A recent review of conventional vs. automated cybersecurity anti-phishing techniques *

Issa Qabajeh[a], Fadi Thabtah[b], Francisco Chiclana[a]

[a] Centre for Computational Intelligence, Faculty of Technology, De Montfort University, Leicester, UK
[b] Applied Business and Computing, NMIT, Auckland, New Zealand

Abstract

In the era of electronic and mobile commerce, massive numbers of financial transactions are conducted online on daily basis, which created potential fraudulent opportunities. A common fraudulent activity that involves creating a replica of a trustful website to deceive users and illegally obtain their credentials is website phishing. Website phishing is a serious online fraud, costing banks, online users, governments, and other organisations severe financial damages. One conventional approach to combat phishing is to raise awareness and educate novice users on the different tactics utilized by phishers by conducting periodic training or workshops. However, this approach has been criticised of being not cost effective as phishing tactics are constantly changing besides it may require high operational cost. Another anti-phishing approach is to legislate or amend existing cyber security laws that persecute online fraudsters without minimising its severity. A more promising anti-phishing approach is to prevent phishing attacks using intelligent machine learning (ML) technology. Using this technology, a classification system is integrated in the browser in which it will detect phishing activities and communicate these with the end user. This paper reviews and critically analyses legal, training, educational and intelligent anti-phishing approaches. More importantly, ways to combat phishing by intelligent and conventional are highlighted, besides revealing these approaches differences, similarities and positive and negative aspects from the user and performance prospective. Different stakeholders such as computer security experts, researchers in web security as well as business owners may likely benefit from this review on website phishing.

Keywords- Classification, Computer Security, Phishing, Machine Learning, Web Security, Security Awareness

1. Introduction

As the largest computer network, the Internet is considered a critical platform for business success and expansion as most commercial trades are being conducted online. With the majority of businesses competing in a global market, they seek to maximise revenue by increasing user accessibility and promotion of products and services through the web-based and mobile platforms. People not only use the internet for socialisation and knowledge but purchasing goods, pay bills, and transfer money, and consequently the internet has become a necessity for many individuals. Since the explosion of mobile commerce applications access to the World Wide Web has

* Accepted for publication in **Computer Science Review** on 28th May 2018.

become a daily essential requirement, with people performing regular financial transactions regardless of their geographical location (Ronald et al., 2007; Aburrous et al., 2010b).

With the advanced development of computer hardware, especially computer networks and cloud technology services, online and mobile commerce have significantly increased in the last few years (Abdelhamid & Thabtah, 2014). Indeed, the number of customers who perform online purchase transactions has dramatically increased and large monetary values are daily exchanged through electronic means, such as private payment gateways, that are usually verified by secure socket layer (SSL) (Sheng et al., 2010). Despite the convenience associated with online transactions from both user and business prospective, an online threat has emerged: phishing.

Phishing involves creating a well designed website that replicates an existing authentic business website in order to deceive users and obtain illegally their credentials and/or login information (Abdelhamid, 2015), so as for phishers to get access to legitimate users' financial information (Afroz, et al., 2011). Unfortunately, the consequences of phishing are fatal because affected legitimate users become vulnerable to identity theft and information breach and no longer trust online commerce and electronic banking (Nguyen, et al., 2015). Phishing typically occurs via email sent to users, from apparent trustworthy sources, urging them to adjust their login information by clicking/following a hyper link within such email (Khadi & Shinde, 2014).

In 2011, Gartner Group published a report (McCall, 2011) that revealed an annual loss over $2.8 billion as a result of phishing activities occurring within the United States. For this reason, an international body that aims to minimise online threats including pharming, spoofing, phishing and malware, the Anti-Phishing Work Group (APWG), was created (Jevans, 2003). APWG periodically disseminates reports for the online community on recent cyber-attacks, with a recent report stating the rapid increase of phishing websites to 17,000 in the month of December 2014 alone (Aaron & Manning, 2014). In the same report, the total number of monthly phishing activities reported in the fourth quarter of 2014 was more than 197,000 websites, and increase of 18% compared with the previous quarter of the same year, with the United States continued to be the top country hosting phishing websites.

It seems imperative that users, as well as businesses, adopt renewable anti-phishing tools or strategies to reduce phishing activities and protect themselves from their potential negative impacts. This is important because phishing attacks are constantly changing and new deceptions are emerging all the time. Anti-phishing solutions adopting DM (ML) are shown to be more practical and effective in combating phishing because they work automatically and are capable of revealing concealed knowledge that online users are not aware of, especially with respect to the relationship among website features and phishing activities. This hidden knowledge, when combined with human experience, can result in an effective shield for protecting users from phishing (add a reference).

In this paper, we investigate the phishing problem and define it in a classification ML context. We then discuss common, traditional, strategies in addition to computerised techniques developed to combat phishing. More importantly, the paper thoroughly investigates traditional and ML anti-phishing classification techniques and critically analyses their benefits and disadvantages theoretically. There have been few former reviews on phishing such as (Suganya, 2016; Mohammad, et al., 2015; Sahu and Dubey, 2014; Basnet, et al., 2008) among others. However, most of these reviews have covered partly one or more of phishing aspects. For instance, Suganya, (2016) and Sahu & Dobey, (2014) briefly reviewed phishing attacks without showing the ways to

combat them or their pros and cons. Mohammad, et al., (2015) discussed in general common solutions of website phishing without providing grounds for recommendations besides not covering specific intelligent approaches. Lastly, Basnet, et al., (2008) compared only few intelligent anti-phishing solutions without on elaborating the other computerised and classic approaches of anti-phishing. Therefore, this article not only comprehensively reviews phishing from wider prospective but also it critically analyses traditional and automated anti-phishing solutions.

This paper serves researchers, organisations' managers, computer security experts, lecturers, and students who are interested in understanding phishing and its corresponding intelligent solutions. This is since wider potential solutions have been critically analysed and experimentally compared besides presenting classic solutions including educational, legal, and software based. This paper is structured as follows: Section 2 presents the phishing problem, its history, and its lifecycle. Section 3 critically analyses common classic methods of combating phishing besides critically analysing them. Section 4 is devoted to intelligent anti-phishing solutions that employ different strategies in deriving the anti-phishing models. Section 5 provides the conclusions.

2. Phishing Background

2.1 Phishing History

Phishing comes from the word "fishing," in which the phisher throws a bait and awaits for potential users to take a bite. Phishing is not recent as an online risk, with its origin rooted in a social engineering method using telephones known as "phone phreaking" (Rader & Rahman, 2015). It was during the 1990s period when the internet community started to grow that phishing was originally observed as an online threat, especially in the United States (Basnet, et al., 2008).

According to McFredies (2016), the first phishing incident was noticed in the mid-1990s when phishers attempted to obtain registered online users' account information from the internet provider America Online (AOL). Phishers during this era frequently utilised instant messages (IM) in AOL chat rooms or emails to reach users so that they would reveal their passwords (Figure 1 illustrates an example of an early phishing attack),

Fig. 1. Example of an early phishing attack (Blogonlymyemail.com)

which were subsequently used by phishers to leverage the victims' accounts and begin emailing spam to other online users. Obviously, phishers realised that they could further trick victims if the IMs and emails requested them to update their billing information. With this realisation, the attackers expanded their aim and using the same electronic means (IMs and emails) attempted to access other financial information from victims such as social security numbers, addresses, credit card information, etc.

One of the most common beliefs that ordinary users have about phishing websites is that grammatical errors and typos are typical within these websites (Rader & Rahman, 2015). This has misled users into believing that a website without grammatical errors must be trustworthy, which has proven not necessarily be true. Phishers in today's cyber world become more innovative and work systematically in groups sometimes even orchestrating phishing campaigns motivated by potential financial gains. In fact, phishers are continuously changing their spoofing methods based on the counterpart security measures taken by organizations, and recently they have directed their attention to the mobile commerce platform. This means that the profile of phishers has changed from egocentric purposes into more organized and serious cybercrime that keeps evolving, which makes it hard to detect.

