TÜRKİYE CUMHURİYETİ
ADANA ALPARSLAN TÜRKEŞ SCIENCE AND TECHNOLOGY UNIVERSITY

GRADUATE SCHOOL
INFORMATION TECHNOLOGIES DEPARTMENT

PHISHING WEBSITE DETECTION BASED ON A NOVEL ARTIFICIAL INTELLIGENCE TECHNIQUE

KEMAL ERDİ

M.Sc.

ADANA 2023

TÜRKİYE CUMHURİYETİ
ADANA ALPARSLAN TÜRKEŞ SCIENCE AND TECHNOLOGY UNIVERSITY

GRADUATE SCHOOL
INFORMATION TECHNOLOGIES DEPARTMENT

PHISHING WEBSITE DETECTION BASED ON A NOVEL ARTIFICIAL INTELLIGENCE TECHNIQUE

KEMAL ERDİ

M.Sc.

THESIS ADVISOR
ASSOC. PROF. DR. FATİH KILIÇ

ADANA, 2023

CERTIFICATION OF APPROVAL

PHISHING WEBSITE DETECTION BASED ON A NOVEL ARTIFICIAL INTELLIGENCE TECHNIQUE

This **M.Sc.** thesis, **completed under the conditions determined by the relevant regulations**, by **Kemal ERDİ with student number** 19700201007 in **Adana Alparslan Türkeş Science and Technology University Institute of Graduate School Information Technologies Department**, has been evaluated by the following academic committee, and approved in accordance with the relevant articles of the "Adana Alparslan Türkeş Science and Technology University Graduate Education Regulations".

Prof. Dr. Ahmet BEYCİOĞLU

Institute of Graduate School Manager

Assoc. Prof. Dr. Fatih KILIÇ

Head of the Department

Assoc. Prof. Dr. Fatih KILIÇ

Academic Advisor

Thesis Advisory Committee Members

Assoc. Prof. Dr. Fatih KILIÇ

Adana Alparslan Türkeş Science and Technology University

Assoc. Prof. Dr. Yasin KAYA

Adana Alparslan Türkeş Science and Technology University

Assoc. Prof. Dr. Fatih ABUT

Çukurova University

Thesis Defense Date

14/11/2023

DECLARATION OF CONFORMITY

In this thesis study, which was prepared following the thesis writing rules of Adana Alparslan Türkeş Science and Technology University Institute of Graduate School, I declare that I provide all the information, documents, evaluations and results in accordance with scientific ethics and moral codes without resorting to any means or assistance that would be contrary to scientific ethics and traditions. I also declare that I refer to all of the articles I used in this study with appropriate references and accept all moral and legal consequences if a situation is found contrary to my statement regarding my work.

14/11/2023

Kemal ERDİ

ÖZET

YENİ BİR YAPAY ZEKA TEKNİĞİNE DAYALI OLTALAMA WEB SİTESİ TESPİTİ

Kemal ERDİ
Yüksek Lisans, Bilişim Teknolojileri Anabilim Dalı
Danışman: Doç. Dr. Fatih KILIÇ
Kasım 2023, 32 sayfa

İnternet kullanımının artması nedeniyle araştırmacılar siber suçların önlenmesine yönelik geniş çapta çalışmalar yapmaktadır. Siber suçlardan biri de oltalama web siteleridir. Siber suçlular giderek daha karmaşık taktikler uyguladıkça, bu tehditleri tespit etmek ve önlemek için yenilikçi yaklaşımlara olan ihtiyaç da artıyor. Bu tezde BGWO, BPSO ve BHHO için zamanla değişen aynalı S-şekilli transfer fonksiyonu uygulanmış ve önerilen modeller oltalama web sitesi veri seti üzerinde test edilmiştir. Önerilen modeller, oltalama web sitesi veri kümesinde umut verici sonuçlara sahiptir.

Anahtar Kelimeler: Oltalama saldırısı, Sürü Zekası, S-Şekilli Transfer Fonksiyonu, Öznitelik Seçimi, Aynalama

ABSTRACT

PHISHING WEBSITE DETECTION BASED ON A NOVEL ARTIFICIAL INTELLIGENCE TECHNIQUE

Kemal ERDİ

M.Sc., Department of Information Technologies

Supervisor: Assoc. Prof. Dr. Fatih KILIÇ

November 2023, 32 pages

Due to the increasing use of Internet, researchers widely study to prevent cybercrimes. One of the cybercrimes is phishing websites. As cybercriminals employ increasingly sophisticated tactics, there is a growing need for innovative approaches to detect and prevent these threats. In this thesis, the time-varying mirrored S-shaped transfer function is applied for BGWO, BPSO, and BHHO, and the proposed models are performed on the phishing websites dataset. The proposed models have promising results on the phishing websites dataset.

