 2004 [11]. Since then, phishing development in cyberspace unfortunately was a great achievement by phishers to make profit.

Total losses due to phishing in 2004 reached more than U.S. $ 2 billion, it was involving more than 15,000 sites that become victims [4]. Jakobsson, et al. mentioned that in the early years of 90's (according to [11] it was around 1995) many hackers would create bogus AOL user accounts with automatically generated fraudulent credit card information [7]. Their intention to give this fake credit card information was to simply pass the validity tests performed by AOL. By the time the tests were passed, AOL was thinking that these accounts were legitimate and resulted to activate them. Consequently, these hackers could freely access AOL resources until AOL tried to actually bill the credit card. AOL realized that these accounts were using invalid billing information, thus deactivated the account.

While creating false AOL user accounts with fake credit card information was not exactly phishing attacks, but AOL's effort to counter against the attacks was leading to development of phishing. This countermeasure includes directly verifying the legitimacy of credit card information and the associated billing identity, forced hackers to pursue alternative way [7]. Hackers would masquerade themselves as AOL's employees asking to other users for credit card information through AOL instant messenger and email system. At this point, we believe the term of phishing attack has been born. Since such attack has not been done before, many of users have been victimized by then. Eventually, AOL enforced warning system to the most of its customers to be vigilant when it comes to sensitive information [11]. At the present day, phishing attacks have evolved not only aim to AOL users, but also any online users motivated by financial gain. Consequently, large number of legitimate institutions such as PayPal and eBay are being spoofed.

2.3 FORMAL DEFINITION OF PHISHING

Before we begin to dig deeper understanding of phishing attacks, we will briefly explore common phishing definition. Currently, there is no consensus definition, since almost in every research papers, academic textbook or journals has its own definition of phishing [7, 8, 12, 1, 10, 6, 2]. Phishing is also constantly evolving, so it might be very challenging to define its universal terminology. There is not so much study that specifically addresses the standard of phishing definition. We will take a look of one particular phishing definition from various sources:

1. Phishing is the act of sending a forged e-mail (using a bulk mailer) to a recipient, falsely mimicking a legitimate establishment in an attempt to scam the recipient into divulging private information such as credit card numbers or bank account passwords [8]

2. Phishing is a form of Internet scam in which the attackers try to trick consumers into divulging sensitive personal information. The techniques usually involve fraudulent E-mail and web sites that impersonate both legitimate E-mail and web sites [12]

3. Phishing is an attack in which victims are lured by official looking email to a fraudulent website that appears to be that of a legitimate service provider [1]

4. In phishing, an automated form of social engineering, criminals use the internet to fraudulently extract sensitive information from businesses and individuals, often by impersonating legitimate web sites [10]

It is noteworthy that the definition described by [2, 5, 6] specifies that the phishers only use email as a communication channel to trick potential victims. While it might be true because using email would greatly cost effective, but we believe that phishing is not only characterized by one particular technological mean, as phishers can also use any other electronic communication to trick potential victims (i.e private message on online social network). This definition is also similar to dictionary libraries [10-12] that mention email as a medium communication between phishers and users.

We believe that standard definition of phishing should be applicable in most of phishing concept that are presently defined. Consequently, the high level of abstraction and is required to build common definition on phishing. We also argued that the definition of phishing should not focus on the technology being used but rather on the methodology how the deception being conducted. Therefore, We follow the definition of phishing by Lastdrager [13] which stated:

> *Phishing is a scalable act of deception whereby impersonation is used to obtain information from a target*

The definition presented above is developed in a comprehensive way by using existing definitions as input and combined them. A systematic review of literature up to August 2013 was conducted along with manual peer review, which resulted in 113 distinct definitions to be analyzed. We thereby agree with Lastdrager [13] that this definition addresses all the essential elements of phishing and we will adopt it as universally accepted terminology throughout our research.

2.4 ECONOMICS IMPACT CAUSED BY PHISHING

It is non-trivial task to find a real cost from phishing damage in term of money or direct cost. This due to phishing economy is consistent with black market economy and does not advertise its successes [1]. On this section, brief explanation of direct cost on phishing attack will be illustrated based on literature reviews.

The difficulty of assessing the damage on phishing attack is caused by unwillingness of many users to share to acknowledge that they have been victimized by phishing attacks. This happens might be because of fear of humiliation, financial loses, or legal liability [1]. However, studies estimate the damage ranging from $61 million [14] to $3 billion per year [15] of direct losses to victims in the US only [16]. The Gartner Group claimed to estimate of $1.2 billion direct losses of phishing attack to US banks and credit card companies for the year 2004 [17]. By the 2007, it escalated to more than $3 billion loss [18]. The estimation also performed by TRUSTe and Ponemon Institute that stated the cost of phishing attack was up to $500 millions losses in the US for the same year [1]. Moreover, figure [num] illustrated the direct cost of phishing in 2013 by the same companies. <still in progress>

2.5 PHISHING MODUS OPERANDI

As we mentioned earlier, Phishing attack is a form of cybercrime. The technique of the attack usually carried out firstly by creating a replica/clone of a legitimate website such as financial website with almost 100% similar. After that, the phishers will try to trick the potential victim to submit important information such as usernames, passwords, PINs, etc. through a fake website that they have created. With the information obtained, they will try to steal money from their victims. Phishers employ variety of techniques to trick potential victims to access their fraudulent website. One of the typical ways is by sending illicit email in a large scale claiming to be from legitimate institution. In the email content, they usually imitate an official-looking logo, using good business language style and often also forge the email headers to make it look like originating from legitimate institution. Typically, the content of the email is to inform the user that the bank is changing its IT infrastructure, and request urgently that the customer should update their data with the consequence of loosing their money if the action does not take place. When the user click the link that was on the email message, they will be redirected to a fraudulent website, which will prompt the victim to fill in the details of their information. While there are various techniques of phishing attack, we will address the common phases of phishing that we are analyzed by literature survey and we will present our own phase.