2.2 Phishing Process

Phishing attacks are initiated through an email sent to potential users. Other ways a phisher may start an attack include Instant Messaging, online blogs and forums, short message services, peer to peer file sharing services, and social media websites (Abdelhamid, et al., 2014). We can summarise the phishing lifecycle as follows (see Figure 2):

1) A link is sent using one of the aforementioned channels to potential victims.
2) When clicked, the link will redirect potential victims to a malicious website.
3) Users become vulnerable as they try to login using their credentials on the malicious website.
4) Login credentials are then transferred to a server, or a key logger is installed into the user's computing device.
5) The phisher can then utilise the credentials to perform additional cybercrimes.

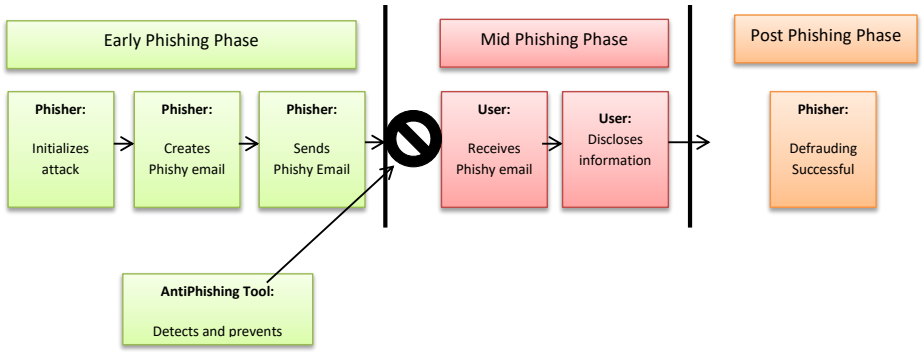

Fig. 2 Phishing life cycle (Abdelhamid, et al., 2014)

2.3 Phishing as A Classification Problem

Generally speaking, websites can be classified by hand-crafted methods based on certain features such as URL length, prefix_suffix, domain, sub_domain, etc. Initially, scholars in the area of online security (Aburrous et al., 2008; 2010b) developed different knowledge bases using their experience and expertise to distinguish phishing from legitimate websites. Recently, there have been studies and proposals aiming at deriving intelligent rules to detect the fine line between legitimate and phishing websites using statistical analysis (Qabajeh and Thabtah, 2015; Mohammad et al., 2014b; Abdelhamid et al., 2014). For instance, Aburrous et al. (2010a) and Mohammad et al. (2014b) defined a number of hand crafted rules based on various website features using simple statistical analysis on websites (instances) collected from different sources including Phishtank and Yahoo directory (Phishtank, 2011). More advanced decision rules have been developed in Abdelhamid et al. (2014) in which the authors used further statistical analysis on a larger phishing dataset collected from varying sources.

ML and DM have proved to be powerful data analysis tools in many application domains such as medical diagnosis, market basket analysis, weather forecasting and events processing, to cite some (Abdelhamid & Thabtah, 2014). This is due to the fact that ML and DM techniques usually reveal concealed meaningful information from large datasets so they can be utilised in management decisions related to development, planning, and risk management. Generally speaking, ML and DM can be seen as an automated and intelligent tool embedded within management information systems to guide decision making processes in both business and scientific domains. Common tasks or problems that ML and DM handle are clustering, association rule discovery, regression analysis, classification, pattern recognition, time series analysis, trends analysis, and multi-label learning (Witten & Frank, 2005).

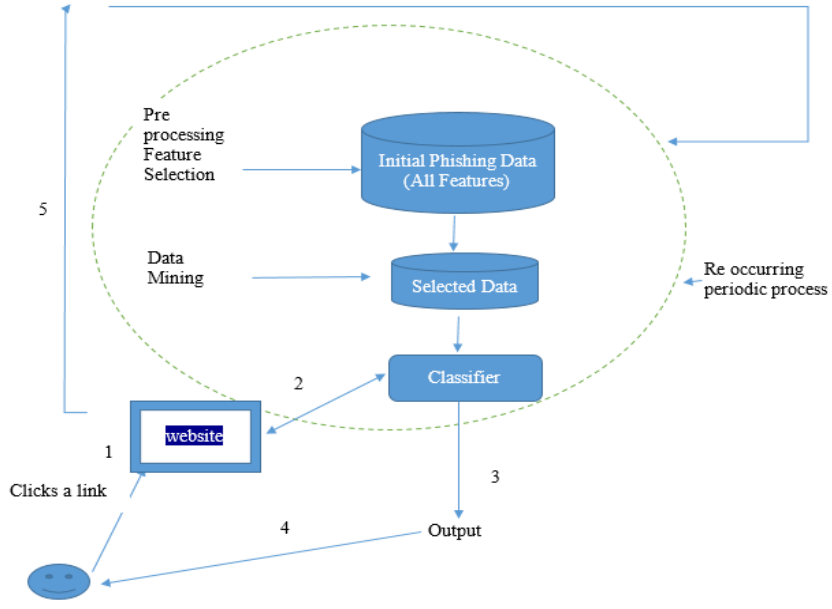

Figure 3. Phishing as a classification process

One of the frequent task of ML is the forecasting of a target variable within datasets based on other available variables (Witten & Frank, 2005). This forecasting process occurs in an automated manner using a classification model, normally named the classifier, that is derived from a labelled training dataset. The goal of the classifier is to "guess" the value of the target variable in unseen data, referred to as the test dataset, as accurately as possible. This task description falls under the umbrella of supervised learning and is known as classification. Abdelhamid and Thabtah (2014) defined classification as the ability to "accurately" predict class attributes for a test instance using a predictive model derived from a training dataset.

Since the problem of website phishing involves automatic categorisation of websites into a predefined set of class values (legitimate, suspicious, phishy) based on a number of available features (variables) then this problem can be considered a classification problem. To be more specific, the training dataset will consist of a set of predefined features and the class attribute and instances are basically the websites' feature values. These instances can be extracted from different sources such as Phishtank and online directories. The aim will be to build an anti-phishing classifier that can predict the type of website based on hidden knowledge discovered from the training set features during the data processing phase. Usually the goodness of the classifier is measured using accuracy, which primarily relies on the correlations of the features and the class (Thabtah et al., 2016a). Figure 3 shows phishing as a classification problem from the ML prospective. As discussed before, phishing websites are dynamics and consequently it can modelled as a dynamic supervised learning classification problem. Therefore, an effective anti-phishing classifier should be adaptable to any new features observed in order to handle and manage the dynamic nature of the problem.

In this review paper, we focus on content based methods that fall under website phishing attacks as shown in Fig. 4. We omit non-content based methods such as domain popularity, restricted from filing, DSN based features, water marking, one time password, layout similarity and crowd sourcing among others. Also email finishing techniques are omitted from the graph since they are out of scope.

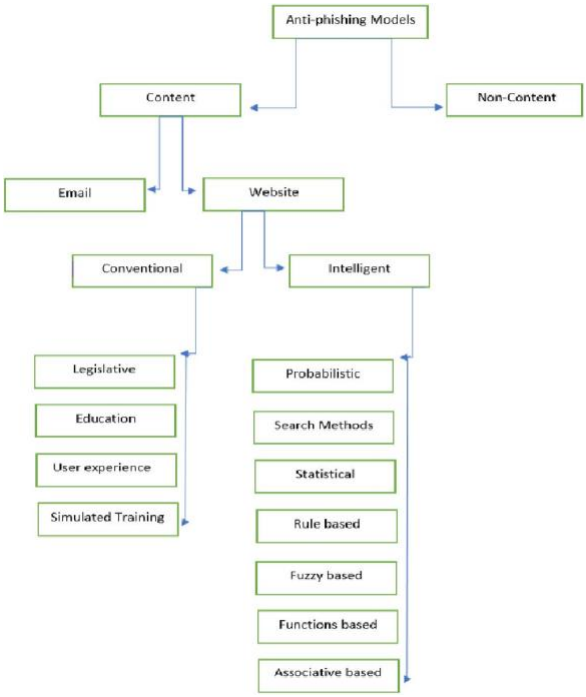

Fig 4. Taxonomy of website phishing approaches.