Keywords: Phishing Attack, Swarm Intelligence, S-Shaped Transfer Function, Feature Selection, Mirrored

Dedication

To my beloved family…

TABLE OF CONTENTS

CURRICULUM VITAE

LIST OF FIGURES

LIST OF TABLES

LIST OF ALGORITHMS

LIST OF ABBREVIATIONS

ABC	: Artificial Bee Colony
ACO	: Ant Colony Optimization
AI	: Artificial intelligence
APWG	: Anti-Phishing Working Group
BCSA	: Binary Crow Search Algorithm
BGWO	: Binary Grey Wolf Optimization
BHHO	: Binary Harris Hawks Optimization
BPSO	: Binary Particle Swarm Optimization
CNNs	: Convolutional neural networks
FS	: Feature Selection
GA	: Genetic Algorithm
GMGWO	: Generalized Mean Gray Wolf Algorithm
GWO	: Grey Wolf Optimization
HHO	: Harris Hawk Algorithm
IDS	: Intrusion Detection System
KNN	: K-nearest Neighbors
PSO	: Particle Swarm Optimization
QBGWO	: Quadratic Binary Grey Wolf Optimization
RF	: Random Forest
RNNs	: Recurrent neural networks
SI	: Swarm intelligence
SIO	: Swarm Intelligence Optimization
SVM	: Support Vector Machine
TV	: Time-Varying
TVMS	: Time-Varying Mirrored S-Shaped
URL	: Uniform Resource Loader

1. INTRODUCTION

1.1. Overview

In recent years, with the rapid spread of the internet and increasing digitalization, cyber security threats for users and organizations are increasing. The general term for security threats is cyber-attack, which can be defined as a deliberate attempt to damage computers, steal data, or enable a compromised computer system to launch new attacks on other computers. Cyber-attacks can occur from different sources. Attackers, sometimes also named as cyber criminals, can be grouped as hackers, non-malicious attackers, and hacktivists. The dangerous ones among these are usually the actors who aim to cause direct damage by accessing the other computers or the information systems in their attacks.

Many fake sites have been created on the internet that imitate real websites in order to obtain sensitive and valuable information from individuals and companies. These types of online attacks, known as phishing attacks, continue to pose a serious threat to consumers, businesses and the financial assets of many business partners. At the same time, many cyber-attacks are successful through phishing. Statistics presented by APWG (Anti-Phishing Working Group) and Phish Tank organizations, which are known and accepted by researchers, show that the number of phishing sites is constantly increasing (APWG, 2023).

According to APWG's 2022 4th quarter report, Figure 1.1 shows that phishing attacks increased by more than 150% annually in the 4-year period. In the same report, they mention the highest monthly figure they have recorded to date, with a sample of 101104 unique e-mail subjects in October 2022 (APWG, 2023).

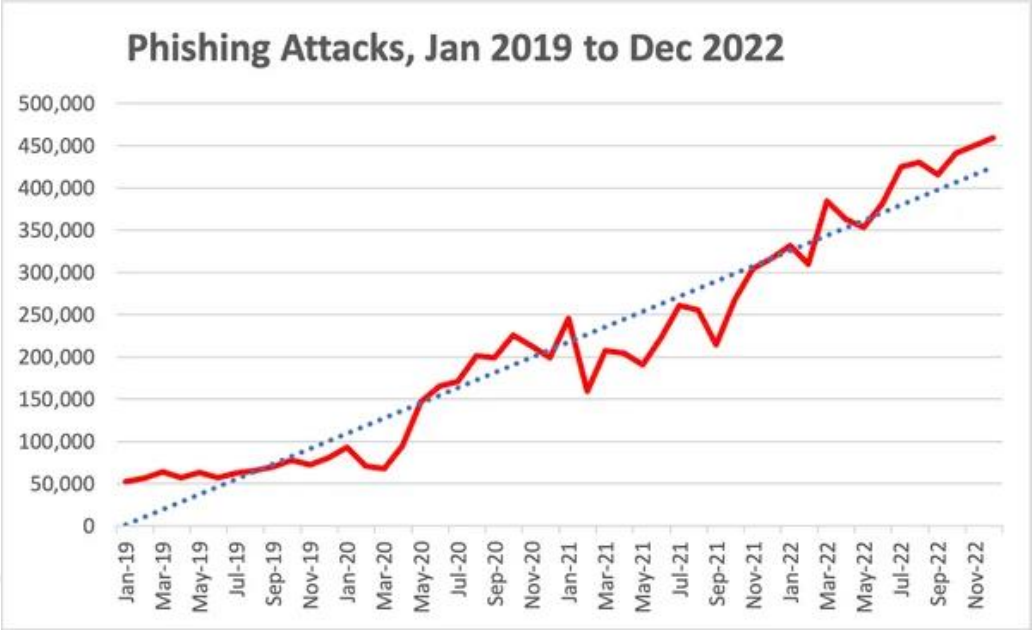

Figure 1.1. Phishing attacks from 2019 to 2022 (APWG, 2023)

As a result, these attacks, which have increased so much, have attracted the attention of many researchers today, so various studies have been carried out and software has been installed on the institution's servers open to the internet to detect deception attempts.

