Review

Methods of Protection Against Phishing and Online Frauds

Nurzhibek Makushova *

Posted Date: 24 November 2025

doi: 10.20944/preprints202511.1732.v1

Keywords: phishing; online fraud; machine learning; DMARC; multi-factor authentication; spear phishing; vishing; cybersecurity

 doi:10.20944/preprints202511.1732.v1

Review

Methods of Protection Against Phishing and Online Frauds

Nurzhibek Makushova

Ala-Too International University, Kyrgyzstan; nurzhibek.makushova@alatoo.edu.kg

Abstract

Today, the digital world is changing incredibly fast — new technologies, apps, services, and AI tools appear every day. But as these technologies grow, so do cyber threats. Phishing, online fraud, and other attacks are becoming more common and more sophisticated. This work explores different ways to protect people and businesses from these threats and evaluates how effective these methods really are. In my research, I used a qualitative comparative method and examined four methods to protect against phishing and online fraud. It is important to regularly update and test security measures, as attackers constantly improve their schemes, create new ones, and adapting to existing protection tools.

Keywords: phishing; online fraud; machine learning; DMARC; multi-factor authentication; spear phishing; vishing; cybersecurity

1. Introduction

Today, information technology is developing at an incredible speed. Every day it evolves, improves, and is used in all areas of human life — both in everyday activities and in professional fields. Information technology has greatly simplified people's lives and increased the efficiency of many processes. However, in addition to the benefits of information technology, it is important not to forget that along with the advantages, there exist also a number of threats. Today, issues of information security are highly relevant as the number of various cyberattacks, including phishing and online fraud, continues to grow.

Phishing is an attack aimed at stealing your money or personal data by tricking you into providing confidential information — such as credit card numbers, banking details, or passwords — on websites that pretend to be legitimate [1]. Cybercriminals often impersonate trusted companies, friends, or acquaintances by sending fake messages that contain a link to a phishing website.

Online fraud is a type of crime that involves the use of digital technologies to commit financial or other fraudulent activities. For example, it may include the use of stolen credit cards, phishing emails, malicious software, and other harmful programs to gain access to personal information or financial accounts [2]. Online fraud can also involve identity theft, where criminals use stolen personal data to open new accounts or make purchases under another person's name.

The present study aims to review modern methods of protection against phishing and online fraud, study their main features, and discuss the most common prevention and detection approaches found in recent research. The purpose of the study is to provide a clear understanding of how different protection mechanisms — including technical tools, organizational measures, and user awareness — help reduce the risks and negative effects of these types of cybercrime.

2. Literature Review

2.1. Phishing and Online Fraud

Phishing has been a dangerous cybersecurity threat for decades. This malicious practice tricks people into handing over sensitive information, such as passwords, banking details, passport details, and so on, through emails, websites, or messages.

 doi:10.20944/preprints202511.1732.v1

Phishing is now widespread, but it wasn't always so. The first mention of phishing was on January 2, 1996, in a Usenet newsgroup called AOHell [3]. With the rapid growth of the internet in the late 1990s and early 2000s, phishing attacks also grew, exploiting the vulnerabilities of the new digital landscape. Cybercriminals began using emails and various fake websites, posing as trusted organizations, such as banks or government agencies, to trick users into revealing their personal information. Because of this, many large companies and individuals lost millions of dollars annually, not to mention the costs of restoration, protection, time, and resources. This significantly impacted the global economy.

As early as 300 BC, Hegestratos and Xenothemis devised a scheme [4] in which they would take out insurance on a boat and then sink it themselves, thus collecting all the insurance premiums. In 1919, Charles Ponzi created an investment scam in which he promised investors profits while extorting money from other participants for new contributions. This is how fraud evolved. Since phishing manifests in different forms depending on the attacker's goals, it is essential to examine the main types of phishing attacks in more detail.

2.2. Types of Phishing Attacks

2.2.1. Spear Phishing

Spear phishing is an attack on a specific employee of an organization aimed at stealing their authentication credentials. Before the attack, the attacker often collects information about the employee, including their name, position, and contact information.