3. Common Traditional Anti-Phishing Methods

Since phishing causes serious breaching of user confidentiality, as well as organizations including government agencies, there have been different methods proposed to combat phishing. These approaches can be categorized into three main categories:

- Education and legal
- Computerized using human-crafted methods
- Intelligent ML methods

In this section, we examine the literature on phishing and critically analyse different techniques based on the above categories. Focus, however, will be on the intelligent anti-phishing solutions since it is believed to be the way forward in shielding the web from phishing threats and promising results have recently been derived by this category in Thabtah, et al. (2016b), Nguyen, et al. (2015), Lee, et al. (2015), Mohammad, et al. (2014a), Jameel and George (2013), Khadi and Shinde (2013), and Sheng, et al. (2010), among others.

3.1 Legal Anti-phishing Legislations

Governments have been slow in responding and opposing started to oppose phishing. California State in the US was the first to issue anti-phishing legislation in 2005 (InformationWeek, 2016). This legislation stated that it is unlawful to use any electronic means such as websites, emails, or any other methods to ask or solicit information from online users by claiming ones self as a business without the authority of that business. Other US States such as Texas have also introduced new cybercrime legislations that include phishing, and in 2005 the General Assembly of Virginia added phishing attacks to their list of computer crimes (General Assembly of Virginia, 2005). These new laws empowered companies such as America Online to file lawsuits in Virginia against phishers in 2006 (Pike, 2006). However, most states in US have not legislated specific laws incriminating phishing and usually prosecute phishers using other computing crime laws such as fraud.

At Federal level in US, lawmakers and congressional representatives have not passed anti-phishing legislation either. There were a few attempts between 2004 and 2006 following the Anti-Phishing Act of 2004 to pass specific bills incriminating phishing and instigating tougher prison sentences, but these bills were stopped at the committee level in Congress. Nevertheless, federal law enforcement can incriminate phishers using other laws that are related to identity theft and fraud such as "18 U.S.C. section 1028" (Pike, 2006). Businesses have also joined the Government in fighting phishing. For example, in 2005 Microsoft filed over 115 lawsuits in Washington's Western District Court accusing a single Internet user of utilising various deceptive methods to access some of the company's users' information (add reference). In mid-2006, the then president George W. Bush established a new cybercrime identity theft task force (Executive-Order-13402, 2006), with a single goal: reduce the risks of cybercrime, especially phishing.

The United Kingdom (UK) has followed the US by strengthening its legal system against severe cybercrimes, including fraud and identity theft. In 2006, the UK introduced the new Fraud Act, which increased prison sentences to up to ten-years for online fraud offences (BBC News, 2005). This same act prevents possession of a phishing website with the intent to deceive users and commit fraud. Further, Microsoft decided to collaborate with

other law enforcement agencies outside US to bring justice to phishers. In doing so, the company signed an agreement with the Australian government to train law enforcement agents in preventing phishing (Government of Australia, 2011). Also, in 2010 Canada introduced an Anti-spam Act that incriminates cybercrime and that aims to protect Canadian online consumers and businesses when globally trading (ClickDimensions, 2014).

3.2 Simulated Training

One of the easy, yet helpful, policies to oppose cybercrimes is to educate users on the ways employed to access their information. When novice users are aware of the circumstances around phishing, they may be able to minimise this risk or stop it as early as possible. Unfortunately, ordinary web browsing users are unaware of how phishing attacks start or how visually to recognise an untruthful website and differentiate it from one that is trustworthy (Mohammad, et al., 2015). Moreover, basic security indicators and anti-phishing software counterparts are still vague for many online shoppers (Qabajeh & Thabtah, 2015). Subsequently, these increase the pace of phishing and motivate phishers to launch further attacks. For instance, a security survey was conducted by Julie, et al. (2007), which revealed the lack of knowledge on cybercrimes, including phishing, held by online users. In addition, some respondents in the survey showed security awareness yet were reluctant in using their financial information for payment purposes, even within trustworthy websites.

There have been a number of studies on educating people as to the severity of phishing. For example, Arachchilage and Love (2013) investigated whether mobile games can be a helpful method for raising awareness of phishing attacks. The authors evaluated learning curves of users who played with a mobile game about phishing developed by Arachchilage and Cole (2011) and assessed whether an interactive mobile platform is effective in educating users in contrast to traditional security training. A comparison of user responsiveness to phishing has also been conducted using the developed mobile game, along with a website designed by APWG. The results showed that users who played the anti-phishing mobile game were able to spot non-genuine websites with a higher rate of accuracy than other users who only used the APWG website.

There are a number of organizations and research studies, such as Arachchilage, et al. (2007) and Ronald, et al. (2007), that have adopted a relative training to warn users of phishing. This training involves sending participants simulated malicious emails from a genuine source to evaluate their exposure to phishing. At the end of the training, participants are given the training material and informed about their vulnerability to phishing.

Embedded training is another way to measure a users' vulnerability to phishing. This training often mimics primary daily business processes performed by employees by experimenting to measure a certain outcome (Arachchilage, et al., 2013). In phishing, the authors of Arachchilage, et al. (2007) used the embedded training methodology to measure phishing awareness at a university. The authors sent malicious emails from the administrator to participants without informing them of the training material content. These emails urged users to click on a link that would redirect them to a malicious website where they would input their login credentials. This aim was directed at identifying the number of users who would actually click on the link. During the experiment, the user was interrupted immediately when he clicked the link and was then provided with the training material. The embedded training proposed by the authors was based on a preliminary pilot study conducted by them on a limited number of university students.

3.3 User Experience: Anti-phishing Online Communities

One of the approaches to reducing the impact of phishing on online users and organisations is to build an anti-phishing community to monitor recent phishing activities and provide news to the different stakeholders. Users' experiences are practical and based on real cases related to different types of phishing. Such efforts by users and organizations have resulted in new proactive online communities and data repositories. These accumulated and useful resources are of interest since they can be employed to study ways to make the Internet safer and free from phishing.

The Monitoring and Takedown (MaT) approach enables individuals who recognise phishing activity to report it via public anti-phishing communities including APWG, PhishTank, Millersmiles, and Symantec among others (Jevans, 2003; PhishTank, 2011; Bright, 2011; Nahorney, 2011). These anti-phishing communities allow users to report phishing content and warn other users and organisations as well. Users can also report phishing content to the Federal Trade Commission's Complaint Department, becoming directly part of the campaign toward combating phishing. Many reputable companies also have an internet fraud department that allows users to report any fraudulent or suspicious activity such as phishing. PhishTank was created in 2003 as a subsidiary of OpenDNS in order to provide the parent company, as well as the online community, with a phishing repository. This large collection of stored phishing websites has given computer security experts, users, researchers and business owners' extensive information about phishing attacks and the features of their associated emails and websites. Another example of a good use of user experiences is Cloudmark, which is an alerting-based anti-phishing method with user rating system (Cloudmark, 2002). When a user is visiting a website and experiencing any kind of threat, they can then rate that website to alert other online users. Finally, Web of Trust (WOT) is another example of an anti-phishing approach based on the user feedback rating model (Web of Trust, 2006).

3.4 Discussion Non Intelligent Anti-phishing Solutions

Legislators in the US, UK and Canada, among others countries, have approved legislative bills that include serious jail sentences for incriminated phishers. This has been made clear in several high profile cases, especially in the US. Nevertheless, these legislative bills have not achieved a decrease of phishing attacks. On the contrary, phishing has now become more severe than ever and businesses as well as individual users have suffered from substantial financial losses as a result. One of the primary reasons for legal actions not to be as effective as expected in minimising phishing is due to the fact that often a phishing website has a short life span (normally about two days), which helps the phisher to disappear quickly once the fraud has been committed, making law enforcement difficult.

As previously mentioned, raising awareness of phishing risks and educating users has shown promising initial results (Ronald et al. 2007). Computer security scholars have adopted different ways to disseminate the seriousness phishing may cause to society, with Jevans (2003) and PhishTank (2011) using web-based material to teach novice users phishing fraud techniques; while others, such as Arachchilage et al. (2007), developing contextual and embedded trainings based on simulated phishing emails coming from genuine sources; or

educational material on phishing based on mobile games in order to increase the motivation factor among (Arachchilage et al. 2009; 2013)..