However, phishing attacks have become increasingly complex, difficult to understand, and complex, and have managed to circumvent the filters created against phishing techniques. Successful phishing sites, on the other hand, often confiscate individual data and inflict financial losses on people. Therefore, it is very important to block these sites by reducing them to the level of personal use.

As a countermeasure, phishing attacks can be detected by raising awareness or using software that detects such attacks. Among the techniques such as web page image processing, web page word processing, source code inspection, URL-based inspection, and the use of third-party services (Khonji, Iraqi, & Jones, 2013) .

1.2. Phishing Attacks

Phishing is a type of attack technique used over the internet to trick users into posing as a trusted entity to obtain sensitive and private information such as usernames, passwords, and financial accounts. The main factor that makes phishing a successful cybercrime is deception. The attacker deceives users by sending fake emails, text messages and making a convincing phone call. The number of phishing attacks has increased in recent years. It has increased not only the number of threats, but also the number of different phishing techniques to deceive users and security systems.

The diagram below Figure 1.2, taken from APWG's report, shows the industries targeted in cyber-attacks in the last quarter of 2022 (APWG, 2023).

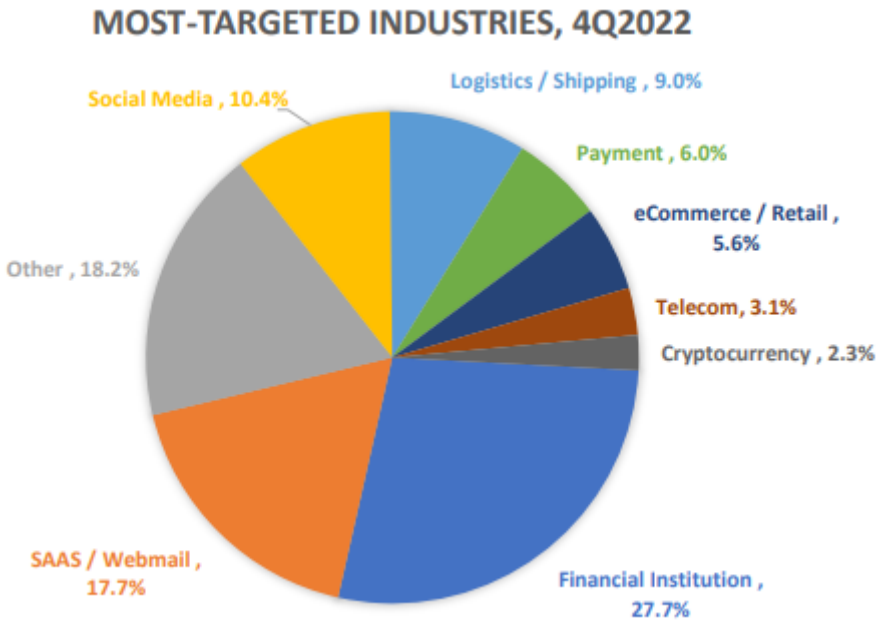

Figure 1.2. Industries targeted in cyber attacks (APWG, 2023)

Phishing attacks have become a permanent threat in the digital environment, targeting both individuals and organizations. Detecting and mitigating these attacks is of high importance for protecting sensitive information and online security. Researchers have explored various

techniques for effectively detecting phishing websites. Research continues on phishing website detection techniques ranging from rule-based methods to machine learning and deep learning methods. Cyber security experts are constantly striving to prevent phishing threats and ensure online security with new methods, and research environments have continued to evolve for many years with the following techniques.

- Traditional Rule-Based Approaches:

Early approaches to phishing website detection relied on rule-based methods that identified patterns and characteristics commonly associated with phishing URLs and content. These techniques often involved analyzing domain names, checking for deceptive characters, and scrutinizing URL structures (Dhamija, Tygar, & Hearst, 2006).

- Machine Learning-Based Approaches:

Phishing website detection data has two classes: phishing websites and non-phishing websites. Machine learning methods are widely used for classification problems. Researchers have proposed robust machine-learning models that have higher accuracy in extracting features from URLs, website content, and user behavior (Fette, Sadeh, & Tomasic, 2007).