Example: An attacker attempted to target an employee of NTL World, a Virgin Media company, using phishing. The attacker claimed the employee was required to sign a new employee handbook. The attacker's goal was to trick the employee into clicking a link and entering personal information [5]

2.2.2. Vishing

Vishing (short for voice phishing) is when an attacker impersonates another person, such as a friend or relative, to steal personal information.

Example: In August 2024, an attacker posing as an IT specialist called an employee of a hospitality company and obtained login credentials. This led to unauthorized access and network reconnaissance. The attack was subsequently detected and stopped by the Darktrace system [6].

2.2.3. Email Phishing

In an email phishing scam, the attacker sends an email that looks legitimate, designed to trick the recipient into entering information in reply or on a site that the hacker can use to steal or sell their data.

Example: In February 2024, European retailer Pepco Group was hit by an email phishing attack: hackers sent fake emails from suppliers with malicious attachments, resulting in the theft of credentials and financial losses of €15.5 million [7].

2.2.4. HTTPS Phishing

An HTTPS phishing attack is sending an email with a link to a fraudulent website to trick the victim into entering their login credentials.

Example: The Scarlet Widow group identifies corporate email addresses and sends nearly blank messages containing malicious HTTPS links. User interaction initiates the phishing sequence [5].

2.2.5. Watering Hole Phishing

In a watering hole phishing attack, a hacker figures out a site a group of users tends to visit. They then use it to infect the users' computers in an attempt to penetrate the network.

Example: From November 2023 to July 2024, the Russian group APT29 (Cozy Bear) compromised Mongolian government websites for a watering hole attack. Visitors with vulnerable iOS and Android devices were infected with spyware via exploits in Safari and Chrome, leading to the compromise of diplomats' and officials' data [8].

 doi:10.20944/preprints202511.1732.v1

2.2.6. Smishing

Smishing is phishing through text messages or SMS.

Example: Fraudsters posing as American Express sent urgent SMS alerts about account issues, directing users to a fake login page [5].

2.2.7. Website Spoofing

With website spoofing, a hacker creates a fake website that looks legitimate. When you use the site to log in to an account, your info is collected by the attacker.

Example: Hackers created a fake Amazon website, identical to the real Amazon.com but using a different URL. Everything else, down to the fonts, looked authentic. The attackers relied on users entering their username and password [5].

2.3. Online Fraud

Romance scams involve creating profiles on dating sites, pretending to be someone else, and using them to communicate, deceive, or extort money. **Government scams** involve deceptive emails that link to fake government websites, among other types of online scams.

Phishing scams

2.4. Machine Learning Approaches for Phishing Detection

As cyber threats become increasingly sophisticated, researchers have explored advanced technological methods to detect and mitigate phishing and fraud, including the use of machine learning algorithms.

Some works have been conducted to investigate how machine learning algorithms can detect phishing attacks. These studies tended to draw their databases from sources such as Alexa, Siri, and PhishTank. These datasets are fed into machine learning models, which can automatically identify useful features and patterns using labeled examples. The detection of phishing attacks can be carried out using machine learning techniques that can broadly be categorized as either **supervised** or **semi-supervised**. Supervised learning algorithms are based on labeled data to construct predictive models, whereas semi-supervised or unsupervised algorithms are able to infer patterns from unlabeled or partially labeled data.

Some of the most popular methods include **Neural Networks (NN)**, **Support Vector Machines (SVM)**, and **Random Forests (RF)**. Neural networks generally consist of an input layer, one or more hidden layers, and an output layer. It has been demonstrated that multi-layer perceptron (MLP) models trained with back-propagation can effectively detect sophisticated phishing patterns. SVMs are used in classification processes to locate the best hyperplanes to distinguish between phishing and non-phishing instances in multi-dimensional feature space. Random Forests, which aggregate the predictions of multiple decision trees, have also been shown to perform well in phishing detection because they are resistant to noisy and high-dimensional data.

The effectiveness of these methods is frequently evaluated based on performance metrics such as **sensitivity**, **specificity**, **accuracy**, and **F1-score**, which are determined using a confusion matrix. Sensitivity is the proportion of correctly identified phishing attacks, whereas specificity is the proportion of correctly identified non-phishing instances. Accuracy reflects the overall correctness of the model, and the F1-score is a balanced measure of precision and recall. These metrics allow researchers to compare the performance of different algorithms and optimize feature selection and model parameters to enhance phishing detection [9].