Even though educating users may positively affect the global efforts of combating phishing, this approach demands high costs and requires users to be equipped with computer security knowledge. Large organizations and governments are periodically investing in the development of anti-phishing materials in both hard and soft forms as well as websites and mobile applications. However, since phishing techniques keep changing/evolving, small to medium enterprises might not have the resources large organisations have to enable them to invest in their users' education. Therefore, a large portion of the online community realistically cannot afford the continuous additional costs to keep updating current anti-phishing material. Furthermore, phishing techniques are becoming more sophisticated because of the group efforts of phishers who employ systematic attack strategies, which make it harder for even security experts and specialised law enforcement agents to keep their skills updated. This makes ordinary users vulnerable, even if they were equipped with basic knowledge about phishing. Thus, more advanced, cheaper and intelligent approaches are needed for their implementation both within educational and legislative solutions to further reduce phishing attacks. We have seen thoughtful attempts that evolved from user experiences, user ratings, and users' social networking (such as Phishtank, Cloudmark, and APWG among others helping novice users and enterprises avoid falling prey to phishing). Effectiveness of these user community based approaches relies mainly on the following factors: (1) User experience; (2) User knowledge; (3) User honesty; and (4) Accessibility and validity of the user community's website data. Unfortunately, these factors are difficult to measure and validate, thus relying on user experience and knowledge alone necessitates carful care and accuracy. We hypothesise that by "only" considering users' experience in judging a websites' legitimacy is not enough to combat phishing, although it can be a supporting approach to a more advanced intelligent solution based on ML/DM.

4. Computerised Anti-Phishing Techniques

There has been development of anti-spam software tools that can block suspicious emails, however, these programs constantly block a large number of genuine emails and classify them as junk emails (Mohammad, et al., 2015). Emails misclassified as spam are simply false positive instances. Thus, one of the ultimate goals of the computerised anti-phishing tool is to reduce false positives and increase true positives so users can be confident of their mailbox's filter results without having to manually check their junk email folder.

4.1 Databases (Blacklist and Whitelist)

A database driven approach to fight phishing, called blacklist, was developed by several research projects (Site Advisor, 2006; Sheng, et al., 2009; Google SB, 2010). This approach is based on using a predefined list containing domain names or URLs for websites that have been recognised as harmful. A blacklisted website may lose up to 95% of its usual traffic, which will hinder the website's revenue capacity and eventually profit (Aburouse et al., 2010b). This is the primary reason that web masters and web administrators give great attention to the problem of blacklisting. According to Mohammad et al. (2015), there are two types of blacklists in computer security:

- **Domain/URL Based**. These are real time URL lists that contain malicious domain names and normally look for spam URLs within the body of emails.
- **Internet Protocol Based**. These are real time URL or domain server blacklists that contain IP addresses who, in real-time, change their status. Often, mailbox providers, such as Yahoo for example, check domain server blacklists to evaluate whether the sending server (source) is run by someone who allows other users to send from their own source.

Users, businesses, or computer software enterprises can create blacklists. Whenever a website is about to be browsed, the browser checks the URL in the blacklist. If the URL exists in the blacklist, a certain action is taken to warn the user of the possibility of a security breach. Otherwise, no action will be taken as the website's URL is not recognized as harmful. Currently, there are a few hundred blacklists which are publicly available, among which we can mention the ATLAS blacklist from Arbor Networks, BLADE Malicious URL Analysis, DGA list, CYMRU Bogon list, Scumware.org list, OpenPhish list, Google blacklist, and Microsoft blacklist (Return Path 2016). Since any user or small to large organisation can create blacklists, the currently public available blacklists have different levels of security effectiveness, particularly with respect to two factors:

1. Times the blacklist gets updated and its consistent availability.
2. Results quality with respect to accurate phishing detection rate.

Marketers, users, and businesses tend to use Google and Microsoft blacklists when compared with other publicly available blacklists commonly use because of their lower false positive rates. A study by Sheng et al. (2009) analysing blacklists concluded that they contain on average 47% to 83% phishing websites.

Blacklists often are stored on servers but can also be available locally in a computer machine as well (Mohammad et al., 2014b). Thus, the process of checking whether a URL is part of the blacklist is executed whenever a website is about to be visited by the user, in which case the server or local machine uses a particular search method to verify the process and derive an action. The blacklist usually gets updated periodically. For example, Microsoft blacklist is normally updated every nine hours to six days, whereas Google blacklist gets updated every twenty hours to twelve days (Mohammad et al., 2015). Hence, the time window needed to amend the blacklist by including new malicious URLs, or excluding a possible false positive URLs, may allow phishers to launch and succeed in their phishing attacks. In other words, phishers have significant time to initiate a phishing attack before their websites get blocked This is an obvious limitation of using the blacklist approach in tracking false websites (Abdelhamid, et al., 2014). Another study by APWG revealed that over 75% of phishing domains have been genuinely serving legitimate websites and when blocked imply that several trustworthy websites will be added to the blacklist, which causes a drastic reduction in the website's revenue and hinder its reputation (Aaron & Manning 2014).

After the creation of blacklists, many automated anti-phishing tools normally used by software companies such as McAfee, Google, Microsoft, were proposed. For instance, The Anti-Phishing Explorer 9, McAfee Site Advisor, and Google Safe Base are three common anti-phishing tools based on the blacklist approach. Moreover, companies such as VeriSign developed anti-phishing internet crawlers that gather massive numbers of websites to identify clones in order to assist in differentiating between legitimate and phishing websites.

There have been some attempts to look into creating whitelists, i.e. legitimate URL databases, in contrast to blacklists (Chen & Guo, 2006). Unfortunately, since the majority of newly created websites are initially identified as "suspicious," this creates a burden on the whitelist approach. To overcome this issue, the websites expected to be visited by the user should exist in the whitelist. This is sometimes problematic in practise because of the large number of possible websites that a user might browse. The whitelist approach is simply impractical since "knowing" in advance what users might be browsing for might be different to those actually visited during the browsing process. Human decision is a dynamic process and often users change their mind and start browsing new websites that they initially never intended to.

One of the early developed whitelist was proposed by (Chen & Guo, 2006), which was based on users' browsing trusted websites. The whitelist monitors the user's login attempts and if a repeated login was successfully executed this method prompts the user to insert that website into the whitelist. One clear limitation of Chen and Guo's method is that it assumes that users are dealing with trustful websites, which unfortunately is not always case.

Phishzoo is another whitelist technique developed by Afroz and Greenstadt (2011). This technique constructs a website profile using a fuzzy hashing approach in which the website is represented by several criteria that differentiate one website from another including images, HTML source code, URL, and SSL certificate. Phishzoo works as follows:

1. When the user browses a new website, PhishZoo makes a specific profile for that website.
2. The new website's profile is contrasted with existing profiles in the PhishZoo whitelist.
    - If a full match is found, the newly browsed website is marked trustful.
    - If partly matching, then the website will not be added since it is suspicious
    - If no match is found but the SSL certificate is matched, PhishZoo will instantly amend the existing profile in the whitelist.
    - If no match is found, then a new profile will be created for the website in the whitelist.

Recently, Lee et al. (2015) investigated the personal security images whitelist approach and its impact on internet banking users' security. The authors utilised 482 users to conduct a pilot study on a simulated bank website. The results revealed that over 70% of the users during the simulated experiments had given their login credentials despite their personal security image test not being performed. Results also revealed that novice users do not pay high levels of attention to the use of personal images in ebanking, which can be seen as a possible shortcoming for this anti-phishing approach.

.

4.2 Intelligent Anti-Phishing Techniques based on ML

Since phishing is a typical classification problem, ML and DM techniques seem appropriate for deriving knowledge from website features that can assist in minimising the problem. The key to success in developing automated anti-phishing classification systems is a website's feature. Since there are a tremendous number of features linked with a website, a necessary step to enhance the predictive system performance is to pre-process the set of features in order to pick up the "most" effective. Feature effectiveness can be measured using different

computational intelligence methods such as information gain, correlation analysis, and chi-square among others (Quinlan, 1979; Hall, 1999; Liu & Setiono, 1995).