- Deep Learning Approaches:

Because of recent advancements in deep learning, Deep learning methods are widely used for classification problems that have huge datasets. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are well-known deep learning models. They have been performed to analyze web page structures, images, and textual content, enabling the identification of phishing websites (Zhang, Xiao, Ghaboosi, Zhang, & Deng, 2012).

- Hybrid and Ensemble Approaches:

The strengths of more than one detection method are combined in a model or multiple sub-stages to increase the accuracy of the models for phishing website detection. Integration of rule-based, machine-learning and deep-learning techniques can be applied to enhance a more comprehensive defense against phishing attacks (Do, Selamat, Krejcar, Herrera-Viedma, & Fujita, 2022).

1.3. Feature Selection

Feature selection methods are used to choose the most relevant features from all features in a dataset for increasing the performance of machine learning models. They can be divided into filtering, wrapper, and hybrid methods. Wrapper methods have a classification algorithm to

evaluate the classification performance for the selected features. However, they have a bigger computation time than filtering methods (Tawhid, 2020). The usage of metaheuristic algorithms to reduce feature size in a dataset is classified as wrapper methods.

If the dataset to be used is optimized for machine learning, it becomes easier and less time consuming to detect the flow or sequence of the process. Optimizing the data set is removing the features of the data set that are unrelated to the task to be learned from the process. This process is known as the feature selection process. The use of feature selection methods can significantly improve the comprehensibility of classification models and generally yield a better generalizing model.

Feature selection methods are implemented to optimize dataset size by separating redundant and irrelevant features. In general, a successful feature selection method provides performance improvements with more effective predictability and less processing time in classification analysis (Liu & Yu, 2005).

The main goal of feature selection is to find a subset of features that yield higher classification accuracy. A reduction in the feature space leads to a more understandable model and simplifies the use of different visualization techniques (Kaytan & Hanbay, 2017).

Feature selection techniques have been applied to various cybersecurity domains, such as intrusion detection, malware analysis, spam detection, and network traffic analysis. In intrusion detection systems, selecting relevant features from network traffic data can enhance the accuracy of detecting and preventing unauthorized access. Similarly, in malware analysis, feature selection aids in identifying distinctive patterns that differentiate malicious software from benign files. The same principle applies to spam detection, where relevant features help distinguish legitimate emails from spam (García, Sánchez, & Mollineda, 2012) .

By selecting only relevant features, the noise introduced by irrelevant or redundant data is minimized, leading to improved detection accuracy in various cybersecurity tasks. Feature selection reduces the dimensionality of data, resulting in reduced computational complexity and resource requirements for training and testing cybersecurity models. A reduced feature set

simplifies the interpretation of model results, enabling cybersecurity professionals to understand the reasoning behind detected threats.

Identifying relevant features can be challenging, especially in rapidly evolving cyber threat landscapes where attackers adopt new tactics. Many cybersecurity datasets suffer from class imbalance, where the number of samples for different classes is skewed. Feature selection must consider this imbalance to prevent bias in detection outcomes. While feature selection can improve interpretability, it might also lead to a trade-off between accuracy and comprehensiveness in some cases.

1.4. Swarm Intelligence Optimization

Swarm intelligence (SI) Algorithms are conducted to search real-world problems and optimize intelligent systems by imitating the collective behavior of social animals such as birds, whales, bees, and wolves. Swarms based on social behaviors to strongly survive in nature have fascinated researchers. Even though the only individual of these colonies are inexperienced individuals, they can perform complex tasks in collaboration. An individual is self-organized according to sharing information from members of swarm without a central control to survive and search for food research (Bonyadi & Michalewicz, 2017). Figure 1.3 shows the general framework of SI algorithm.

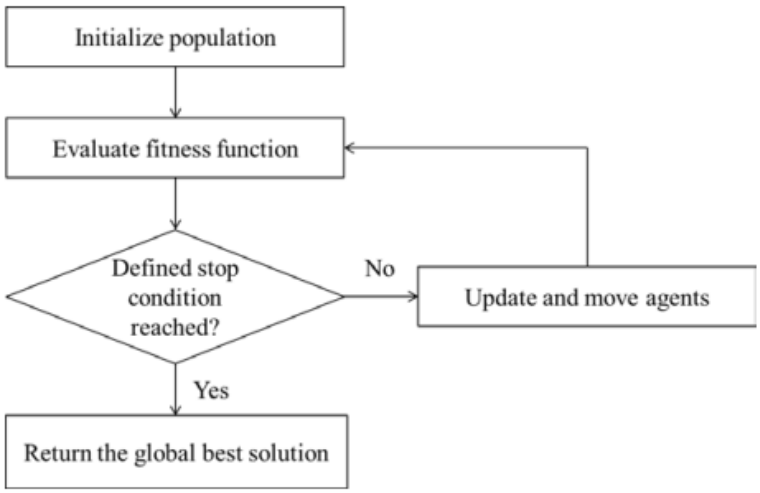

Figure 1.3. General Framework of SI algorithm (Brezočnik, Fister Jr, & Podgorelec, 2018)

Dozens of algorithms based on swarm intelligence have been published over the years, following the best-known PSO, which was proposed in 1995 (Kennedy & Eberhart, Particle swarm optimization, 1995). Particle swarm optimization was modeled by the social behavior of animals and insects. Each individual swarm member is treated as a particle. Collaboration and learning enable the collective intelligence of these scattered particles (Kennedy, Particle Swarm Optimization, 2010). Quantity of algorithms proposed among the years are shown in Figure 1.4.