Based on the example scenario explained earlier, we believe that phishing attacks may consist of several phases. J. Hong [16] argued that there are three major phases:

1. Potential victims receive a phish.
2. The victim may take a suggested action in the message.

1 http://www.theregister.co.uk/2004/09/29/phishing_survey/

Figure 1: Example of fake ING logo in phishing email

3. The phisher monetizes the stolen information.

Frauenstein, et al. [19] suggested that typically there are five main processes are used to perform phishing attack. On the first process, a phisher usually will do some reconnaissance on how would the attack is executed and what information would be obtained from the victim. The first process is called planning. On the second process, a phisher typically deliver its message via email. This email is desired by the phisher to look as legit as possible to potential victim. For this purpose, target institutions logo, trademark, symbol, etc. are used to make the content look official to the victim. The author called this process as Email Design. Figure 1 illustrates the example of fake ING bank logo in a phishing email to create "legitimate" feel[2].

On the third process, phisher fabricate a story to make potential victim think that email is important. To achieve users attention, usually phisher build up a story about system upgrade, account hijacked, security enhancement, etc. so that the victim would feel obliged to be informed. This technique is commonly known as reverse social engineering. Moreover, we believe this process also corresponds with Cialdini [20] that suggested reciprocation as one of the technique to persuade people. On the fourth process, a phisher usually include threatening tone or explain the urgency and consequences if the potential victim chooses not to take action desired by the phisher (for example; account removal, account blocked, etc.). Consequently, users may fear of their account being deleted. This process also corresponds with the theory of persuasion called authority [20]. The last process involved with fraudulent website that has been created by the phisher. Users may falsely believe to the message given in the email and may click a URL/link that is embedded in the email. Subsequently, the URL would redirect users to this fraudulent website that typically prompt users' sensitive information. Furthermore, the

2 http://www.martijn-onderwater.nl/wp-content/uploads/2010/03/ing-phishing.jpg

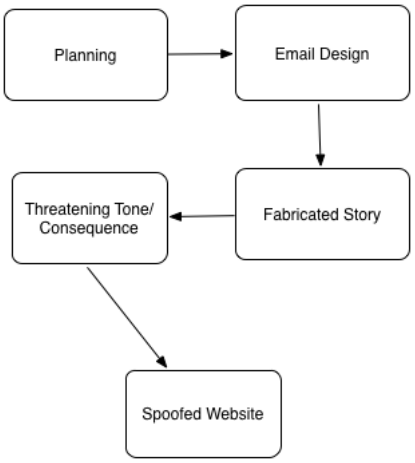


Figure 2: Phishing processes based on Frauenstein[5]

website is created to be as similar as possible to the target institution's website, so that potential victim may still believe that it is authentic. We will explain more on Cialdini's six basic tendencies of human behavior in generating positive response to persuasion [20] in a later section. To make a better understanding, we have created a diagram of these processes based on Frauenstein's typical five main processes [19] in Figure 2.

Considering that phishing attack is a process, Wetzel [21] suggested a taxonomy to make sense of the complex nature of the problem by mapping out a common attack lifecycle, and a possible set of activities attackers engage in within each phase. The taxonomy is illustrated in Figure 3.

A study suggest that there are several phases involved in phishing attack [5]:

1. The attacker obtains E-mail addresses for the intended victims. These could be guessed or obtained from a variety of sources.
2. The attacker generates an E-mail that appears legitimate and requests the recipient to perform some action.
3. The attacker sends the E-mail to the intended victims in a way that appears legitimate and obscures the true source.
4. Depending on the content of the E-mail, the recipient opens a malicious attachment, completes a form, or visits a web site.
5. The attacker harvests the victim's sensitive information and may exploit it in the future.

The phases described by [5] are also analogous with the information flow explained by [22] represented in Figure 4.

Quoted from Emigh [22], the information flow of phishing attack is described by the following phases without countermeasures part:

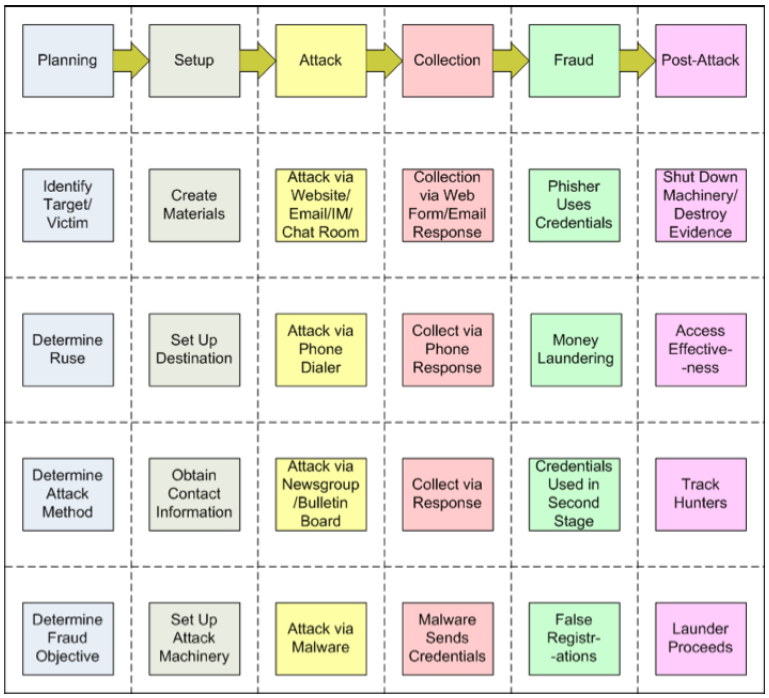


Figure 3: Phishing attack taxonomy and lifecycle[13]

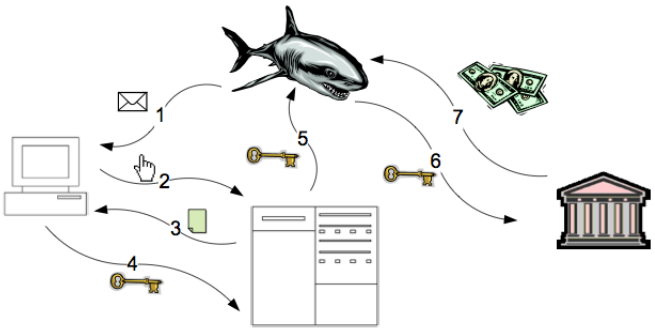


Figure 4: Flow of information in phishing attack [3]

1. A malicious payload arrives through some propagation vector.
2. The user takes an action that makes him or her vulnerable to an information compromise.
3. The user is prompted for confidential information, either by a remote web site or locally by a Web Trojan.
4. The user compromises confidential information.
5. The confidential information is transmitted from a phishing server to the phisher.
6. The confidential information is used to impersonate the user.
7. The phisher engages in fraud using the compromised information.

