Improving Phishing Detection with the Grey Wolf Optimizer

1st Aws Naser Jaber
Artificial Intelligence Lab
OsloMet – Oslo Metropolitan University
Oslo, Norway
awsalzar@oslomet.no

2nd Lothar Fritsch
Department of Computer Science
OsloMet – Oslo Metropolitan University
Oslo, Norway
lotharfr@oslomet.no

3rd Hårek Haugerud
Department of Computer Science
OsloMet – Oslo Metropolitan University
Oslo, Norway
haugerud@oslomet.no

***Abstract*—With the recent epidemic of COVID-19-themed scam and phishing, the efficient automated detection of such attacks is crucial. Although many anti-phishing solutions, such as lists and similarity and heuristic-based approaches detect attacks, methods still can be improved. Classification accuracy is highly dependent on the feature selection method used to select appropriate features for classification. In this article, a multi-objective grey wolf optimizer is used to select proper features for classifying phishing websites through a variational autoencoder. Our results indicate the superiority of the classification rate compared with related work: A classification rate of 97.49%, is obtained, thereby suggesting the feasibility of evaluating our work.**

***Index Terms*—cybersecurity; phishing website detection; anti-phishing; machine learning**

I. Introduction

Phishing exploits the most severe cybersecurity vulnerability and the weakest link in the security chain [1] - the human being. This was obvious in recent incidents where the attackers exploited human error and emotions (e.g., anxiety and stress) during the COVID-19 pandemic outbreak by targeting them using phishing emails, applications, and websites related to the coronavirus [2]. Attackers impersonated healthcare organizations to deliver fake COVID-19 related news and information, while other well-known healthcare specialists, using phishing attacks to get access to vaccine information. Moreover, they impersonated using fabricated video communications platforms, such as Zoom, Google Meet, and Skype, to mislead users to download malware, for example ransomware, or compromise data security [3]. More than 50% of their experts have been targeted by phishing attacks when they worked from home during the outbreak [4].

Phishing blacklists are a popular defense strategy aimed at protecting people from phishing attacks. These blacklists typically contain known phishing URLs, providing an access control list. Google Safe Browsing (GSB) and PhishTank are two of the most popular phishing blacklist providers [5].Phishing attacks against financial institutions were still the most common in the second quarter of 2021, according to APWG founding member OpSec Security, with 29.2% of all attacks, up from 22.5 percent in Q2020. Phishing attempts against cryptocurrency targets, such as cryptocurrency exchanges and wallet providers, increased from 2% to 7.5% in Q2. However, attacks on the most-targeted sectors increased significantly in 2021 as shown in Fig. 1.

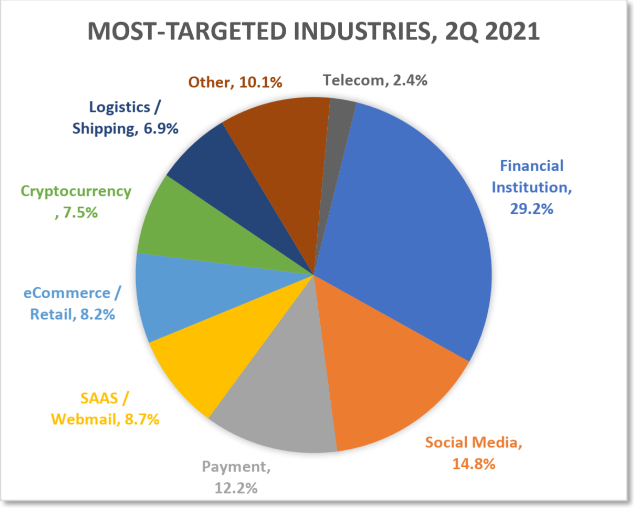

Fig. 1. Most-Targeted phishing sectors, Q1 2021

A. Motivation

Such incidents prove that the current security approach against phishing attacks is inefficient and fails to detect such attacks, in particular when encountering the new techniques and tactics used by attackers [6]. However, these current approaches highly depend on blacklist- and whitelist-based, tool bar, visual-similarity, and heuristic-based techniques, which fail to discover the new attack patterns [7]. For example, attackers can make a small change in the URL to bypass the list-based technique, or they can discover the features used in heuristic approaches to bypass the detection model [8]. To address these challenges, the usability of machine learning techniques has attracted attention and made security researchers design more intelligent anti-phishing approaches [9]. This is because they have high capabilities to deal with

The first author thanks the OsloMet Artificial Intelligence Lab (Norway) for full financial support. The work of the second and third authors was partially supported by the Research Council of Norway under Grant Agreement No. 303585 for the CyberHunt project and the OsloMet AI lab.