2.5. Technical Protection Methods for Email Security

DMARC(Domain-based Message Authentication, Reporting, and Conformance) is an authentication standard designed to prevent early sending address spoofing. It allows the domain owner to determine what to do with emails that did not pass SPF or DKIM verification, and to receive reports

 doi:10.20944/preprints202511.1732.v1

about emails and errors. Since DMARC works together with SPF and DKIM, it provides additional protection.

DMARC may contain personal data from senders and recipients. From the point of view of data protection [10], when using this standard, it is important to comply with GDPR — personal data should be anonymized where possible and used accordingly.

DMARC is an effective technical method of protection against phishing and online fraud; It increases domain security and reduces the risk of hacking.

2.6. Organizational Measures: User Awareness and Training

Another crucial method of protection against phishing attacks is user education and awareness training. Regardless of how advanced technical measures are, untrained or careless users can create vulnerabilities that bypass even the most sophisticated firewalls and security systems. Therefore, regular and effective training should be an integral part of an organization's cybersecurity strategy.

The main goal of user training is to help employees recognize different types of phishing and online fraud, understand the potential risks, and respond appropriately in suspicious situations. Training programs should combine both theoretical knowledge and practical exercises. For example, simulated phishing emails allow users to practice identifying malicious messages and reporting them correctly. Regular sessions and updates on emerging threats enhance the overall effectiveness of these programs.

It is important to note that the best defense strategy is multi-layered. Combining technical measures, organizational policies, and continuous user training ensures maximal protection against cyberattacks. By integrating user awareness into the security framework, organizations can safeguard both their systems and personnel, reducing the likelihood of successful phishing attacks.

2.7. Multi-Factor Authentication (MFA) Measure

Multi-Factor Authentication (MFA) is considered a security-enhancing strategy because it involves a person presenting two or more verification factors during the authentication process. These are normally knowledge factors (e.g., passwords or PINs), possession factors (e.g., security tokens or smartphones) and inherence factors (e.g., biometrics or facial recognition). The MFA may be applied by multiple methods, such as the issuance of a one-time password (OTP) using an SMS or any other biometric identifier such as fingerprints or facial features; the use of a hardware token that produces time-sensitive identifiers. MFA implementation has a tremendous positive impact on improving security through minimizing the risk of phishing and password theft and assists organizations to adhere to the regulatory aspects that require more robust authentication measures. It has been demonstrated that MFA can effectively reduce the success of phishing attacks by a significant margin (usually over 90 percent) because attackers are not able to bypass numerous authentication stages easily [11]. It has been applied in online banking systems and has been associated with a significant decrease in fraud. The use of biometric authentication, especially, offers a high protection because it is hard to duplicate or steal physical characteristics unlike the traditional password or token. Nonetheless, there are still the problems of user resistance, higher implementation costs and user education. Nevertheless, the MFA is generally accepted as the key component of a system of security measures designed to reduce fraudulent activity over the internet.

 doi:10.20944/preprints202511.1732.v1

Table 1. Comparative Analysis of Phishing Protection Methods

Method	Description and Principle of Operation	Advantages	Limitations / Disadvantages
Machine Learning Approaches for Phishing Detection	Uses algorithms like Neural Networks, SVM, and Random Forests to classify phishing and legitimate instances, automatically learning patterns from datasets (e.g., PhishTank, Alexa).	- High accuracy with quality and sufficient data - Can detect previously unseen attack patterns - Adaptive and self-improving models	- Requires labeled data and large datasets - Vulnerable to adversarial examples or evolving attacks - Computationally intensive
Technical Protection Methods (DMARC, SPF, DKIM)	Authentication protocols verify domain legitimacy and prevent sender address spoofing. DMARC integrates with SPF and DKIM to enforce domain-level security policies.	- Strong protection against spoofed or forged emails - Enhances domain reputation and trustworthiness - Provides automated reporting and filtering	- May include personal data (GDPR considerations) - Requires proper configuration and continuous maintenance
Organizational Measures (User Awareness and Training)	Regular training and awareness programs educate users to recognize phishing attempts and respond correctly. Simulated phishing tests improve user vigilance.	- Improves human factor resilience - Cost-effective and widely applicable - Promotes a culture of cybersecurity awareness	- Effectiveness depends on training quality and frequency - Human errors remain possible
Multi-Factor Authentication (MFA)	Uses two or more verification factors (knowledge, possession, or inherence) such as passwords, tokens, or biometrics. Strengthens authentication security and reduces phishing success rates.	- Significantly decreases successful phishing attempts (over 90%) - Increases overall system security - Supports compliance with security regulations	- Implementation and maintenance costs - User resistance and usability challenges - May require additional hardware or software infrastructure