Phishing attack steps are also addressed by [23]. In their study, a successful phishing attack involves several phases:

1. Preparation
2. Delivery of the Lure
3. Taking the Bait
4. Request for Confidential Information
5. Submission of Information
6. Collection of Data
7. Impersonation
8. Financial Gain

As we can see from the earlier phases, there is a high amount of similarity between each other. We believe our phases are also analogous with other earlier studies. We thereby present our own phases of phishing attack and its categorization such as the lure, the hook and the catch. From Figure 5, we believe that the phishers generally characterized in the following way:

- The lure

1. Phishers prepare the attack
2. Deliver initial payload to victim
3. Direct spoofed website

- The hook

4. Prompt for confidential information

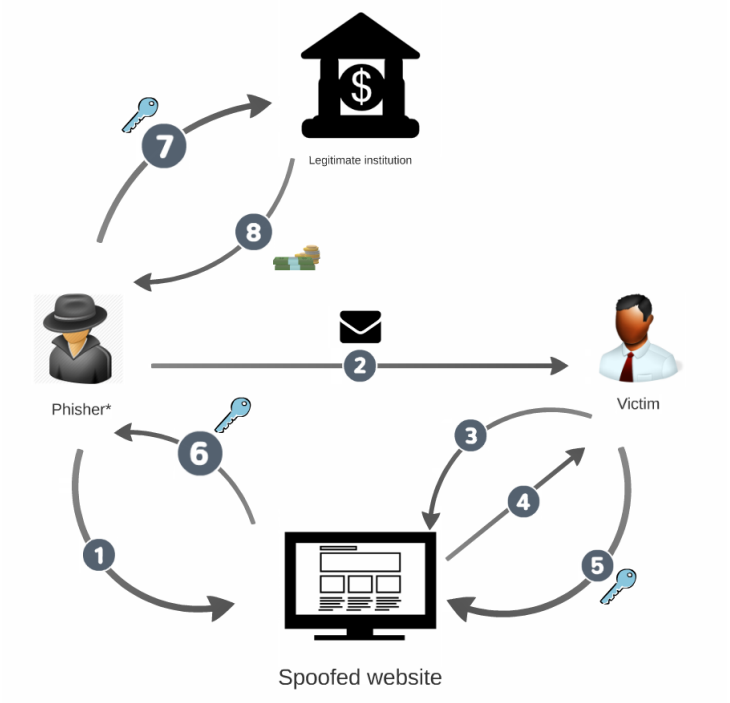


Figure 5: Information flow phishing attack

5. Disclosed confidential information
6. Collect stolen information

- The catch

7. Impersonates victim
8. Received pay out from the bank

2.6 COMMON TYPE OF PHISHING

In January 2014, 8300 patients data are being compromised in medical company in the US[3]. The data includes names, addresses, date of birth and phone numbers were being stolen. Other than demographic information, clinical information associated with this data was also stolen, including social security numbers. In the April 2014, phisher has successfully stole US$163,000 from US public school based on Michigan[4]. It has been said that the email prompted to transfer money is coming from the finance director of the school. In March 2014, Symantec has discovered phishing attack aimed at Google drive

3 http://www.scmagazine.com/medical-staffers-fall-for-phishing-emails-data-on-8300-compromised/article/340590/
4 http://www.scmagazine.com/phishing-scam-targets-michigan-public-schools/article/343177/

users[5]. The attack was carried firstly with incoming email asking for opening document hosted at Google docs. Users that have clicked on the link are taken to fraudulent Google login page prompted Google users credentials. Interestingly, the URL seems very convincing because it hosted on Google secure servers. We believe even more phishing incidents on financial area as well, but sometimes the news is kept hidden due to creditability reason.

One may ask, what type of phishing are these? What type of phishing commonly used nowadays? The threat of phishing attacks is still alarming until today, and will be evolving in the future with more sophisticated technique of attacks. For this reason, we believe that it is necessary to provide a brief insight on popular variants of phishing that currently exist.

2.6.1 *Deceptive phishing*

There are many variations based on deceptive phishing schemes. Typical scenario of deceptive phishing schemes is to send a large amount of illicit emails containing call to action asking recipients to click embedded links [1]. These variations include cousin domain attack. For example, legitimate PayPal website addressed as www.paypal.com, this cousin domain attacks confuse potential victims to believe that www.paypal-security.com is a subdivision of the legitimate website due to identical looking addresses. Similarly, homograph attacks create a confusion using similar characters to its addresses. For example, www.paypal.com and www.paypa1.com, both addresses look the same but on the second link, it uses "1" instead of "l".

Moreover, phishers may embed a login page directly to the email content. This suggests the elimination of the need of end-users to click on a link and phishers do not have to manage an active fraudulent website. IP addresses are often used instead of human readable hostname to redirect potential victim to phishing website and JavaScript is used to take over address bar of a browser to make potential victims believe that they are communicating with the legitimate institution. We will also see few examples of malicious JavaScript on our preliminary analysis section.

Another type of deceptive phishing scheme is rock-phish attacks. They held responsible for half a number of reported incidents worldwide in 2005 [24]. These attacks evade email filters by utilizes random text and GIF images which contain the actual message. Rock phish attacks also utilize a toolkit that capable to manage several fraudulent websites in a single domain. Sometimes, deceptive phishing schemes lead to installation of malware when users visit fraudulent website

5 http://www.scmagazine.com/phishing-scam-aimed-at-google-docs-drive-users/article/338369/

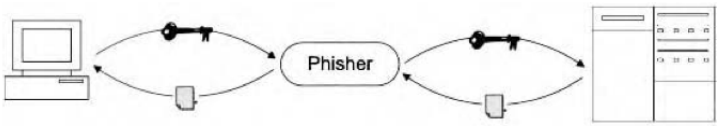


Figure 6: Man in the middle phishing[7]

and we will describe malware based phishing scheme in the next section.