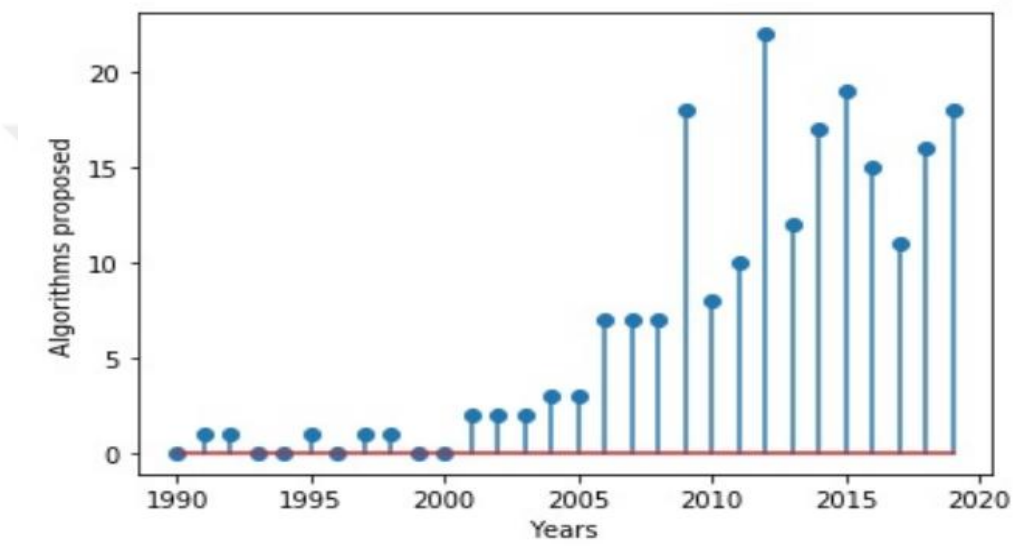

Figure 1.4. Quantity of new proposals for year *(Torres-Treviño, 2021)*

2. LITERATURE REVIEW

The popular solution approaches for phishing detection in cyber security areas are given in this section. A significant number of studies on solving such models by using feature selection techniques are mainly emphasized. In order to show a gap in the literature, research on the use of feature selection in cyber security, development efforts, and the development of feature selection with swarm intelligence are described in this section.

2.1. Feature selection in Cybersecurity

Cyber security consists of protecting systems related to computers, software, mobile, industrial systems, and computer networks. There are two important requirements in the operation of most phishing and intrusion detection systems in cyber security. The first of these is fast access in a real-time environment, and the second is a high detection rate. The mechanisms generally used to detect attacks and phishing are decision-making using machine learning methods. It is important to prepare and provide inputs that facilitate the learning of machine learning systems. Important studies using feature selection and machine learning models in the cyber security area are presented in Table 2.1 (Ali & Ahmed, 2019).

Table 2.1. Important studies using intelligent techniques to detect the phishing websites (Ali & Ahmed, 2019)

Year	**Publication Title**	**ML Approaches**	**FS Methods**
2011	an efficient phishing web page detector	SVM	12 features are generated from web sites for each sample.
2012	an assessment of features related to phishing websites using an automated technique	none	frequency analysis-based feature evaluation
2013	intelligent phishing detection and protection scheme for online transactions	Neuro-Fuzzy	Based on the find function to find and analyze the most frequent features that appear often in a lot of web pages

2014	predicting phishing websites based on self-structuring neural network	self-structuring neural networks	17 features are chosen by frequency analysis-based feature evaluation
2014	intelligent rule-based phishing website classification	some rule-based classification algorithms: C4.5, RIPPER, PRISM, and CBA	frequency analysis-based feature evaluation
2014	an experimental study for assessing email classification attributes using feature selection methods	OneRule, JRip, Part, and J48	CFS, IG, and Chi-square
2014	phishing website detection based on supervised machine learning with wrapper feature selection	BPNN, SVM, C4.5, RBFN, kNN, NB, and RF	wrapper FS
2014	phishing detection-based associative classification data mining	MCAC	–Chi-Square measure
			–frequency analysis-based feature evaluation
2016	new rule-based phishing detection method	SVM	two feature sets related with the page resource identity and the access protocol of page resource elements were extracted through examining the content and the page document object model (DOM)
2018	phishing-aware: a neuro-fuzzy approach for anti-phishing on fog networks	Neuro-Fuzzy	six features are chosen according to user's experience