3. Methodology

This study is a qualitative literature review with comparative analysis of existing protection methods against phishing and online fraud. No primary data collection or experiments were conducted. The analysis is based exclusively on peer-reviewed articles, industry reports, and publicly documented attack cases published between 2019 and 2025

3.1. Research Type

- **Type of study**: Qualitative research with elements of comparative analysis.
- **Justification**: This design allows the comparison of different protection methods: technical, organizational, without the need for empirical testing. It helps understand research questions related to effectiveness, costs, scalability, and risks.

The research process consists of the following steps:

1. Searching for and selecting relevant academic literature and industry reports.
2. Classifying protection methods into categories.
3. Conducting comparative analysis based on predefined evaluation criteria.

3.2. Objects of the Study

Since this is a secondary study, it does not involve human participants. The objects of analysis include:

- peer-reviewed scientific articles,
- reports from cybersecurity organizations,
- documented real-world phishing and fraud cases from public sources.

 doi:10.20944/preprints202511.1732.v1

3.3. Materials and Tools

- **Databases**: Google Scholar and other academic search engines for scientific publications; official websites of cybersecurity companies for industry reports.
- **Reports and documents**: Publications by Fortinet, GlobalSecurityMag, Microsoft, Hostragons Global Limited, the Certified Senders Alliance, and other organizations.

3.4. Data Collection

1. Generate search queries relevant to phishing, online fraud, and protection methods.
2. Search selected databases.
3. credibility and authority of the source (peer-reviewed journals; official reports from Microsoft, Fortinet, the Certified Senders Alliance, etc).
4. relevance (publication period 2019–2025);
5. Read, annotate, and code selected literature to extract key findings on protection methods, results, and application conditions.

3.5. Data Analysis Methods

- **Classification and coding**: All identified protection methods are grouped into predefined categories: machine learning approaches, technical protocols (DMARC, SPF, DKIM), organizational measures, and multi-factor authentication.
- **Comparative analysis**: Each category is assessed using three criteria: effectiveness, scalability, and limitations or risks.
- **Synthesis of results**: Findings are summarized, highlighting common patterns, advantages, and disadvantages. A comparative table is produced and supplemented with narrative interpretation.

3.6. Ethical Aspects

All data was taken exclusively from open sources—scientific articles, company reports, and cybersecurity news. No personal data was used in the study. Examples of real-life attacks are provided without identifying victims unless officially published. The study fully complies with ethical standards and GDPR requirements.

3.7. Limitations

Since I worked only with secondary data, my conclusions depend on the quality and completeness of the available sources. Phishing and security methods evolve rapidly—new technologies or vulnerabilities may emerge after this article is completed that are not considered here. Also, some company reports may be biased toward promoting their products.

4. Results

After analyzing scientific articles, cybersecurity reports, and real-world attack cases, I was able to identify clear patterns in how different protection methods against phishing and online fraud actually perform. It turned out that technical solutions, machine learning, employee training, and multi-factor authentication differ significantly in their effectiveness, reliability, and long-term practical impact. Each of these approaches works well in its own phase of an attack. Below, I present the key findings that emerged directly from this analysis.