2.6.2 *Malware-based phishing*

Generally, malware based phishing refers to any type of phishing which involves installing malicious piece of software onto users' personal computer [1]. Subsequently, this malware is used to gather confidential information from victims instead of spoofing legitimate websites. This type of phishing incorporates malwares such as keyloggers/screenloggers, web Trojans and hosts file poisoning.

2.6.3 *Man in the middle phishing*

Man-in-the-middle phishing attacks refer to the phishers that positioned in the middle between a legitimate institution and an end user [1]. The objective of these attacks is to make end users and legitimate institution (eg. Internet banking website) believe that they are truly communicating with each other. According to [1], Figure 6 illustrates the information flow of this type of attacks. Confidential information intended for legitimate site will be passed to the phishers before they can forward it to the legitimate site. Similarly, information from legitimate site to end-users will be passed to the phishers first before the phishers can forward it to them. Unfortunately, man-in-the-middle attacks are also capable to exploit two-factor authentication system (sometimes called two steps verification) that ensures the authenticity of sender and recipient (Eg. Popular ABN Amro case in 2007[6]).

2.7 BAD NEIGHBORHOODS ON PHISHING

In the real world, there are parts of certain area that have higher crime rates than others (eg. Bronx in the US). Evidently, it is statistically more likely that a crime will occur compared to other locations [25]. To better illustrate this analogy, the police department in Belgium[7] has put up statistical information regarding crime rates in the country. Figure 7 shows an example in 2013; there were up to 5525 car theft

6 http://www.theregister.co.uk/2007/04/19/phishing_evades_two-factor_authentication

7 http://www.polfed-fedpol.be/

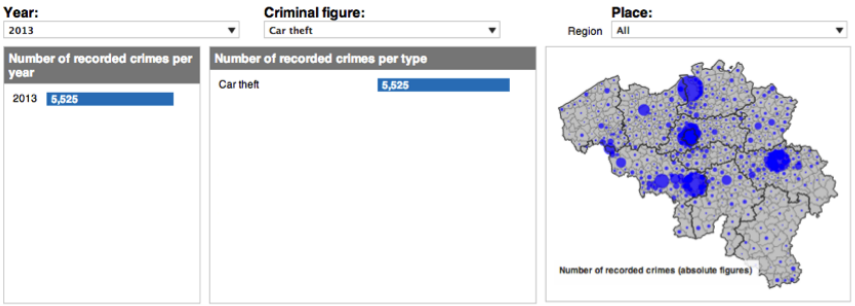


Figure 7: Belgium police record on car theft incidents in 2013

incidents recorded and we can see there are certain areas that have higher probability that a car got stolen. For example, Antwerp had 415 incidents whereas Berlare had only 1 car theft incident[8]. The data is not necessarily based on cities, it can be based on the residential area within a city. This holds true that much higher crime rates in a concentrated location compared to any other locations. It is called bad neighborhood.

To reduce the crime rates in a bad neighborhood, it makes sense that the authority should put more enforcement in this location. Moreover, the citizen should avoid this location as much as they can if they want to feel much safer. Evidently, the existence of bad neighborhood phenomenon also occurs in the Internet world called "Internet bad neighborhoods". There are certain networks of Internet infrastructure that contain more malicious activities that other networks. For our preliminary analysis, we will adopt formal definition of Internet bad neighborhoods or Internet Badhoods by [25] which states:

> *Internet bad neighborhood is a set of IP addresses clustered according to an aggregation criterion in which a number of IP addresses perform a certain malicious activity over a specified period of time*

Several studies have suggested that the source of the Internet Badhoods tend to be concentrated in certain portions of IP address space [26-28]. A dissertation study argues that Internet Badhoods do not always only based on network prefixes level (e.g. /24, /32, etc..) but it can be aggregated into several levels (ISPs, Countries) [25]. Moreover, Internet Badhoods may vary depending on which application exploited. While spam is most likely distributed all around the world, however, phishing Badhoods are most likely concentrated in developed countries (e.g. US) [25]. This suggests that phishing sites are required to have more reliable hosts in term of availability, while spams are not. We will see on section [num] our preliminary analysis on phishing Badhoods.

8 http://www.polfed-fedpol.be/crim/crim_statistieken/app_crimestat/app_crimestat_dashboard_crimfig_misdrijven_nl.php

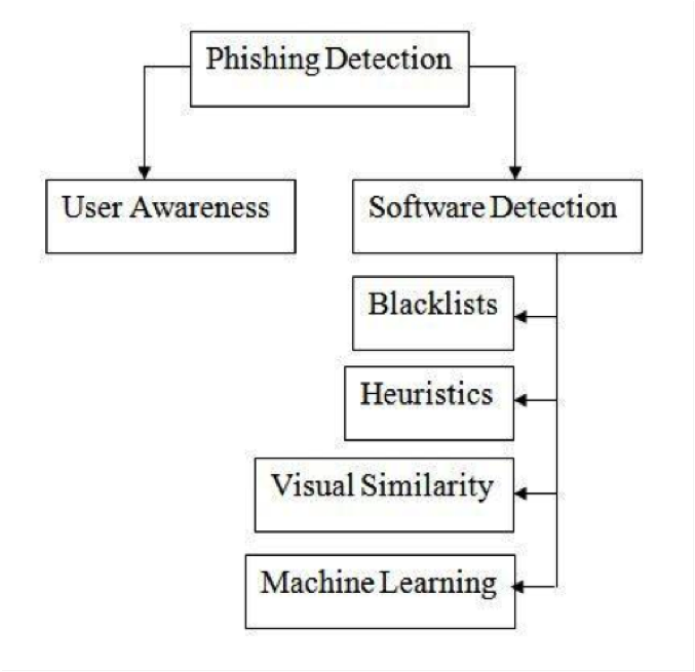


Figure 8: Phishing detection[9]

2.8 CURRENT COUNTERMEASURES OF PHISHING ATTACKS

There are various types of phishing countermeasures. However, since all the approaches reviewed so far all are preventive in nature, we believe phishing detection also aims to prevent user's confidential information to be successfully transmitted onto the wrong hands. A study suggests that phishing detection can be classified into two types; user training approach and software classification approach [29]. Figure 8 categorizes the current countermeasures that presently developed. Table 1 also indicates the advantages and the disadvantages of each category[9]. However, we do not aim to give a complete wide range literature survey on all available phishing countermeasures. We also do not evaluate the effectiveness of current anti-phishing technology. We will only discuss our result of literature review on Phishtank as a blacklist and machine learning classification, and anti-phishing training as prevention.