DOI: https://doi.org/10.1109/ICEIC54506.2022.9748592

hidden attack patterns in data. They also provide a high performance in phishing attack detection and fewer false alarms compared with other techniques [10].

The discriminative modeling is to learn a predictor given the observations, the goal of generative modeling is to tackle the more general problem of learning a joint distribution over all variables [3]. However, it may aid the development of useful world abstractions that can be applied to a range of prediction problems in the future. Unsupervised representation learning is the search for disentangled, semantically meaningful, statistically independent, and causal sources of variation in data, and the variational autoencoder (VAE) has been widely used for this purpose.

However, the anti-phishing-based machine learning approaches are still facing two key challenges: 1) selecting the most relevant features that can help in identifying such attacks efficiently as they have a strong influence on the decision-making process, and 2) selecting the best machine learning algorithm that power the decision engine. In recent years, deep learning (DL) has become increasingly significant in machine automation. Deep learning systems' success is heavily dependent on the availability of vast volumes of training data as well as the architecture used. The number of layers, the number of neurons in each layer, and the number of connections between layers are all factors to consider when choosing an architecture. Another drawback of deep learning systems is their reliance on gradient descent-based training methods, which can fail to find a global optimum solution in some cases.

Deep neural networks (DNN) provide a local optimum solution in the majority of cases due to poor parameter selection. There is no one-size-fits-all solution to these issues, and most architectures are chosen by trial and error. It is, however, impossible to find all potential parameter combinations via brute force searching. Parameters must be carefully chosen, since too few neurons can lead to under-fitting and too many neurons can lead to over-fitting. The use of evolutionary computation is one proposed answer to the challenge of architecture selection. Evolutionary computation technologies effectively solve optimization challenges by mimicking nature's optimized processes and behaviors. The ideal parameters for a DNN architecture may be found using evolutionary techniques, which can be phrased as an optimization problem. For the past few years, evolutionary algorithms have been utilized to solve this problem, although many academics are still unaware of the most up-to-date methods for DNN optimization.

B. Contribution

We propose a Multi-Objective Grey Wolf Optimizer (MOGWO) for VA evolution to improve the accuracy rate ”testing ” for detecting phishing websites. The following are the study's main contributions:

- For both optimized and generalized DNN architecture selection, a MOGWO-based algorithm is proposed. This algorithm is applicable to any phishing dataset and requires no domain knowledge. The architectures that were created represent CNNs which can be used for classification or generation tasks
- In MOGWO-VA, fitness is evaluated using both supervised and unsupervised learning objectives. If only a limited amount of labeled data is available, unsupervised data is used for training, and a limited amount of labeled data has been used in training to ensure meaningful

II. Background

A. Phishing attacks overview

Phishing is a combination of social engineering and technical exploitation designed to mislead the victims and convince them to provide personal and sensitive information such as user names, Identity (ID), password, and credit cards and with the objective of using it fraudulently [11]. Typically, phishing attacks consists of various methods as shown in Fig. 2, including:

- E-mail phishing: attackers send a fake email that seems to be a legitimate one to ask the users to change or update their information by following the provided link [12].
- Website-phishing: attackers create a fake website replicated of legitimate site and redirect the users to this malicious website using advertisements or using a link posted in social networks such as Facebook and Twitter [13].
- Key-loggers: This can be fake software such as fake mobile APP or flash player update and its target collect the user's keyboard information such as bank accounts [14].
- Content-injection phishing: attackers inject or replace part of the content of legitimate site using malicious code to mislead the users and obtain their credentials [15].
- Domain Name System (DNS) phishing, also known as “Pharming” where the attackers modify the company's host files or DNS to return a fake address and redirect the users to a malicious website [16].
- Session hijacking: this attack happens when the attackers monitor the end-user session in a legitimate website when they sign in to their account and do their transaction [17].
- Engine phishing: attackers create an attractive website with sounding offers and index it legitimately with search engines and thus the end-users find it easily in their normal search [18].
- Phone and voice phishing: attackers use fraudulent phone calls to convince people to give their personal information [19].