4.1. Machine Learning Approaches for Phishing Detection

Analysis of the reviewed sources shows that Machine Learning (ML) based systems consistently demonstrate high accuracy in phishing detection, especially when processing large volumes of incoming mail. In many examined studies, models utilizing Neural Networks (NN), Support Vector Machines (SVM), Random Forest, and other algorithms successfully identified both typical and new, previously unknown phishing campaigns, including those where traditional signature and heuristic filters proved ineffective.

 doi:10.20944/preprints202511.1732.v1

However, a problem exists in that real-world effectiveness significantly drops due to issues with datasets, languages, and, crucially, targeted attacks using adversarial techniques and the rapid change in attacker tactics. As soon as the data distribution shifts (data drift), the model requires regular retraining and adjustment. Thus, Machine Learning remains one of the most promising areas, but only on the condition of continuous monitoring, training, and the availability of resources for model updates and the collection of current data.

4.2. Technical Protection Methods for Email Security

The analysis of SPF, DKIM, and DMARC protocols revealed that when properly configured, they reliably protect against sender domain spoofing. In all reviewed reports and case studies, domains with an enabled DMARC policy (especially p=reject or p=quarantine) were practically immune to successful spoofing—attackers simply could not send emails 'on behalf of' the organization.

However, there is a drawback: effectiveness is entirely dependent on the quality of implementation. Errors in DNS records are very common (incorrect selectors, superfluous or missing domains in SPF, or a non-existent or incorrect DMARC policy). Because of this, protection weakens, works only partially, or fails entirely. Furthermore, these protocols are powerless against emails sent from legitimate but already compromised accounts, or through third-party services that the organization has officially authorized in its SPF records.

Then, technical email authentication protocols are a powerful and relatively simple first barrier, but they achieve their maximum effectiveness only with correct configuration and periodic auditing.

4.3. Organizational Measures: User Awareness and Training

The reviewed articles all indicated improvement in the levels of measurable behavior of those who were exposed to structured training programs, phishing simulations, and ongoing awareness exercises. Those organizations that used periodic training based on scenarios found a progressive reduction in click-through in simulated phishing attacks. Nevertheless, the comparison also revealed discrepancies in the long-term maintenance of acquired skills.

One of the trends that kept surfacing was the loss of user vigilance a few months after training should no reinforcement mechanisms be used. Also security audit reports revealed that the employees were still vulnerable to high-quality social engineering attacks, which utilized a sense of urgency, signification of authority, or contextual personalization. The evidence presented in the reviewed article indicates that organizational controls can contribute to user resilience but are not able to completely change the human factor aspect of phishing risks.

4.4. Multi-Factor Authentication (MFA) Measure

The assessed sources repeatedly highlighted the high protection role of multi-factor authentication (MFA) in deterring an unauthorized access after the credential compromise. The hardware-based MFA systems and biometric-based MFA systems were determined as less prone to interception and replay. Even in instances where credentials were obtained using phishing, MFA blocked a sizable percentage of attempted access.

Nevertheless, other results also indicated case-specific weaknesses. Certain implementations of MFA were still vulnerable to SIM swap attacks, real-time proxy phishing, or user fatigue that resulted in clicking on a confirm of fraudulent login requests. Regardless of these shortcomings, MFA has shown high empirical results as an extra check measure in debased credentialing cases.

4.5. Cross-Method Effectiveness

A cross-method comparison revealed that different protection strategies excel at different stages of the phishing attack chain. ML systems and technical protocols demonstrated the strongest performance in early detection and filtering stages. Organizational measures primarily influenced user behavior at the interaction stage, while MFA mitigated risks in post-compromise authentication attempts.

 doi:10.20944/preprints202511.1732.v1

4.6. Comparative Evaluation Table

Table 2 summarizes the extracted results in three main protection categories using comparative criteria identified in the methodology: effectiveness, scalability, and limitations or risks. The table reflects aggregated patterns across all reviewed sources.

Table 2. Comparative evaluation.

Protection Method	Effectiveness	Scalability	Limitations / Risks
Machine Learning Approaches for Phishing Detection	High detection accuracy; adaptive to new patterns	High in automated environments; dependent on infrastructure	Susceptible to adversarial evasion; performance drops with data drift
Technical Protection (DMARC, SPF, DKIM)	Strong against spoofing and unauthorized domain use	Very high once deployed	Misconfiguration reduces effectiveness; ineffective against phishing from legitimate domains
Organizational Measures (Training, Awareness)	Moderate; improves recognition and reduces errors over time	Medium; effectiveness varies across individuals	Human error persists; performance declines without reinforcement; vulnerable to sophisticated social engineering
Multi-Factor Authentication (MFA)	Very high for preventing account compromise	High in most organizations	Susceptible to SIM swap and real-time phishing; user resistance to adoption

5. Discussion

The results of the study show that no single security method can completely protect against phishing and online fraud, but each approach has specific advantages in different cases.