2.8.1 *Phishing detection*

2.8.1.1 *Phishtank*

One of the common approaches to detect phishing attack is blacklisting. Phishtank is a blacklisting company specifically for phishing URLs and it is a free community web based where users can report, verify and track phishing URLs. Phishtank stores phishing URLs in its database and the data is widely available for use by other compa-

Detection techniques	Advantages	Disadvantages
Blacklist	• Requiring low resources on host machine • Effective when minimal FP rates are required	• Mitigation of zero hour phishing attack • Can result in excessive queries with heavily loaded servers
Heuristics and visual similarity	• Mitigate zero hour attacks	• Higher FP rate than blacklists • High computational cost
Machine Learning	• Mitigate zero hour attacks • Construct own classification model	• Time consuming • Costly • Huge number of rules

Table 1: Advantages-disadvantages detection technique[9]

nies. Some of the big companies that are using Phishtank's data includes; Yahoo Mail, McAfee, APWG, Web Of Trust, Kaspersky, Opera and Avira. In this section, we will discuss how the current literatures have to do with the data provided by Phishtank. Table 2 summarizes the papers selected and its relevancy with Phishtank.

Table 2: Summary phishtank studies

Paper title	**First author**	**Country**	**Relevancy with phishtank**
Evaluating the wisdom of crowds in assessing phishing website [30]	Tyler Moore	United Kingdom	Examine the structure and outcomes of user participation in Phishtank. The authors find that Phishtank is dominated by the most active users, and that participation follows a power law distribution and this makes it particularly susceptible to manipulation.
Re-evaluating the wisdom of crowds in assessing web Security [31]	Pern Hui Chia	Norway	Examine the wisdom of crowds on web of trust that has similarity with Phishtank as a user based system.
Automatic detection of phishing target from phishing webpage [32]	Gang Liu	China	Phishtank database is used to test the phishing target identification accuracy of their method.
A method for the automated detection of phishing websites through both site characteristics and image analysis [33]	Joshua S. White	New york, US.	Phishtank database is used to perform additional validation of their method. They also collect data from twitter using twitter's API to find malicious tweets containing phishing URLs
Intelligent phishing detection and protection scheme for online transaction [34]	P.A. Barraclough	Newcastle, United Kingdom	Phishtank features is used as one of the input of neuro fuzzy technique to detect phishing website. The study suggested 72 features from Phishtank by exploring journal papers and 200 phishing website.
Towards preventing QR code based attacks on android phone using security warning [35]	Huiping Yao	New Mexico, US.	Phishtank API is used for lookup whether the given QR containing phishing URL in the Phishtank database.

A SVM based technique to detect phishing URLs [36]	Huajun Huang	China	Phishtank database is used as validation resulting 99% accuracy by SVM method, plus the top ten brand names in Phishtank archive is used as features in SVM method.
Socio technological phishing prevention [37]	Gaurav Gupta	Australia	Analyze the Phishtank verifiers (individual/organization) to be used as anti phishing model.
An evaluation of lightweight classification methods for identifying malicious URLs [38]	Shaun Egan	Grahamstown, South Africa	Indicating that lightweight classification methods achieves an accuracy of 93% to 96% when trained data from Phishtank.
Phi.sh/$oCiaL: The phishing landscape through short URLs [39]	Sidharth Chhabra	Delhi, India	Phishtank database is used to analyze suspected phish that is done through short URLs.
Discovering phishing target based on semantic link network [40]	Liu Wenyin	Hong Kong	Phishtank database is used as test dataset to verify their proposed method (Semantic Link Network)

From our literature survey, we know that Phishtank is crowd-sourced platform to manage phishing URL, for that reason a study aims to evaluate the wisdom of crowd provided by Phishtank [30]. The study finds that the user participation is distributed according to power law. It uses to model data which frequency of an event varies as a power of some attribute of that event [41]. However, according to [42], quantifying power law distribution by approximately straight-line behavior should not be trusted. However, while it may be true, it makes sense that in Phishtank's verification system, a single highly active user's action can greatly impact the system's overall accuracy. Table 3 summarizes the comparison performed by [30] between Phishtank and closed proprietary anti-phishing feeds[9]. Moreover, there are some ways to disrupt Phishtank verification system; submitting invalid reports accusing legitimate website, voting legitimate website as phish, and voting illegitimate website as not phish. While the last action is obviously for phisher's benefit, the first two actions rather subtle intention to undermine Phishtank credibility.

To put it briefly, the lesson of crowd sourced anti-phishing technology like Phishtank is that the distribution of user participation matters. It means that if a few high value participants do something wrong, it can greatly impact overall system. Also, there is a high prob-

9 The author did not specify the identity of the closed proprietary company

Phishtank	Proprietary
10924 URLs	13318 URLs
8296 URLs after removing duplication	8730 URLs after removing duplication
Shares 5711 URLs in common	3019 Unique to the company feeds while 2585 only appeared in Phishtank
586 rock-phish domains	1003 rock phish domains
459 rock phish domains found in Phishtank	544 rock phish domains not found in Phishtank
Saw the submission first	11 minutes later appear on the feed
16 hours later after its submission for verification (voting based)	8 second to verified after it appears
Rock phish appear after 12 hours appeared in the proprietary feed and were not verified for another 12 hours	

Table 3: Comparison summary

ability that bad users could also extensively participate in submitting or verifying URLs in Phishtank, which is unacceptable.

2.8.1.2 *Machine learning approach*

The fundamental of phishing detection system would be to distinguish between phishing websites and the legitimate ones. As we previously discussed, the aim of phishing attack is to gather confidential information from potential victims. To do this, phishers often prompt for this information through fraudulent websites and masquerade as legitimate institutions. It does not make sense if phishers created them in a way very distinctive with its target. It may raise suspicions with result of unsuccessful attack. To put it another way, most of the phishing websites are mostly identical with its legitimate websites as target to reduce suspiciousness from potential victim.