B. Artificial intelligence (AI) in Phishing attacks

Artificial intelligence (AI) continues to grow as a powerful tool for predicting and detecting phishing websites [20]. However, implementing AI in phishing attack detection is still suffering from too many false alarms as it is highly affected by noisy and irrelevant features and data. Several studies have been presented to improve the detection and prediction approaches' performance and to design more intelligent solutions to deal with the continuous evolution of phishing techniques

Fig. 2. Phishing attacks overview

[21]. For example, in [22] a Cyber Threat Intelligence (CTI) Model is proposed that defines the capabilities required for cyber security practitioners. The CTI capabilities in this model are described in terms of analytical component capability, contextual response capability, and experiential practice capability. This research is an excellent guide for cyber security practitioners and sets the theoretical basis for the CTI model. However, there should be a communication factor in the CTI capabilities model along with the other three terms, because for several threats, such as social engineering and phishing threats, communication is an important factor.

A feasible computational cost in Meta-heuristic algorithms need to find near-optimal solutions. The core algorithms for Meta-heuristic used agents for logical computation, for example op, Particle Swarm Optimizer (PSO), Genetic Algorithm (GA), Ant Colony Optimization Algorithm(ACO) and Grey Wolf Optimizer (GWO) particles. Due to the fact for their searching best solution, which led to figure -out the local or the global optimum solution. While, DNN may uses back-propagation (BP) or feed-forward neural network modify the weight but that mean it will not produce the best optimal solution. Hence, optimized weighted has improved in learning, testing stages of DNN. By intersect the weights for the Meta-heuristic and enhance the slope of the error rate as example.

III. Related Works

According to [23], similar character illusions can be harmful in terms of social engineering attacks and plagiarism detection evasion. In their research, they successfully analyzed the homoglyphs or similar characters using a machine learning approach with an average precision of 97%.

Their research is useful for defending against phishing and social engineering attacks. However, more improvements are required because composite homoglyphs and multiple unicode characters were not considered in their research.

A hybrid ensemble feature selection method for phishing attack detection was proposed by [24]. The cumulative Distribution Function gradient (CDF-g) is used to reduce the primary feature subsets and then pass them to ensemble data perturbation, where a single feature selection technique is applied to multiple data subsets. The output of the ensemble data perturbation was fed to an ensemble function perturbation where multiple feature selection techniques were applied to the same feature sunset. The main objective of using CDF-g is to define the cut-off of features ranking and address the challenges of overstating or understating features numbers. This proposed method was evaluated with several machine learning algorithms. Their results proved its superiority in terms of accuracy compared with other single filter-based selection methods. However, it showed less performance compared with the full feature sets in the case of random forest as classifier. In a related phishing detection work, a study proposed by [25] utilized two feature selection methods, Information Gain (IG) and Chi-square for choosing the most relevant features for phishing attacks. To identify the appropriate number of selected features, the authors proposed a threshold-based rule set. In this method, the two successive features for which they have at least 50% difference in values of IG and Chi-square was adopted as cut-off ranks identifier. Nevertheless, the proposed methods performance still needs to be improved.

A comprehensive study for evaluating and comparing various feature selection methods and phishing classifiers was presented by [26]. This includes filter measures such as IG and Relief-F, Correlation-Based Feature Selection (CFS), and the wrapper method. The experimental results showed that the wrapper evaluator and the best-first: forward searching method achieved the best performance in terms of accuracy, which proved their efficiency in selecting the best feature subset. In addition, the study also investigated the best classifier performance where the random forest achieved the best performance compared with other classical machine learning techniques. The authors concluded in the end that machine learning techniques can be extremely effective in designing anti-phishing approaches.