Once an initial features set is chosen, the intelligent algorithm can be applied on the selected features to come up with the predictive system. There are many ML and DM algorithms for classification that have been developed by scholars in the last two to three decades as covered in Chapter 2. Most of these algorithms use one of the following major classification approaches in deriving their predictive systems:

1) Decision trees (ID3, C4.5 and successors)(Quinlan, 1993).
2) Probabilistic models (Naïve Bayes, Bayesian Network and successors)(Duda & Hart, 1973; Friedman et al., 1997).
3) Rule-based classification
    a. Associative classification (AC)
        i. Classification based Association (CBA and successors) (Liu et al., 1998).
        ii. Classification based on Multiple Association (CMAR and successors) (Li et al., 2001).
        iii. Multiclass Classification-based Association (MCAR and successors) (Thabtah et al., 2005).
    b. Rule induction such as FOIL, RIPPER and successors (Quinlan, 1979; Cohen, 1995).
    c. Covering or greedy, such as PRISM (Cendrowska, 1987) and eDRI (Thabtah, et al., 2016).
4) Neural Networks (NN) methods and their successors (Grossberg, 1988).
5) Support Vector Machine (SVM) (Joachims, 1999)
6) Fuzzy Logic (FL) (Zadeh, 1965)
7) Boosting and paging methods, and their successors (Freund & Schapire, 1997).
8) Search methods such as Genetic Algorithms (GA) (Goldberg, 1989)

Among the aforementioned classification approaches, rule-based classification systems are more suitable as an anti-phishing tool because: (1) their proven merits in predicting target values in many domains, including medical diagnoses, stock market analysis, email classification, text categorization, etc; and (2) the content of rule-based classification systems is simply human knowledge that novice users can easily understand and apply when necessary. Often the knowledge is formed as “If-Then” rules in which the antecedent of the rule (“If” part) consists of the conjunction of attribute values (Feature values) and the consequent of the rule (“Then” part) containing the target attribute value (Website type).

The rest of this section critically analyses intelligent anti-phishing attempts based on ML. We show how these approaches derive a classification anti-phishing system along with their benefits and weaknesses.

4.2.1 Decision Trees and Rule Induction

Fette et al. (2007) explored email phishing utilising the C4.5 decision tree classifier among other methods including Random Forest, SVM and Naïve Bayes. As a result, a new Random Forest method called "Phishing Identification by Learning on Features of Email Received" (PILFER) was developed. Experiments on a set of 860

phishy and 695 ham emails were conducted. Various features for distinguishing phishing emails identified included: IP URLs, time of space, HTML messages, number of connections inside the email, and JavaScript. The authors claim that PILFER can be improved towards grouping messages by joining all ten features discovered in the classifier apart from "Spam filter output".

Mohammad et al. (2014b) investigated a number of rule induction algorithms on the problem of website phishing classification. The authors compared RIPPER, C4.5 (Rules), CBA, and PRISM on a security dataset they collected containing 2500 instances and 16 features. A special hand crafted rule to collect the data was developed by the authors based on simple statistical analysis performed on the initial dataset's features. Experiments of the four rule-based classification methods showed that there are eight effective features that can be employed by the classification algorithm in combating phishing: SSL and HTTPS, Domain-age, Site-traffic, Long-URL, Request-URL Sub-domain, Multi—sub-domain, Suffix-prefix, and IP-address.

Khadi and Shinde (2013) studied the problem of email-based phishing and proposed a potential solution based on combining a RIPPER classifier with fuzzy logic. The role of fuzzy logic is to pick the main features of the email and rank them based on a probability score. Meanwhile, the role of RIPPER is to automatically use these features to classify the type of emails as ham or phishy. Two components of the email were utilized by Khadi and Shinde: the email message (spelling errors, embedded link) and URL (IP address, Length, Long URL, Suffix_Prefix, Crawler URL, Non matching URL). Moreover, very limited data consisting of just 100 instances from phishtank was in experiments involving the WEKA software tool. No comparison with other fuzzy logic or rule-based classifications was conducted by the authors. Results showed that there are twelve rules generated by RIPPER from the dataset with an 85.4% prediction rate.

Aburrous et al. (2010a) investigated rule induction methods to seek their applicability for categorising websites based on phishing features. Website features were initially manually classified into six criteria as described in an earlier report on phishing by Aburrous et al. (2008). Using WEKA, a number of experiments with four classification algorithms (RIPPER, PART, PRISM, C4.5) were conducted against 1006 instances downloaded from Phishtank. The focus of the experiments was the classification accuracy of the classifiers produced. Results revealed that rule induction is a promising approach because it was able to detect, on average, 83% of phishing websites. The authors suggested that results obtained could be further enhanced if a careful feature selection were employed.

4.2.2 Associative Classification (AC)

The two AC methods CBA and MCAR have been evaluated on a Phishtank dataset to seek their applicability in cracking phishing (Liu et al., 1998; Thabtah et al., 2005; Abourrous et al., 2010b). Abourrous et al. (2010b) used a dataset consisting of over 1000 instances with 27 different features and applied CBA, MCAR, and four other rule-based classifiers using the WEKA DM tool. The aim was to assist security managers within organisations by building an intelligent anti-phishing tool within browsers that can detect phishing as accurately as possible. Experimental results of the six ML algorithms revealed that AC methods generated more rules than the rest of the algorithms, yet had higher predictive classifiers. More specifically, the AC systems produced showed high correlations among features linked with three major criteria: URL, Domain Identity, and Encryption.

Nevertheless, the massive number of rules derived by MCAR and CBA may overwhelm end-users since they might not be able to control the anti-phishing system. Furthermore, the authors did not implement the AC rules within a browser to evaluate its real performance, which does not facilitate measuring the success or failure of their classification systems.

Recently, more domain specific AC anti-phishing systems have been created (Abdelhamid et al., 2014; Abdelhamid, 2015). These new models take into account not only two class values of the phishing problem (legitimate, phishy) but also considers a harder case to detect: the "suspicious" class label. Instances that cannot be fully classes as phishy nor as legitimate are very hard to detect by typical ML algorithms, thus increasing their false positive rates. Abdelhamid et al. (2014) and Abdelhamid (2015) have therefore enhanced current intelligent classification systems by including two distinct advantages: (1) extending the phishing problem to include suspicious cases, making it more realistic; and (2) proposing a new multi-label learning phase that can discover disjunctive in addition to conjunctive rules. These additional disjunctive rules are tossed out by existing AC methods. This new multi-label phase enhances predictive power and provides more useful knowledge to the end-user. The authors used a dataset that has 16 features and over 1500 instances, comparing the performance of their classifiers with other rule-based classifiers with respect to the knowledge derived and its accuracy. The authors employed the chi-square testing method to measure the features goodness and discriminate among features with respect to their impact on phishing. Processed data results showed high competitive performance of the new multi-label associative classifiers when compared with CBA, MCAR, rule induction, and decision trees.

4.2.3 Neural Network (NN)

One of the common ways to train a NN is trial and error (Mohammad et al., 2014a). However, this methodology has been criticised because of the time spent to tune the parameters and the requirement of an available domain expert. Thabtah et al. (2016b) proposed a NN anti-phishing model based on self-structuring the classification system rather than using trial and error. The algorithm proposed by the authors updated several parameters, like the learning rate, in a dynamic way before adding a new neuron to the hidden layer. The process of updating these NN features is performed during the building of the classification model and based on the network environment, behaviour of the desired error rate, and the computed error rate at that point. The dynamic NN model was applied to detect phishing on a large dataset from UCI containing over 11000 websites (Mohammad et al., 2015b). Experiments using different epoch sizes (100, 200, 500, 1000) have been conducted, and the results obtained exhibited better predictive systems when compared to Bayesian Network and Decision Trees.