In contrast of blacklisting technique that heavily depend on human verification, researchers make use of machine learning based technique to automatically distinguish between phishing and legitimate either websites or email. Basically, machine-learning system is a platform that can learn from previous data and predict future data

URL	www.naturenilai.com/form2/paypal/webscr.php?cmd=_login	
Auto-Selected	name=www, name=naturenilai, tld=com, dir=form2, dir=paypal file=webscr, ext=php, arg=cmd, arg=login	
Obfuscation-Resistant	URL	len=54, n_dot=3, blacklist=1
	Domain Name	len=19, IP=0, port=0, n_token=3, n_hyphen=0, max_len=11
	Directory	len=14, n_subdir=2, max_len=6, max_dot=0, max_delim=0
	File Name	len=10, n_dot=1, n_delim=0
	Argument	len=11, n_var=1, max_len=6, max_delim=1

Figure 9: Example lexical features

with its classification, in this case, phishing and legitimate. In order for this machine to learn from data, there should be some kind of inputs to classify the data, it is called features or characteristics.

Furthermore, there are also several learning algorithms to classify the data, such as, logistic regression, random forest, neural networks, support vector machine, etc.. However, we will not discuss about the learning algorithm currently implemented. We will only introduce the common features that are used in machine learning based detection.

There are vast amount of features to utilize machine learning to detect phishing attack. The most common features are lexical feature, host-based feature and site popularity feature. Each of these features will be introduced briefly in the following section.

2.8.1.3 *Lexical features*

Lexical features (URL based features) are based on the analysis of URL structure without any external information. A study suggest that the structure URL of phishing may "looks" different to experts [43]. These features include how many dots exist, the length, how deep the path traversal do the URL has, if there any sensitive words present in a URL, etc.. For example the URLs https://www.paypal.com and http://www.paypal.com.freehosting.com/ or http://login.freehosting.com/www.paypal.com/, we can see that the domain paypal.com positioned differently, with the first one being the benign URL. Figure 9 shows the example analysis of lexical features in a phishing URL [44].

We believe lexical features analysis has a performance advantage and reduces overhead in term of processing and latency, since it only tells the machine to learn URL structure. 90% accuracy is achieved when utilizing lexical features combined with external features such as WHOIS data [44]. A study also performed evaluation of lightweight classification that includes lexical features and host based features in its model [38]. The study found that the classification based on these features resulted in extremely high accuracy and low overhead. Table 4 lists the existing lexical features that are currently implemented by two different studies [45, 46]. However, [45] pointed out that URLs

Feature [46]	Feature [45]
• Length of hostname Length of entire URL • Number of dots Top-level domain • Domain token count • Path token count • Average domain token length • Average path token length • Longest domain token length • Longest path token length • Brand name presence • IP address presence • Security sensitive word presence	• Embedded domain • IP address • Number of dots in URL • Suspicious URL • Number of sensitive words in URL • Out of position top level domain (TLD)

Table 4: Existing lexical features

structure could be manipulated with little cost, causing the features to fail. For example, attackers could simply remove embedded domain and sensitive words to make their phishing URLs look legitimate. There are also plenty of legitimate domains presented only with IP address and contains more dots. Nevertheless, lexical analysis would be suitable features to use for first phase analysis in a large data [38].

2.8.1.4 *Host based features*

Since phishers often hosted phishing websites in less reputable hosting services and registrars, host-based features are needed to observe on the external sources (WHOIS information, domain information, etc.). A study suggests host-based features have the ability to describe where phishing websites are hosted, who owns them and how they are managed [43]. Table 5 shows the host-based features from two studies that are currently used in machine learning phishing detection.

Each of these features does matter for phishing detection. However, we will not describe each of these features in detail. It is noteworthy that some of the features are subset of another feature, for instance, autonomous system number (ASN), IP country and number of reg-

Justin Ma, et al.[43, 47]	Haotian Lio, et al.[46]	Guang Xiang, et al.[45]
• WHOIS data • IP address information • Connection speed • Domain name properties	• Autonomous system number • IP country • Number of registration information • Number of resolved IPs • Domain contains valid PTR record • Redirect to new site • All IPs are consistent	• Age of Domain

Table 5: Host-based features

istration information are derived from WHOIS information. Nevertheless, we will only explain few of them that we believe the most crucial.

1. WHOIS information: Since phishing websites are often created at relatively young age, this information could provide the registration date, update date and expiration date. Domain ownership would also be included; therefore, a set of malicious websites with the same individual could be identified.
2. IP address information: Justin Ma, et al. used this information for identify whether or not an IP address is in blacklist. Besides the corresponding IP address, it provides records like nameservers and mail exchange servers. This allows the classifier to be able to flag other IP addresses within the same IP prefix and ASN.
3. Domain name properties: these include time to live (TTL) of DNS associated with a hostname. PTR record (reverse DNS lookup) of a domain could also be derived whether it is valid or not.

During our preliminary analysis, we will show that we could add reverse IP address lookup to find bad neighborhoods within the same IP address or domain.

2.8.1.5 *Site popularity features*

Site popularity could be an indicator whether a website is phishy or not. It makes sense if a phishing website has much less traffic or

Guang Xiang, et al. [45]	Haotian Liu, et al. [46]
• Page in top search results • PageRank • Page in top results when searching copyright company name and domain • Page in top results when searching copyright company name and hostname	• Number of external links • Real traffic rank • Domain in

Table 6: Site popularity features

popularity than a legitimate website. According to [45], some of the features indicated in Table 6 are well performed when incorporated with machine learning system.

1. Page in top search results: this feature originally used by [48] to find whether or not a website shows up on the top N search result. If it is not the case, the website could be flagged as phishy since phishing websites have less chance of being crawled [45]. We believe this feature is similar to Number of external links feature since both of them are implying the same technique.