IV. The Proposed Framework

Based on studies on existing solutions, we see the need to design a new hybrid phishing website prediction framework. Our basic idea is that a MOGWO will improve the VA for feature selection. MOGWO is selected because it shows the best results in comparison with other metaheuristic algorithms. Furthermore, MOGWO is chosen as the base algorithm because it comprises a feasible mechanism that supports the weakness of VA by feeding their weights through MOGWO ternaries. The proposed framework is shown in Fig. 3 and will be explained in detail later in this section.

The MOGWO algorithm is selected as the candidate in this research instead of other classical search algorithms, such as Particle Swarm Optimization and genetic algorithm, because its advantages are embedded in the search mechanism. The key advantages are threefold, namely, the employment of multiple elite leaders, the adaptive transition from exploration to exploitation, and the stochastic nature in determining the trajectory to approach or diverge from the elicit signals. These advantageous characteristics endow MOGWO with enhanced exploration capability and search diversity while maintaining its efficient computational cost. Comparatively, other meta-heuristic algorithms are more likely to be trapped in local optima, owing to the dictation of the global best solution and the lack of diversification in its guiding signals over the entire iterative process. However, the hybrid method of MoGWO with VA begins by assigning weights in the initial layer for feature extraction. These values can be obtained through iterative learning. In this case, MOGWO-VA can obtain better performance than standard VA itself.

A. *MOGWO*

MOGWO is an enhanced GWO based on a multi-objective meta-heuristic [27]. To allow performing multi-objective optimization, two new components were developed and added to the GWO algorithm [28]. The first component was an archive for retrieving the best non-dominated obtained solution so far during optimization storing. The second component was a module that updated the position of omega wolves from the archive through a leader selection method that required MOGWO to select alpha, beta, and delta wolves. The unique feature of the archive was the high emphasis on the maintenance of the updating method. The MOGWO has improved the non-dominated solutions in the archive effectively. The proposed leader selection mechanism allowed the MOGWO algorithm to show superior coverage and convergence simultaneously. In the following section, mathematical models about the hunting techniques and social hierarchy of grey wolves are given to assist in MOGWO and optimization steps. The MOGWO design considerations to model the social hierarchy of grey wolves mathematically are given as follows:

- Alpha (α) is the fittest solution.
- Beta (β) is the second fittest solution.
- Delta (δ) is the third best solution.
- The rest of the pack is omega (ω) as seen in Fig. 4.
- Hunting (optimization) is guided by alphas, betas, and deltas.
- Omegas follow the three other wolves in the hunting.

Modelling the "encircling prey" step of hunting and the mathematical foundation of the algorithms is explained in detail in the original work [27].

Grey wolves can recognize their prey's location and encircle them. The alpha usually guides this hunt along with beta and delta occasionally but an abstract search space, the optimum (prey) cannot be located easily. For practical mathematical simulations, alpha, beta, and delta have been employed to know better about the prey's location. Thus, these three are the best solutions, and the omegas are obliged to update their positions accordingly. Summing up, the search for prey begins with the creation of the random population of the grey wolves in the MOGWO algorithm. In most iterations, the leaders of the pack estimate where a prey might be located.

B. *Hybird MOGWO-VAE for feature selection*

The variational autoencoder takes the input and returns parameters for a probability density. For example, Gaussian gives the mean and co-variance matrix. It can sample from this distribution to obtain random values of the lower-dimensional representation of z. Implemented via a neural network, each input x provides a vector mean and diagonal covariance matrix that determines the Gaussian density. The parameters for the NN must be learned to set up a loss function. In the meantime, the decoder takes the latent variable z and returns the parameters for a distribution. This finding gives the mean and variance for each phishing attack feature in the output. Reconstruction is produced by sampling and implemented via a neural network, and the NN parameters are learned.