2018	towards detection of phishing websites on client-side using machine learning-based approach	Random forest (RF)	the features are based on URL and content of the web page.
2018	detection of phishing websites using an efficient feature-based machine learning framework	J48 tree, MLP, SMO, LR, RF, BN, AdaBoostM1 (AM1), and SVM	the features are chosen based on the authors experience
2019	a new hybrid ensemble feature selection framework for machine learning-based phishing detection system	NB, SVM, C4.5, PART, and JRip	a function perturbation ensemble and distribution function gradient

In addition to the studies in Table 2.1, the studies using feature selection-based swarm optimization are given as follows.

Mohammad et al. presented a model using feature selection and self-structuring neural network.to detect phishing websites (Mohammad, Thabtah, & McCluskey, Predicting phishing websites based on self-structuring neural network, 2014). They also used a dataset with 17 features in their study. In 2019, Almseidin et al. presented a model to strengthen the decision-making of the machine learning methods they used by hybridizing filtering and wrapping feature selection methods (Almseidin, Zuraiq, Al-Kasassbeh, & Alnidami, 2019). The observation of a decrease in false alarm rates and an increase in detection rates and accuracy after feature selection with the PSO algorithm was in the IDS which was used by Kunhare et al. (Kunhare, Tiwari, & Dhar, 2020).

Over the next few years, it was seen that similar studies were carried out in the literature. A study about GA-GWO hybrid models was presented by Davahli et al. (Davahli, Shamsi, & Abaei, 2020). This model was aiming to reduce the dimensionality of wireless traffic, which is the main problem of IoT IDS. The team, whose goal was to make the correct feature selection, reduced the data set from 154 features to 92 features. These results noticeably reduced the computation time of the IoT IDS, from 16,530 seconds to 12,075 seconds (Davahli, Shamsi, & Abaei, 2020).

A phishing detection system using Binary Modified Equilibrium Optimizer (BMEO) with an AV-shape transfer function was proposed by (Minocha & Singh, 2022). The use of the transfer functions improves the performance of classification by enhancing the search algorithm's exploration capabilities. They observed that the transfer function which is used in the model enhanced the algorithm's exploration capabilities and classification performance (Minocha & Singh, 2022). In 2022, Alzaqebah et al. proposed to use a modified algorithm called GWO to select the best feature set and eliminate irrelevant and noisy features (Alzaqebah, Aljarah, Al-Kadi, & Damaševičius, 2022). Another study is presented by Moizuddin and Jose (Moizuddin & Jose, 2022). They stated that they used IDS based on GMGWO (Generalized Mean Gray Wolf Algorithm) and ECAE (ElasticNet Contractive AutoEncoder). In the system which consists of two stages, it has been observed that when GMGWO undertakes the feature selection task, the performance of the IDS system, which undertakes the classification task, increases (Moizuddin & Jose, 2022).

In (Alzaqebah, Aljarah, & Al-Kadi, 2023), a swarm intelligence approach for feature and parameter optimization of the extreme learning machine is proposed. By extending the Harris Hawk's optimizer for fast convergence and wide search space, the team observed increased performance across different network attacks.

2.2. Swarm optimization techniques for feature selection

In recent decades, the majority of SI methods have been widely used to select the most relevant feature subset because of the SI's power search mechanism. The popular swarm intelligence approaches for feature selection problems are mentioned in this section.

Mafarja et al. proposed Binary Dragonfly Algorithm using eight time-varying S-shaped and V-shaped transfer functions (Mafarja, et al., 2018). Binary Crow Search Algorithm with time-varying flight length to improve the balance between exploration and exploitation is presented by Chaudhuri & Sahu for feature selection (Chaudhuri & Sahu, 2021). The Authors evaluated the proposed model with 8 different transfer functions and 21 benchmark datasets. In 2021, a time-varying V-shaped transfer function was used as a requirement to replace traditional transfer functions in the BHHO algorithm in (Chantar, Thaher, Turabieh, Mafarja, & Sheta,

2021). The model was performed for eighteen well-known data sets. The results showed that the time variable obtained better results (Chantar, Thaher, Turabieh, Mafarja, & Sheta, 2021).