The ANN Back Propagation algorithm (Rumelhart, et al., 1986) was investigated on a security dataset concerning website phishing by Mohammad et al. (2013). The authors collected a dataset with over 2000 instances from different legitimate and phishing sources. Processing the dataset, they tried to measure the correlation between the features and target attributes using basic univariate statistical analysis (frequency of features values and the target attribute values). Finally, they applied the Back Propagation ANN algorithm to derive anti-phishing models. The results of the study indicated that ANN is a promising approach for combatting phishing, particularly

since the results showed increased accuracy of the models generated from the Back Propagation algorithm when compared with decision trees and probabilistic.

Mohammad et al. (2014a) have developed an anti-phishing NN model that relies on constantly improving the learned predictive model based on previous training experiences. Since phishers continuously update their deception methods, new features become apparent while others become insignificant. In order to cope with these changes, the authors proposed a self-structuring NN classification algorithm that deals with the vitality of phishing features. The algorithm employs validation data to track the performance of the constructed network model and make the appropriate decision based on results obtained against the validation dataset. For instance, when the achieved error against the network is lower than the minimum achieved error, the algorithm saves the network's weights and continues the training process. However, when the achieved error is larger than the minimum achieved error so far, the algorithm continues the training process without saving the weights. Other important network parameters are also updated when necessary during the building of the classification model without waiting until the model has been entirely built. Results obtained against a phishing dataset of thirty features and over 10000 instances showed that the self-structuring NN model is able to generate anti-phishing models more accurately than traditional classification approaches such as C4.5 and probabilistic approaches.

Feed Forward NN (FFNN) was applied on an email phishing classification problem by Jameel and Georg (2013). Basic implementation of a multilayer FFNN based on Back Propagation was used to differentiate suspicious from legitimate emails. Eighteen binary features were extracted from the email (header and HTML body) and made available as the training dataset attributes. These features were given values based on human rules developed by security domain experts. To derive the NN models, 6000 emails were used. The results obtained showed that FFNN is able to categorise emails with high speed and with an error rate below 2%. However, the authors have not yet embedded their FFNN into browsers for live testing.

In 2007, an experimental study contrasting five ML algorithms on the problem of classifying emails as ham or suspicious was conducted by Abu-Nimeh et al. (2007). The authors chose Classification and Regression Trees (CART), NN, Random Forests (RF), Bayesian Additive Regression Trees (BART), and Logistic Regression (LR) to measure the most successful approaches in email phishing detection. A training dataset consisting of 2889 emails and 43 email's features was used. To produce the results, the testing method employed was ten-fold cross validation and the evaluation measures used were precision, recall, and harmonic mean. Results revealed that RF achieved a lower error rate while NN generated the highest error rate among the tested classifiers. Moreover, despite RF generating the highest predictive classifiers, it derived the least false positive rate among all contrasted algorithms. The authors suggested though that more carefully chosen features may improve the performance of the anti-phishing email tool.

4.2.4 Support Vector Machine (SVM)

Proposed by Pan and Ding (2006), the SVM classification method evaluates the discrepancy between a website's identity, its HTTP transactions, and structural features. The anti-phishing solution proposed contains two layers:

- Website Identity: The set of characters appearing inside the domain name.
- Structural Features Classifier: Features that are related to the website identity and HTTP transactions.

Once a new website identity and its structural features are captured (Abnormal URL, Abnormal anchors, Server Form Handler, Abnormal certificate in SSL, Abnormal DNS, Abnormal cookies), then a SVM algorithm is trained on a historical data set consisting of the same features in order to derive the new website type. Experimental results on six features using the proposed SVM indicated that the first helps toward increasing the detection rate since malicious websites are not correlated. Furthermore, the SVM model achieved just over 83% prediction rate, and therefore more investigation is needed into the feature selection phase by including other features that could improve the performance of the classifier.

4.2.5 Fuzzy Logic

Phishing in electronic banking (Ebanking) applications has been investigated by Aburrous et al. (2008) utilizing Fuzzy Logic. A simulated phishing email was sent by the authors with the help of the security manager at Jordan Ahli Bank to measure security indicators of phishing among a sample of 120 employees after obtaining the necessary authorization (www.ahlionline.com.jo). The email urged the chosen employees to reactivate their accounts by logging in because server maintenance conducted the previous two days required account reactivations. Shocking results were obtained: 37% of the targeted employees submitted their credentials without investigation, of which 7% were Information Technology employees. The authors' goal with the simulated email was to determine features that users may look for inside the email when they suspect phishing to be used within a FL system to help in differentiating types of email.

FL has been used as an anti-phishing model to help classify websites into legitimate or phishy in Aburrous et al. (2008). The authors claimed that FL could be effective in identifying phishing activities because it provides a simple way of dealing with intervals rather than specific numeric values. Their proposed FL classification model was built manually to categorise websites using the six criteria listed in Table 1. Each of those criteria contains a number of phishing indicators as described in the same table. Each feature in the dataset was assigned three possible values by the authors: Phishy, Genuine, and Doubtful. Limited results indicated that there are two effective indicators to distinguish phishiness in websites: Domain Identity and URL.

A fuzzy based ANN model was proposed in 2015 by Nguyen et al. (2015) to classify websites based on a smaller set of phishing features related to the website's URL (PrimaryDomain, SubDomain, PathDomain) and its rank (PageRank, AlexaRank, AlexaReputation). The proposed fuzzy ANN model does not use any rules set, rather it employs a computational function to split data instances (websites) into "genuine" and "non-genuine" categories. Their model was tested against 21600 websites from legitimate and phishing sources such as Phishtank and DMOZ. They also compared the generated results with that of Aburrous et al. (2010a) and Zhang and Yuan (2008). It was discovered that their fuzzy NN model was able to slightly enhance the phishing detection rate.

Table 1: Phishing Features per category (Aburrous, et al., 2008)

Criteria	N	Phishing Indicators
URL	1	IP address
	2	Abnormal request URL
	3	Abnormal URL of anchor
	4	Abnormal DNS record
	5	Abnormal URL
Encryption	1	Using SSL certificate (Padlock Icon)
	2	Certificate authority
	3	Abnormal cookie
	4	Distinguished names certificate
Source Code	1	Redirect pages
	2	Straddling attack
	3	Pharming attack
	4	OnMouseOver to hide the Link
	5	Server Form Handler (SFH)
Page Style & Contents	1	Spelling errors
	2	Copying website
	3	Using forms with *Submit* button
	4	Using pop-ups windows
	5	Disabling right-click
Web Address	1	Long URL address
	2	Replacing similar char for URL
	3	Adding a prefix or suffix
	4	Using the @ Symbol to confuse
	5	Using hexadecimal char codes
Human	1	Emphasis on security
	2	Public generic salutation
	3	Buying time to access accounts

4.2.6 CANTINA Term Frequency Inverse Document Frequency Approach

Carnegie Mellon Anti-phishing and Network Analysis Tool (CANTINA) is a content based anti-phishing method that determines suspicious websites using the statistical measure of Term Frequency Inverse Document Frequency

(TF-IDF). Term Frequency (TF) is a statistical formula that measures keyword significance in a document while Inverse Document Frequency (IDF) measures the importance of that keyword across a large collection of documents (Witten & Frank, 2005). CANTINA evaluates the website content (links, anchor tags, forms tags, images, text, etc) for TF-IDF to produce a lexical signature of the website. This signature (top ranked TF-IDF key words) will be passed into the search engine to seek their rank in domain names and decide the type of the website. The description of the CANTINA based classification process is as follows:

1. Parse the webpage.

2. Compute the TF-IDF for the common terms of the website.

3. Select the top five terms according to the computed scores of all TF-IDF terms.

4. Add the top five terms to the URL to locate the lexical signature.

5. Input the lexical signature into a search engine.

6. Check whether the domain name of the current website matches the domain names of the top N search results (often N=30).

7. Return "Legitimate" when there is a match or "Phishy" when there is no match.

When the search results in an empty set, the current website is classified as "phishy". To overcome the "no results" problem the authors merged TF-IDF with other content features such as "IP Address," "domain age." "suspicious Images," "suspicious Link," and "suspicious URL."