$$L(\theta, \emptyset, x) = E_{q\emptyset(z|x)}[log_{p\theta}(x,z) - log_{q\emptyset}(z|x)] \tag{1}$$

$$= E_{q\emptyset(z|x)}[log_{p\theta}(X|Z) + log_{p\theta}(Z) - log_{q\emptyset}(Z|X)] \tag{2}$$

$$= -D_{KL}[q\emptyset(Z|X)||p\theta(Z)] + E_{q\theta(z|x)}[log_{p\theta}(X|Z)] \tag{3}$$

Algorithm 1 The procedure of MOGWO-AV

```
Require: Initialize the MOGWO population and set-up each
    CNN layer;
Ensure: Calculation process: weights (w), biases (b) loss
    function (fx)
  Solution vector (x): w and b on the last layer;
  while Termination criterion is not satisfied do
    For total numbers of x' do;
    x'=x+△x, $ℒ x'
    if ℒ x' ≥ l (x') then
      $x \Longleftarrow x'$
    else
      Indicate transition probability $p(x) \Longleftarrow x'$ ;
      All new layers update through x
    end if
  end while=0
```

V. Experimental setup

Initially, the experimental environment is presented in this section. Next, the experimental design and dataset description are provided. Finally, the performance evaluation is described and explored. The CPU of the computer used in this experiment is an Intel(R) Core (TM) Intel Core i7 @ 2.80GHz, the amount of RAM is 32 GB, and the operating system is 64-bit Windows 10. The simulation software is MATLAB 2021b.

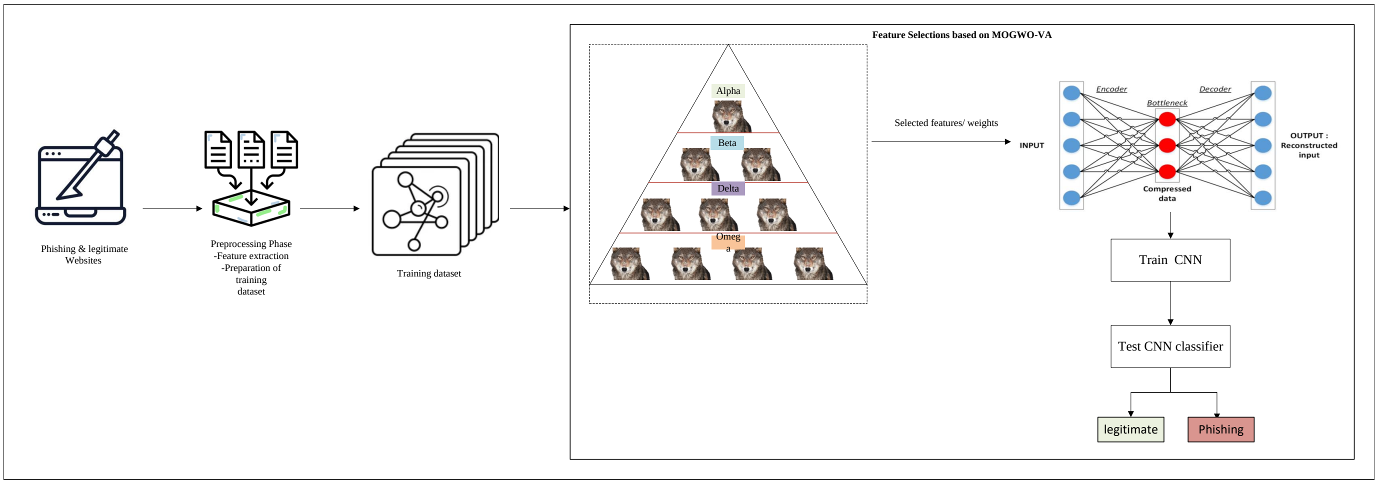
Phishing & legitimate Websites
Preprocessing Phase -Feature extraction -Preparation of training dataset
Training dataset
Feature Selections based on MOGWO-VA
Alpha
Beta
Delta
Omeg a
Selected features/ weights
INPUT
Encoder
Bottleneck
Decoder
OUTPUT : Reconstructed input
Compressed data
Train CNN
Test CNN classifier
legitimate
Phishing