In 2023, Singh et al. presented techniques that integrated mutual information, correlation, and tree-based feature selection to offer an innovative approach to PSO (Singh, Kumar, & Kumar, 2023). They aim to accurately distinguish between legitimate and phishing websites by selecting relevant attributes. Using a comprehensive database of 111 features, their study shows that efficiency and classification performance are improved with PSO optimization. When the researchers found that they achieved a 94.69% accuracy rate on the dataset they used, they proposed a PSO-based feature selection technique combined with machine learning classifiers for phishing detection (Singh, Kumar, & Kumar, 2023) .

In (Mohd Yusof, Muda, Pratama, & Abraham, 2023), Yusof et al. proposed a Binary Whale Optimization algorithm that uses a new time-varying transfer function to overcome high-dimensional molecular descriptors and low classification accuracy. As a result of their experiments, they found that the time-variable algorithm they proposed was better than the other algorithm in terms of convergence speed, classification performance, number of selected features and stability quality (Mohd Yusof, Muda, Pratama, & Abraham, 2023).

2.3. Motivation

The reviewed literature shows that machine learning and deep learning feature selection methods for phishing detection have better performance according to accuracy and run-time. The time-varying mirrored S-shaped transfer function is performed in (Beheshti, 2020; Kılıç, Kaya, & Yildirim, 2021) and the studies presented promising results for feature selection and binary problems. Furthermore, the time-varying mirrored S-shaped transfer function is not applied in the related cybersecurity literature. In this thesis, the main contributions:

- The time-varying mirrored S-shaped transfer function is applied for BGWO, BPSO, and BHHO.
- The proposed models are performed on the phishing website dataset.
- The proposed models have promising results on the phishing website dataset.

3. MATERIALS AND METHODS

In this study, a novel feature selection model with SI on web fishing dataset is proposed. Time-Varying Mirrored S-Shape (TVMS) transfer function proposed by (Beheshti, 2020) is adapted in HHO, PSO, and GWO for detection of web fishing activities. This model has not been implemented for cyber security studies. This study mainly aims to improve accuracy with a lower feature number and investigate state of art feature selection methods for cyber security.

Web phishing detection methods commonly consist of the following steps:

I. **Data Preparation and Preprocessing:** Prepare the phishing website dataset by extracting relevant features, and labeling instances as legitimate or phishing. Firstly, this process commonly generates dataset from the website's front-end code and web address. Secondly, the dataset is converted to a suitable structure for web phishing detection systems.

II. **Web Phishing Systems:** Machine learning techniques are implemented to detect phishing websites. These systems can have multiple sub stages. For example, artificial intelligence (AI) models such as ANN, SVM, and Decision Tree directly can be used after learning and testing processes. Generally, data reduction (feature selection) and optimization techniques are performed to improve AI models.

    a. **Feature Selection Methods:** These methods aim to select the most relevant feature subset to decrease the feature number and increase the accuracy of the models. In this study, the latest binary transfer function is applied to Binary PSO, HHO, and GWO algorithms.

    b. **Hyper-Parameter Optimization:** AI models have some crucial parameters. For example, ANN has a learning rate, momentum rate, and epoch number. These parameter values must be carefully selected to obtain better results.

    c. **Evaluation:** Evaluate the performance of the proposed models and methods on a test dataset without train dataset using metrics like accuracy, F1, precision, recall, and minimum number of selected features.

3.1. Phishing Websites Dataset

Although website phishing studies are widely investigated in the literature, there are a few website phishing datasets and most of them are not available or published for public usage. In this study, the dataset published by Mohammad et al. under the CC BY 4.0 license, an open source, is used to compare the proposed models. This data set was collected from the Phish Tank archive, Miller Smiles archive, and Google search operators and published in (Mohammad & McCluskey, 2015). The authors prepared this dataset because they encountered the difficulty of finding a reliable dataset in their research.

The data set has 11055 rows with 31 attributes. Dataset phishing criteria are divided into 4 sections (Abnormal Based Features, Address Bar based Features, Domain based Features, HTML, and JavaScript based Features). These features are given in the Table 3.1 (Mohammad, Thabtah, & McCluskey, Phishing websites features, 2015).

Table 3.1. List of features for Phishing Websites Dataset

Address Bar based Features	Abnormal Based Features	HTML and JavaScript based Features	Domain based Features
Using the IP Address	Request URL	Website Forwarding	Age of Domain
Long URL to Hide the Suspicious Part	URL of Anchor	Status Bar Customization	DNS Record
Using URL Shortening Services “TinyURL”	Links in <Meta>, <Script> and <Link> tags	Disabling Right Click	Website Traffic
URL’s having “@” Symbol	Server Form Handler (SFH)	Using Pop-up Window	PageRank
Redirecting using “//”	Submitting Information to Email	IFrame Redirection	Google Index
Adding Prefix or Suffix Separated by (-) to the Domain	Abnormal URL		Number of Links Pointing to Page
Sub Domain and Multi Sub Domains			Statistical-Reports Based Feature
HTTPS (Hyper Text Transfer Protocol with Secure Sockets Layer)			

The Existence of "HTTPS" Token in the Domain Part of the URL			
Favicon			
Domain Registration Length			
Using Non-Standard Port			

3.2. Optimization algorithms

3.2.1. Particle Swarm Optimization

Particle Swarm Optimization (PSO) was introduced by (Kennedy & Eberhart, Particle swarm optimization, 1995) in 1995. PSO involves a swarm of particles moving through a solution space to find the good solution near to optimal solution. In classification problems, PSO can optimize model parameters, such as weights and biases of neural networks, to improve classification accuracy.