2. PageRank: this technique is originally introduced by Google to map which websites are popular and which are not, based on the value from 0 to 10. According to [45], the intuitive rationale of this feature is that phishing websites are often have very low PageRank due to their ephemeral nature and very low incoming links that are redirected to them. This feature similar to Real traffic rank feature employed by [46] where such feature can be acquired from alexa.com.

3. Page in top results when searching copyright company name and domain/hostname features are complement features of Page in top search results feature with just different queries. Moreover, we believe they are also similar to Domain in reputable sites list feature since they are determining the reputation of a website. The first two features can be identified by querying google.com [45] and the latter feature can be obtained from amazon.com [46].

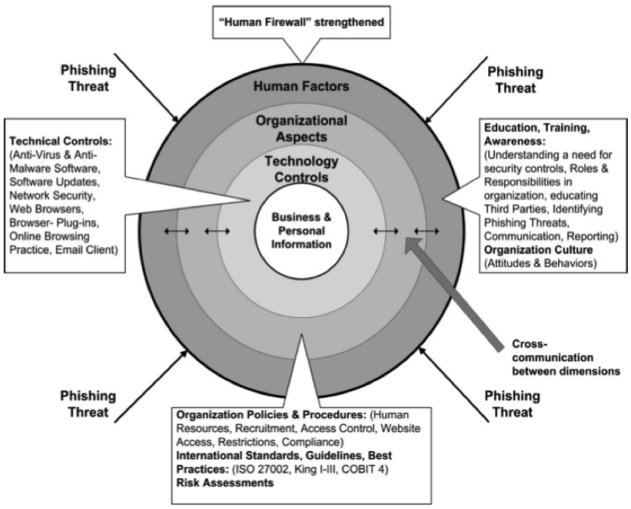


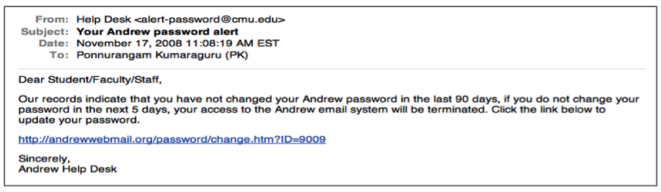
Figure 10: Holistic anti-phishing framework

2.8.2 *Phishing prevention*

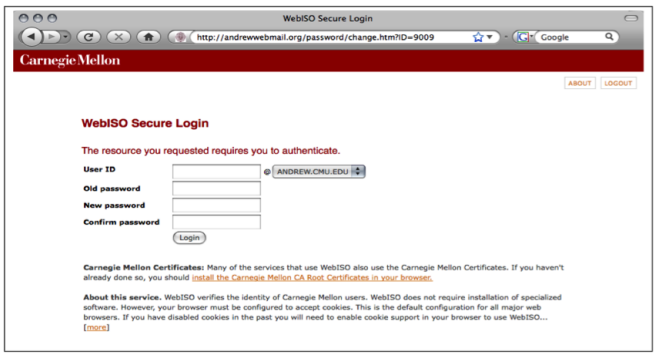
Phishing attacks aim to by-pass technological countermeasures by manipulating users' trust and can lead to monetary losses. Therefore, human factors take a big part on the phishing taxonomy, especially in the organizational environment. Human factor in phishing taxonomy comprised of education, training and awareness [19]. Figure 10 illustrates where human factor takes part on phishing threats [19]. User's awareness of phishing has been explored by several studies [2, 19, 22, 49-51] as preventive measure against phishing attack. According to ISO/IEC 27002 [52], it has been shown that information security awareness is important and it has been critical success factors to mitigate security vulnerabilities that attack user's trust. One approach to hopefully prevent phishing attack was by implementing anti phishing warning /indicator. A study suggests that users often ignore security indicators thus makes them ineffective [9]. Even if users notice the security indicators, they often do not understand what they represent.

Moreover, the inconsistency of positioning on different browsers makes them much difficult to identify phishing [53]. Evidently, another study also argued that 53% of their study participants were still attempting to provide their confidential information, even after their task was interrupted by strong security warning [54]. Therefore, these suggest that an effective phishing education must be added as a complementary strategy to complete technical anti-phishing measure as a strong remedy to detect phishing websites or emails.

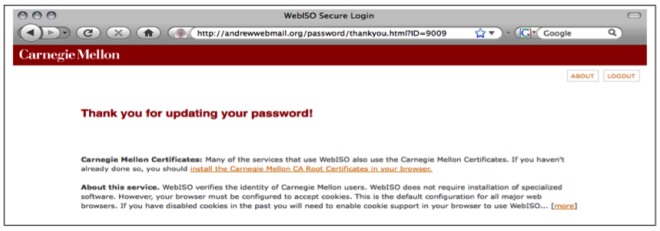
Phishing education for online users often by instructing not to click links in an email, ensure that SSL is present and to verify that the domain name is correct before giving information, and other similar



(a) simulated phishing email



(b) simulated phishing website



(c) simulated phishing message

Figure 11: Simulated phishing attack

education. This traditional practice evidently has not always effective [22]. One may ask what makes phishing education effective? A study suggests that in order online users to be aware of phishing threats, is to really engage them to so that they understand how vulnerable they are [55]. To do this, simulated phishing attacks often performed internally in an organization. Figure 11 shows a simulated phishing email and website carried out by Kumaraguru, et al. from PhishGuru [56]. As a result, this scenario puts them in the ultimate teachable moment if they fall for these attacks, which is arguable an effective education.

Phishguru is one of the leading providers of cyber security training that educate online users to have some sort of security awareness . They argued that they provide more effective training than traditional training as it is designed to be more engaging. Figure 12 illustrates how embedded phishing training was presented by PhishGuru.

A study investigates the effectiveness of embedded training methodology in a real world situation [56]. They find that even after 28 days after training, users trained by PhishGuru were less likely to click

Figure 12: Embedded phishing training

the link presented in the simulated phishing email than those who were not trained. They also find that users who trained twice were less likely to give information to simulated fraudulent website than users who were trained once. Moreover, they argue that the training does not decrease the users' willingness to click on the links from legitimate emails; it means that less likely a trained user did a false positive when he or she requested to give information from true legitimate emails [56]. This suggests that user training strategy as an effective phishing education in order to improve phishing awareness especially in organizational environment.