Fig. 3. The proposed framework

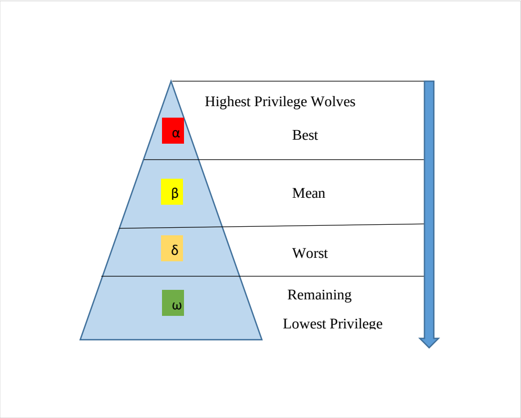

Fig. 4. Different grades of grey wolves

A. *Preprocessing and Feature Extraction*

These data consist of a collection of legitimate, as well as phishing website instances. Each website is represented by the set of features that denote whether the website is legitimate or not. The dataset serves as input for the machine learning process. In total, it features 111 attributes with 88647 instance, excluding the target phishing attribute, which denotes whether the particular instance is legitimate (value 0) or phishing (value 1). We prepared two variations of the dataset, one where the total number of instances is 58,645 and the balance between the target classes in more or less balanced with 30,647 instances labeled as phishing websites and 27,998 instances labeled as legitimate. The attribute evaluator used CfsSubsetEval, which evaluates the worth of a subset of attributes by considering the individual predictive ability of each feature along with the degree of redundancy between them. The search method used multi-objective evolutionary search, which explores the attribute space equivalent to evolutionary non-dominated radial slot-based algorithm or the NSGA-II multi-objective evolutionary algorithm. Fig. 5 shows how the URL is separated into four parts.

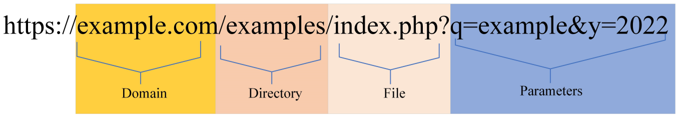

Fig. 5. Separation of the URL [13]

VI. Comparative Analysis

The dataset used to train the classifiers is based on PhishTank, which have 112 features. It consists of 100,000 URLs that were collected in October 2021 and were applied in the training phase with a 80/20 split between training and testing data, respectively. However, the feature selections have selected 17 feature numbers from 10,15,20,21,39,41,44,46,49,59,64,67,87,95,100,101, and 111 . After the feature selection and extraction, we went to the CNN for training and testing to determine whether the results are legitimate or phishing. Given the feasibility of feature selection, our classification rate shows better result than other related works, as shown in Table I.

TABLE I
MOGWO-VA comparison with ML, DL and Metaheuristic

Algorithm/Method	Type	Dataset	Accuracy
MOGWO-VA	Hybrid	PhishTank	97.49%
Adam optimizer and DNN [29]	Deep learning	UCI	90.38
RNN and CNN [30]	Deep learning	phishing and normal URLs website	95.79%
CNN [31]	Deep learning	Phishing website datasets.	95.02%
Auto encoder + NIOSELM [32]	Hybrid	Websites (PhishTank, Alexa, DMOZ)	94.60%
Grey wolf optimizer + SVM [33]	Hybrid	Websites (PhishTank, Yahoo)	90.38%
Genetic algorithm (GA) + DNN [34]	Deep learning	UCI	89.50%
Convolutional auto encoder + DNN [35]	Deep learning	Websites (PhishTank)	89.00%

VII. Further Discussions and Conclusions

Many aspects of a website, including the URL, page, content features, domain features, source code, and so on, are used to detect it. Thus, determining which features can be employed to train a model to improve detection accuracy is challenging. The prediction results may not be reliable if only a single feature is used for detection. The use of a website's numerous features provides more information about the site, which can aid in detection. The paper presents a hybrid machine learning approach for detection of websites phishing attacks based on Multi-Objective Grey Wolf Optimizer. In this paper, work in progress has shown a feasible result in terms of accuracy. While, the limitations of the study need to perform more ML metrics with a large phihsing data set. In the future, we will plan to work more in real time to make the MOGAO-VA the rule for intrustion detection systems.