Sanglerdsinlapachai and Rungsawang (2010) have used CANTINA TF-IDF and added a few more features such as "Forms" and "Top pages' similarity linked with the domain," and removed features such as "domain age" and "known images." A dataset consisting of 200 websites was used in the experiments, and three DM methods were applied to the dataset. Results obtained, despite being limited, revealed that the reduced features set maintained a similar detection rate with that of the CANTINA features set. Moreover, adding the new features slightly enhanced the detection rate for most of the learning methods considered in the experiments. Table 2 shows a brief summary of the common anti-phishing approaches that are based on automated learning along with the name of the method, the learning approached used, the first author, and their reference details

Table 2 Common recent anti-phishing methods based on ML

Method name	ML technique	First Author	Reference
Dynamic rule induction	Rule induction learning	Qabajeh Issa	Qabajeh, et al., 2014
Enhanced Dynamic rule induction	Rule induction and covering approaches	Thabtah Fadi	Thabtah, et al., 2016
Classification based association	AC	Aburrous Maher	Aburrous, et al., 2010
Multi-label Classifier based Associative Classification	AC	Abdelhamid Neda	Abdelhamid, et al., 2014
Self-structuring neural network	NN	Mohammad Rami	Mohammad, et al., 2014
Neural Network trained with Back-Propagation	NN	Mohammad Rami	Mohammad, et al., 2013
Feed Forward Neural Network	NN	Jameel Noor Ghazi	Jameel & George, 2013
Fuzzy DM	Fuzzy logic	Aburrous Maher	Aburrous, et al., 2008
Fuzzy DM	Fuzzy logic	Khadi Anindita	Khadi & Shindi, 2014
PILFER	Decision tree	Fette Ian	Fette, et al., 2007
Page classifier	SVM	Pan Ying	Pan & Ding, 2006
CANTINA	Term frequency and inverse document frequency	Sanglerdsinlapachai Nuttapong	Sanglerdsinlapachai & Rungsawang, 2010
Biased SVM, LIBSVM, ANN, Self-Organizing Map	NN, SVM and other ML techniques	Basnet Ram	Basnet, et al., 2008

5. Conclusions

Website phishing classification is a fundamental problem due to the very large online transactions performed by businesses, individuals and governments. While many users are vulnerable to the phishing attacks, playing catch-up to the phishers' evolving strategies is not an option. There have been different approaches to combat phishing ranging from legal, educational, simulation, online community forums, black lists and machine learning among others. Unlike existing phishing reviews that were based around only intelligent techniques such as machine learning and data mining this paper focuses on raising awareness and educating users on phishing from training and legal prospective. This indeed will equip individuals with knowledge and skills that may prevent phishing on a wider context within the community. In this paper, we review conventional anti-phishing approaches such as law enforcement, user training, and education and then critically analyses their different methods. Then the attention is directed to review predictive ML method particularly rule-based methods, decision trees, associative classification, SVM, NN, and computational intelligence. We contrast the ways these methods detect phishing activities, their performance and their advantages and disadvantages.

While many countries such as the USA have taken a lead to criminalise phishing activities and put together more sever legislations, it is still hard to find attackers basically since phishing attacks have a short life span. Despite this limitation, it is still crucial that law enforcement agencies improve their information sharing work as well as jurisdiction. Moreover, educating novice users using visual cues can partly improve their abilities to detect phishing; however, many novice users still not paying high attention to visual cues when browsing the internet which make them vulnerable to phishing attacks. Users need to be exposed to repetitive training about phishing attacks since phishers continuously change the deception tactics.

Online phishing communities gather data that allow users to share information about phishing attacks such blacklisted URLs, which is useful information centre for users. However, this approach necessitates good awareness about web security indicators besides blacklisted URLs become outdated as updates are not performed in real-time.

Finally, anti-phishing methods based around ML especially AC and rule induction are suitable to combat phishing due to their high detection rate and more importantly the easy to understand outcomes they offer (If-Then rules). These rules empower novice users as well as security experts to understand and manage security indicators. However, adding a visualization layer into ML learning methods is advantageous to novice users as they may react quickly to visual cues.

In near future we are intend to design and implement a knowledge base using rule induction that can on real time warns online users of any possibility of phishing attacks

References

1. Aaron, G., Manning, R. (2014) APWG Phishing Reports. http://docs.apwg.org/reports/apwg_trends_report_q4_2014.pdf [Accessed March 20 2016].
2. Abdehamid N. (2015) Multi-label rules for phishing classification. Applied Computing and Informatics 11 (1), 29-46.
3. Abdelhamid N., Thabtah F., Ayesh A. (2014) Phishing detection based associative classification data mining. Expert systems with Applications Journal. 41 (2014) 5948–5959.
4. Abdelhamid N., Thabtah F., (2014) Associative Classification Approaches: Review and Comparison. Journal of Information and Knowledge Management (JIKM). Vol. 13, No. 3 (2014) 1450027.
5. Abu-Nimeh, S., Nappa, D., Wang, X. and Nair (2007) A Comparison of Machine Learning Techniques for Phishing Detection. In The 2nd annual Anti-Phishing Working Groupse Crime researchers, eCrime '07. New York, NY, USA, 2007. ACM.
6. Aburrous M.., Hossain M., Dahal K.P. and Thabtah F. (2010A) Experimental Case Studies for Investigating E-Banking Phishing Techniques and Attack Strategies. Journal of Cognitive Computation, Springer Verlag, 2 (3): 242-253.
7. Aburrous M., Hossain M., Dahal K.P. and Thabtah F. (2010B) Associative Classification techniques for predicting e-banking phishing websites. Proceedings of the 2010 International Conference on Information Technology, Las Vegas, Nevada, USA, 2010, pp. 176-181.
8. Aburrous M., Hossain A., Dahal K., Thabtah F. (2008) Intelligent Quality Performance Assessment for E-Banking Security using Fuzzy Logic. Proceedings of the 7th IEEE International Conference on Information Technology (ITNG 2008). Las Vegas, USA.
9. Afroz, & Greenstadt, R. (2011) PhishZoo: Detecting Phishing Websites by Looking at Them. In Fifth International Conference on Semantic Computing (September 18- September 21). Palo Alto, California USA,

2011. IEEE.

10. Arachchilage NAG, S Love S. (2013) A game design framework for avoiding phishing attacks. Computers in Human Behavior 29 (3), 706-714
11. Arachchilage NAG., Cole M. (2011) Design a mobile game for home computer users to prevent from "phishing attacks". Information Society (i-Society), 2011 International Conference on, 485-489.
12. Arachchilage NAG,, Rhee Y., Sheng S., Hasan SH., Acquisti A., Cranor L. F., Hong J. (2007) Getting users to pay attention to anti-phishing education: evaluation of retention and transfer. In eCrime '07 Proceedings of the anti-phishing working groups 2nd annual eCrime researchers summit. Pittsburgh, PA, USA, 2007. ACM.
13. BBC News, 2005.http://news.bbc.co.uk/2/hi/uk_news/england/lancashire/4396914.stm [Accessed 11 April 2016].
14. Basnet R., Mukkamala S., Sung AH (2008) Detection of phishing attacks: A machine learning approach (2008) Soft Computing Applications Industry, pp. 373-383.
15. Bright, M. (2011) MillerSmiles. [Online] Available at: http://www.millersmiles.co.uk/ [Accessed 09 January 2016].
16. Cendrowska, J. (1987) PRISM: An algorithm for inducing modular rules. International Journal of Man-Machine Studies, Vol.27, No.4, 349-370.
17. Chen, J. & Guo, C. (2006) Online Detection and Prevention of Phishing Attacks (Invited Paper). In First International Conference on Communications and Networking in China. ChinaCom '06. Beijing, 2006. IEEE.
18. ClickDimensions (2014) www.clickdimensions.com/sites/default/files/PDF/WhitePaper-CASL.pdf [Accessed 12 May 2016].
19. Cloudmark Org. (2002) Cloudmark. http://www.cloudmark.com/en/home [Accessed 10 February 2016].
20. Cohen, W.W., 1995. Fast Effective Rule Induction. In In Proceedings of the Twelfth International Conference on Machine Learning. Tahoe City, California, 1995. Morgan Kaufmann.
21. Executive-Order-13402, 2006. Executive Order 13402. http://www.gpo.gov/fdsys/pkg/FR-2006-05-15/pdf/06-4552.pdf [Accessed May 19, 2016].
22. General Assembly of Virginia, 2005. CHAPTER 827. http://leg1.state.va.us/cgi-bin/legp504.exe?051+ful+CHAP0827 [Accessed April 01 2016].
23. Google Safe-Browsing, 2010. Google Safe Browsing. http://code.google.com/p/google-safe-browsing/ [Accessed 10 April 2016].
24. Government of Australia (2011) Hackers, Fraudsters and Botnets: Tackling the Problem of Cyber Crime. Report on Inquiry into Cyber Crime, 2011.
25. Grossberg (1988) Nonlinear neural networks: Principles, mechanisms, and architectures. Neural Networks, 1(1), p.17–61.
26. Hall M., Frank E., Holmes G., Pfahringer B., Reutemann P., Witten I. (2009) The WEKA Data Mining

Software: An Update; SIGKDD Explorations, Volume 11, Issue 1.