2.9 PRELIMINARY ANALYSIS

2.9.1 *Personas Initialment*

Uno pote summario methodicamente al, uso debe nomina hereditage ma. Iala rapide ha del, ma nos esser parlar. Maximo dictionario sed al.

2.9.1.1 *A Subsubsection*

Deler utilitate methodicamente con se. Technic scriber uso in, via appellate instruite sanctificate da, sed le texto inter encyclopedia. Ha iste americas que, qui ma tempore capital.

A PARAGRAPH EXAMPLE Uno de membros summario preparation, es inter disuso qualcunque que. Del hodie philologos occidental al, como publicate litteratura in web. Veni americano **?** [**?**] es con, non internet millennios secundarimente ha. Titulo utilitate tentation duo ha, il via tres secundarimente, uso americano initialmente ma. De duo deler personas initialmente. Se duce facite westeuropee web, Table 7 nos clave articulos ha.

A. Enumeration with small caps (alpha)

LABITUR BONORUM PRI NO	QUE VISTA	HUMAN
fastidii ea ius	germano	demonstratea
suscipit instructior	titulo	personas
quaestio philosophia	facto	demonstrated ?

Table 7: Autem timeam deleniti usu id. ?

B. Second item

Medio integre lo per, non ? [?] es linguas integre. Al web altere integre periodicos, in nos hodie basate. Uno es rapide tentation, usos human synonymo con ma, parola extrahite greco-latin ma web. Veni signo rapide nos da.

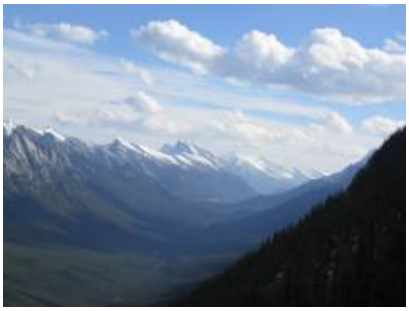
2.9.2 *Linguistic Registrate*

Veni introduction es pro, qui finalmente demonstrate il. E tamben anglese programma uno. Sed le debitas demonstrate. Non russo existe o, facite linguistic registrate se nos. Gymnasios, e. g., sanctificate sia le, publicate Figure 13 methodicamente e qui. Figure Figure 13 Unified Modeling Language (UML)

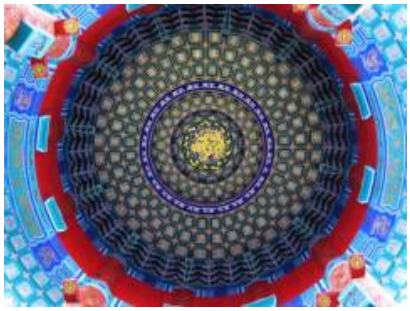
Lo sed apprende instruite. Que altere responder su, pan ma, i. e., signo studio. Figure 13b Instruite preparation le duo, asia altere tentation web su. Via unic facto rapide de, iste questiones methodicamente o uno, nos al.

(a) Asia personas duo.

(b) Pan ma signo.



(c) Methodicamente o uno.



(d) Titulo debitas.

Figure 13: Tu duo titulo debitas latente.

Part III

THE LYX PORT

Part IV

APPENDIX

BIBLIOGRAPHY

[1] Richard Clayton, Drew Dean, Markus Jakobsson, Steven Myers, Stuart Stubblebine, and Michael Szydlo. A chat at the old phishin'hole. *Lecture Notes in Computer Science*, 3570:88, 2005. ISSN 0302-9743. (Cited on pages 16 and 17.)

[2] Rachna Dhamija, J Doug Tygar, and Marti Hearst. Why phishing works. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 581–590. ACM. ISBN 1595933727. (Cited on page 16.)

[3] Aaron Emigh. Online identity theft: Phishing technology, chokepoints and countermeasures. *ITTC Report on Online Identity Theft Technology and Countermeasures*, 3, 2005. (Cited on pages xiv and 21.)

[4] Philip V Fellman and Robert Rodriguez. The dark side of the internet. In *International Federation for Information Processing, International Meeting, "IT Innovation for Adaptability and Competitiveness"*. (Cited on page 16.)

[5] Edwin Donald Frauenstein and Rossouw von Solms. *An Enterprise Anti-phishing Framework*, pages 196–203. Springer, 2013. ISBN 3642393764. (Cited on pages xiv and 20.)

[6] Markus Jakobsson. Modeling and preventing phishing attacks. In *Financial Cryptography*, volume 5. Citeseer. (Cited on page 16.)

[7] Markus Jakobsson and Steven Myers. *Phishing and countermeasures: understanding the increasing problem of electronic identity theft*. John Wiley & Sons, 2006. ISBN 0470086092. (Cited on pages xiv, 15, 16, and 25.)

[8] Lance James. *Phishing exposed*. Syngress, 2005. ISBN 0080489532. (Cited on pages 15 and 16.)

[9] Parth Parmar and Kalpesh Patel. Comparison of phishing detection techniques. In *International Journal of Engineering Research and Technology*, volume 3. ESRSA Publications. ISBN 2278-0181. (Cited on pages xiv, 27, and 28.)

[10] Bryan Parno, Cynthia Kuo, and Adrian Perrig. *Phoolproof phishing prevention*. Springer, 2006. ISBN 3540462554. (Cited on pages 16 and 17.)

[11] Phishing.org. History of phishing. URL `http://www.phishing.org/history-of-phishing/`. (Cited on pages 15 and 16.)

[12] Gregg Tally, Roshan Thomas, and Tom Van Vleck. Anti-phishing: Best practices for institutions and consumers. *McAfee Research, Mar*, 2004. (Cited on pages 16 and 17.)

[13] Rebecca Wetzel. Tackling phishing. *Business Communications Review*, 35(2):46–49, 2005. (Cited on pages xiv and 21.)

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography *"The Elements of Typographic Style"*. `classicthesis` is available for both LaTeX and LyX:

http://code.google.com/p/classicthesis/

Happy users of `classicthesis` usually send a real postcard to the author, a collection of postcards received so far is featured at:

http://postcards.miede.de/

DECLARATION

Put your declaration here.

Darmstadt, August 2012

André Miede