References

[1] Valentin Mullet, Patrick Sondi, and Eric Ramat. A Review of Cybersecurity Guidelines for Manufacturing Factories in Industry 4.0, 2021.

[2] Bernardi Pranggono and Abdullahi Arabo. COVID -19 pandemic cybersecurity issues . *Internet Technology Letters*, 4(2), 2021.

[3] Ugur Ozker and Ozgur Koray Sahingoz. Content Based Phishing Detection with Machine Learning. In *2020 International Conference on Electrical Engineering, ICEE 2020*, 2020.

[4] Harjinder Singh Lallie, Lynsay A. Shepherd, Jason R.C. Nurse, Arnau Erola, Gregory Epiphaniou, Carsten Maple, and Xavier Bellekens. Cyber security in the age of COVID-19: A timeline and analysis of cyber-crime and cyber-attacks during the pandemic. *Computers and Security*, 105, 2021.

[5] Simon Bell and Peter Komisarczuk. An Analysis of Phishing Blacklists: Google Safe Browsing, OpenPhish, and PhishTank. In *ACM International Conference Proceeding Series*, 2020.

[6] Yuosuf Al-Hamar, Hoshang Kolivand, Mostafa Tajdini, Tanzila Saba, and Varatharajan Ramachandran. Enterprise Credential Spear-phishing attack detection. *Computers and Electrical Engineering*, 94, 2021.

[7] Muna Al-Hawawreh, Nour Moustafa, Sahil Garg, and M. Shamim Hossain. Deep Learning-enabled Threat Intelligence Scheme in the Internet of Things Networks. *IEEE Transactions on Network Science and Engineering*, 2020.

[8] Routhu Srinivasa Rao and Syed Taqi Ali. PhishShield: A Desktop Application to Detect Phishing Webpages through Heuristic Approach. In *Procedia Computer Science*, volume 54, 2015.

[9] S. Priya and S. Selvakumar. Detection of phishing attacks using probabilistic neural network with a novel training algorithm for reduced Gaussian kernels and optimal smoothing parameter adaptation for mobile web services. *International Journal of Ad Hoc and Ubiquitous Computing*, 36(2), 2021.

[10] A. Aldo Tenis and R. Santhosh. Challenges and Security Issues of Online Social Networks (OSN). In *Lecture Notes on Data Engineering and Communications Technologies*, volume 68. 2022.

[11] Rana Alabdan. Phishing attacks survey: Types, vectors, and technical approaches, 2020.

[12] Priyanka Verma, Anjali Goyal, and Yogita Gigras. Email phishing: text classification using natural language processing. *Computer Science and Information Technologies*, 1(1), 2020.

[13] Grega Vrbančič, Iztok Fister, and Vili Podgorelec. Datasets for phishing websites detection. *Data in Brief*, 33, 2020.

[14] Thorsten Balke, Alejandra Vovides, Christian Schwarz, Gail L. Chmura, Cai Ladd, and Mohammad Basyuni. Monitoring tidal hydrology in coastal wetlands with the ”mini Buoy”: Applications for mangrove restoration. *Hydrology and Earth System Sciences*, 25(3), 2021.

[15] V. Srijyothi Mopidevi, K. V.D. Kiran, and Dinesh Hirawat. A tool for analyzing & mitigating application vulnerabilities in any web application. *Journal of Advanced Research in Dynamical and Control Systems*, 12(2), 2020.

[16] Kang Li, Xiangzhan Yu, and Jiujin Wang. A Review: How to Detect Malicious Domains. In *Communications in Computer and Information Science*, volume 1424, 2021.

[17] Anuj Kumar Baitha and Prof. Smitha Vinod. Session Hijacking and Prevention Technique. *International Journal of Engineering & Technology*, 7(2.6), 2018.

[18] Sourena Maroofi, Maciej Korczyński, and Andrzej Duda. Are You Human?: Resilience of Phishing Detection to Evasion Techniques Based on Human Verification. In *Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC*, 2020.