Machine learning (ML) methods improve detection accuracy, especially when processing large volumes of electronic or network data, identifying both old and new attack patterns. However, their effectiveness depends on the majority of datasets and requires constant updating and training of models for new attacks.

DMARC, SPF, and DKIM are effective in preventing domain spoofing and unauthorized email delivery, but their effectiveness depends on configuration and ongoing maintenance, and they limit protection against attacks originating from pseudonymous accounts.

Organizational measures, including training and employee visibility, increase resilience to human error and reduce the risk of falling victim to phishing and online fraud. However, without additional and ongoing training, their effectiveness deteriorates over time.

Multi-factor authentication (MFA) is recognized as the most effective method for preventing account compromise, even in the event of stolen credentials, with hardware and biometric systems providing additional protection. However, certain vulnerabilities, such as SIM swapping, proxy phishing, and user resistance, can reduce its effectiveness under normal circumstances.

A comparative analysis of various methods showed that a combination of technological, organizational, and multi-factor authentication methods provides effective protection. Machine learning and technical protocols can drive early detection and filtering, user training influences behavior at the point of interaction, and MFA reduces risks after credential compromise. The results are consistent with previous research emphasizing the level of multi-layered collaboration in cybersecurity and underscore the need to adopt these measures for the highest level of protection.

While the study emphasizes secondary sources, the quality and comprehensiveness of which may be limited, it provides insight into the strengths, limitations, and practical implications of existing

 doi:10.20944/preprints202511.1732.v1

security methods. Future research should focus on empirically evaluating comprehensive defense strategies, exploring common AI-generated fishing strategies, developing interpretable machine learning models to improve user trustworthiness, and exploring behavioral interventions to maintain greater vigilance.

According to the study's findings, the most effective method of protecting against phishing and online fraud is a combination of two or more defense methods, such as user education and technical measures, multifactor authentication(MFA), and machine learning.

6. Conclusion

The analysis of contemporary literature and real-world cases confirms that none of the examined protection methods — machine learning models, email authentication protocols (SPF, DKIM, DMARC), user awareness training, or multi-factor authentication — can provide comprehensive protection against phishing and online fraud when implemented in isolation.

The highest level of security is achieved only through a defence-in-depth strategy that integrates all four categories:

- early detection and filtering (ML + DMARC/SPF/DKIM),
- reduction of human error (continuous training and simulations),
- robust authentication (phishing-resistant MFA),
- regular auditing and updating of all components.

As phishing techniques continue to evolve rapidly in 2025 and beyond, organizations must treat cybersecurity as a dynamic, multi-layered process rather than a one-time implementation.

References

1. Microsoft. Protect yourself from phishing, n.d.
2. Finscore. Types of online fraud, n.d.
3. Jason, J. History of phishing: A deep dive into its global impact, n.d.
4. APAC Insider. The evolution of online fraud and how to stay safe, 2024.
5. Fortinet. Types of phishing attacks, n.d.
6. Rushanth, R. From call to compromise: Darktrace's response to a vishing-induced network attack, 2024.
7. Slavin, B. A roundup of the top phishing attacks in 2024 so far, 2024.
8. Raza, M. What is a watering hole attack? Detection and prevention, 2025.
9. Gori, M.; Visumathi, J.; Mahdal, M.; Anand, J.; Elangovan, M. An effective and secure mechanism for phishing attacks using a machine learning approach. *Processes* **2022**, *10*, 1356. https://doi.org/10.3390/pr10071356.
10. Certified Senders Alliance. Protection contre le phishing: DMARC & RGPD sont-ils compatibles?, 2018.
11. Abill, R.; Adaan, A.; Billy, E. Investigating the effectiveness of multi-factor authentication against financial fraud, 2025.