27. Information Week (n.d.) http://www.informationweek.com/california-enacts-tough-anti-phishing-law-/d/d-id/1036636?. [Accessed March 17 2016].
28. Jameel N. Gh., George L. (2013) Detection of Phishing Emails using Feed Forward Neural Network. Journal of Computer Applications 77(7):10-15, September 2013.
29. Jevans D. (2003) Anti-Phishing Working Group (APWG): http://www.antiphishing.org/ [Accessed June 20th 2016].
30. Joachims H. Making (1999) Large-scale support vector machine learning practical, Advances in kernel methods: support vector learning, MIT Press, Cambridge, MA, 1999.
31. Julie S., D., Mandy, H. & Cranor, L.F., 2007. Behavioral Response to Phishing Risk. In The Anti-Phishing Working Groups, 2nd annual eCrime researchers summite, Crime '07. New York, NY, USA, 2007. ACM.
32. Khadi A., Shinde S. (2014) Detection of phishing websites using data mining techniques. International Journal of Engineering Research and Technology, Volume 2(12).
33. Lee J., Bauer L., Mazurek L. M. (2015) The Effectiveness of Security Images in Internet Banking. IEEE Internet Computing (Volume:19 , Issue: 1).Pp. 54 – 62.
34. Liu, H. and Setiono, R. (1995) Chi2: Feature Selection and Discretization of Numeric Attribute. Proceedings of the Seventh IEEE International Conference on Tools with Artificial Intelligence, November 5-8, 1995, pp. 388.
35. Liu, B., Hsu, W., and Ma, Y. (1998) Integrating classification and association rule mining. Proceedings of the Knowledge Discovery and Data Mining Conference- KDD, 80-86. New York.
36. McCall (2011) Gartner, Inc. http://www.gartner.com/newsroom/id/565125 [Accessed June 5th 2016].
37. McFredies, P (n.d.) Phishing. http://www.wordspy.com/words/phishing.asp [Accessed May 15th 2016].
38. Mohammad R., Thabtah F., McCluskey L., (2015A) Tutorial and critical analysis of phishing websites methods. Computer Science Review Journal. Volume 17, August 2015, Pages 1–24 Elsevier.
39. Mohammad R., Thabtah F., McCluskey L. (2015B) Phishing websites dataset. Available: https://archive.ics.uci.edu/ml/datasets/Phishing+Websites Accessed January 2016.
40. Mohammad R., Thabtah F., McCluskey L., (2014A) Predicting Phishing Websites based on Self-Structuring Neural Network. Journal of Neural Computing and Applications, 25 (2). pp. 443-458. ISSN 0941-0643. Springer.
41. Mohammad R., Thabtah F., McCluskey L., (2014B) Intelligent Rule based Phishing Websites Classification. Journal of Information Security (2), 1-17. ISSN 17518709. IET.
42. Mohammad, R. M., Thabtah, F. & McCluskey, L. (2013)Predicting Phishing Websites using Neural Network trained with Back-Propagation. Las Vigas, World Congress in Computer Science, Computer Engineering, and Applied Computing, pp. 682-686.

43. Nguyen L. A. T., To B. L., and Nguyen H. K. (2015) An Efficient Approach for Phishing Detection Using Neuro-Fuzzy Model. Journal of Automation and Control Engineering Vol. 3, No. 6, December 2015

44. Pan Y., and Ding X. (2006) Anomaly Based Web Phishing Page Detection. In The 22nd Annual Computer Security Applications Conference (ACSAC). Miami Beach, Florida, USA, 2006. IEEE.

45. PhishTank, 2011. PhishTank. http://www.phishtank.com/ [Accessed January 16 2016].

46. Pike, G. H. (2006). Lost data: The legal challenges. Information Today, 23(10), 1–3.

47. Platt J. (1998) Fast training of SVM using sequential optimization, (Advances in kernel methods – support vector learning, B. Scholkopf, C. Burges, A. Smola eds), MIT Press, Cambridge, 1998, pp. 185-208

48. Nahorney, B. (2015) The MessageLabs Intelligence Annual Security Report: 2009 Security Year in Review. http://www.symantec.com/content/en/us/enterprise/other_resources/intelligence-report-06-2015.en-us.pdf [Accessed June 09 2016].

49. Qabajeh I., Thabtah F., Chiclana F. (2015) Dynamic Classification Rules Data Mining Method. Journal of Management Analytics.Volume 2, Issue 3, pp. pages 233-253. Wiley.

50. Quinlan, J. (1993) C4.5: Programs for machine learning. San Mateo, CA: Morgan Kaufmann.

51. Rader M., Rahman S. (2015) Exploring historical and emerging phishing techniques and mitigating the associated security risks. International Journal of Network Security & Its Applications (IJNSA), Vol.5, No.4, July 2013 DOI : 10.5121/ijnsa.2013.5402 23.

52. Ronald, D.J.C., Curtis, C. & Aaron, F.J., (2007) Phishing for user security awareness. Computers & Security, 26(1), pp.73-80.

53. Retun Path (2016) https://blog.returnpath.com/blacklist-basics-the-top-email-blacklists-you-need-to-know-v2/ [Accessed 22 March 2016].

54. Rumelhart, David E.; Hinton, Geoffrey E.; Williams, Ronald J. (1986). Learning representations by back-propagating errors. Nature 323 (6088): 533–536.

55. Fette I., Sadeh N., Tomasic A. (2007) Learning to detect phishing emails. Proceedings of the 16th international conference on World Wide Web. 649-656.

56. Sahu KR., Dubey J. (2014). A Survey on phishign attacks. International Journal of Computer Applications (0975 – 8887).Volume 88 – No.10, pp. 42-45. February 2014.

57. Sanglerdsinlapachai N., and Rungsawang, A (2010) Using Domain Top-page Similarity Feature in Machine Learning-based Web. In Third International Conference on Knowledge Discovery and Data Mining., 2010. IEEE.

58. Sheng S., Holbrook M., Arachchilage NAG., Cranor L. Downs J. (2010) Who falls for phish?: a demographic analysis of phishing susceptibility and effectiveness of interventions. In CHI '10 Proceedings of the 28th international conference on Human factors in computing systems. New York, NY, USA, 2010. ACM.

59. Suganya, V. (2016). A Review on Phishing Attacks and Various Anti PhishingTechniques. International Journal of Computer Applications (0975 – 8887).Volume 139 – No.1, pp. 20-23. April 2016.

60. Thabtah F., Mohammad R., McCluskey L. (2016A) A Dynamic Self-Structuring Neural Network Model to Combat Phishing. In the Proceedings of the 2016 IEEE World Congress on Computational Intelligence. Vancover, Canada.

61. Thabtah F., Qabajeh I.., Chiclana F. (2016B) Constrained dynamic rule induction learning. Expert Systems with Applications 63, 74-85.

62. Thabtah, F., Cowling, P., and Peng, Y. (2005) MCAR: Multi-class classification based on association rule approach. Proceedings of the 3rd IEEE International Conference on Computer Systems and Applications , 1-7

63. Witten I. H. and Frank E. (2005). Data Mining: Practical Machine Learning Tools and Techniques.

64. WOT (2006) Web of Trust. http://www.mywot.com/ [Accessed 24 March 2016].