[19] JaeSeung Song and Andreas Kunz. Towards Standardized Prevention of Unsolicited Communications and Phishing Attacks. *Journal of ICT Standardization*, 2021.

[20] Yazan Ahmad Alsariera, Victor Elijah Adeyemo, Abdullateef Oluwagbemiga Balogun, and Ammar Kareem Alazzawi. AI Meta-Learners and Extra-Trees Algorithm for the Detection of Phishing Websites. *IEEE Access*, 8, 2020.

[21] Abdelhakim Hannousse and Salima Yahiouche. Towards benchmark datasets for machine learning based website phishing detection: An experimental study. *Engineering Applications of Artificial Intelligence*, 104, 2021.

[22] Bongsik Shin and Paul Benjamin Lowry. A review and theoretical explanation of the 'Cyberthreat-Intelligence (CTI) capability' that needs to be fostered in information security practitioners and how this can be accomplished, 2020.

[23] Perry Deng, Cooper Linsky, and Matthew Wright. Weaponizing Unicodes with Deep Learning -Identifying Homoglyphs with Weakly Labeled Data. In *Proceedings - 2020 IEEE International Conference on Intelligence and Security Informatics, ISI 2020*, 2020.

[24] Kang Leng Chiew, Choon Lin Tan, Kok Sheik Wong, Kelvin S.C. Yong, and Wei King Tiong. A new hybrid ensemble feature selection framework for machine learning-based phishing detection system. *Information Sciences*, 484, 2019.

[25] Fadi Thabtah and Neda Abdelhamid. Deriving correlated sets of website features for phishing detection: A computational intelligence approach. *Journal of Information and Knowledge Management*, 15(4), 2016.

[26] Mahmoud Khonji, Andrew Jones, and Youssef Iraqi. An empirical evaluation for feature selection methods in phishing email classification. *Computer Systems Science and Engineering*, 28(1), 2013.

[27] Seyedali Mirjalili, Shahrzad Saremi, Seyed Mohammad Mirjalili, and Leandro Dos S. Coelho. Multi-objective grey wolf optimizer: A novel algorithm for multi-criterion optimization. *Expert Systems with Applications*, 47:106–119, 4 2016.

[28] Seyedali Mirjalili, Shahrzad Saremi, Seyed Mohammad Mirjalili, and Leandro Dos S. Coelho. Multi-objective grey wolf optimizer: A novel algorithm for multi-criterion optimization. *Expert Systems with Applications*, 47, 2016.

[29] L. Lakshmi, M. Purushotham Reddy, Chukka Santhaiah, and U. Janardhan Reddy. Smart Phishing Detection in Web Pages using Supervised Deep Learning Classification and Optimization Technique ADAM. *Wireless Personal Communications*, 118(4), 2021.

[30] Yongjie Huang, Qiping Yang, Jinghui Qin, and Wushao Wen. Phishing URL detection via CNN and attention-based hierarchical RNN. In *Proceedings - 2019 18th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/13th IEEE International Conference on Big Data Science and Engineering, TrustCom/BigDataSE 2019*, 2019.

[31] Brij B. Gupta, Krishna Yadav, Imran Razzak, Konstantinos Psannis, Arcangelo Castiglione, and Xiaojun Chang. A novel approach for phishing URLs detection using lexical based machine learning in a real-time environment. *Computer Communications*, 175, 2021.

[32] Liqun Yang, Jiawei Zhang, Xiaozhe Wang, Zhi Li, Zhoujun Li, and Yueying He. An improved ELM-based and data preprocessing integrated approach for phishing detection considering comprehensive features. *Expert Systems with Applications*, 165, 2021.

[33] Sagnik Anupam and Arpan Kumar Kar. Phishing website detection using support vector machines and nature-inspired optimization algorithms. *Telecommunication Systems*, 76(1):17–32, 2021.

[34] Waleed Ali and Adel A. Ahmed. Hybrid intelligent phishing website prediction using deep neural networks with genetic algorithm-based feature selection and weighting. *IET Information Security*, 13(6), 2019.

[35] S T Deepa. Phishing Website Detection Using Novel Features And Machine Learning Approach. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(7):2648–2653, 2021.