Particle swarm optimization has been considered one of the most effective optimization approaches for decades due to its ease of use, number of parameters that can be used for modification, rapid convergence, and scalability. Additionally, PSO (Particle Swarm Optimization) is one of the oldest swarm-based algorithms. This algorithm allows the people in a simple living structure to join forces to create a more intelligent structure (Kennedy & Eberhart, Particle swarm optimization, 1995). Algorithm 3.1 shows the pseudo code of PSO (Kennedy & Eberhart, Particle swarm optimization, 1995; Kılıç, Kaya, & Yildirim, 2021).

Initialization: v_i , x_i ,p_{best}, g_{best}, *maxIter*
Generate random particles P
for *each particle (i)* ***do***
    Compute fitness value (fi)
    Update (p^i_{best}) , (g_{best})
End
while *t < maxIter* ***do***
    for *each particle (i)* ***do***
        uptade (x_i,v_i)
        Calculate fitness function (f_i)
        Update (p^i_{best}) , (g_{best})
    End
    Increment *t*
End

Algorithm 3.1. PSO algorithm (Kennedy & Eberhart, 1995)

According to the particle swarm optimization algorithm stages, there are three prominent elements: x, v particle and fitness function parameters. Correct specification of the fitness function is a prerequisite for using metaheuristic algorithms. The fitness function shows the quality of the solution. The accuracy value is selected for the fitness function in this study. v is the current velocity of the particle and x is its position. x is a vector consisting of 0 or 1. If a feature is selected, corresponding position is 1 otherwise 0. v and x vectors of each particle are adjusted according to personal best (p_{best}^{i}) position and population global best position (g_{best}). v and x are computed using Eqs. (2) and (4) respectively. S(θ) is S-Shape function and is used to convert from a number to a value in between 0 to 1.

$$w^t = w_{max} - \frac{(w_{max} - w_{min}).t}{maxIter} \tag{1}$$

$$v_i^{t+1} = w^t.v_i^t + c_1.rand.(p_{best}^i - x_i) + c_2.rand.(g_{best} - x_i) \tag{2}$$

$$S(\theta) = \frac{1}{1 + e^{-\theta}} \tag{3}$$

$$x_i = \begin{cases} 1\ rand < S(v_i^t) \\ 0, otherwise \end{cases} \tag{4}$$

In general, these two elements influence the convergence of the particle. These two factors are very important because they determine whether the particle eventually comes to a solution. In addition, the number of particles is a parameter that directly affects the solution. In fact, as the number increases, the result improves, but the algorithm works more slowly.

A binary version of PSO (BPSO) using a sigmoid transfer function was introduced to solve discrete optimization problems (Kennedy & Eberhart, 1997). In this study, BPSO and TMVS-BPSO are adapted for phishing website detection. Algorithm 3.4 presents TMVS-BPSO proposed in (Beheshti, 2020) .

3.2.2. Grey Wolf Optimization

The GWO method is introduced by taking inspiration from the leadership hierarchy and hunting behaviors of grey wolves. Types of grey wolves in nature are divided into alpha, beta, delta, and omega according to behaviors and leader levels (Mirjalili, Mirjalili, & Lewis, 2014).

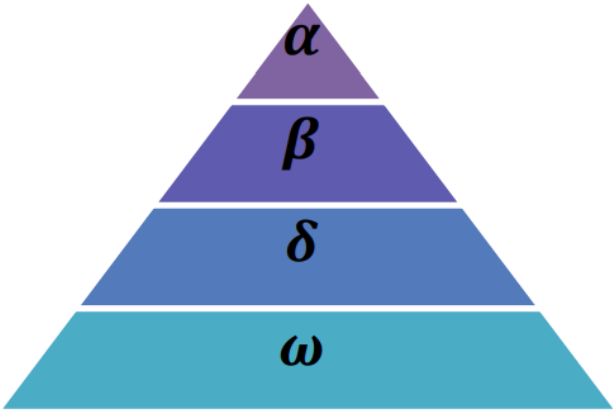

Figure 3.1. Hierarchy of grey wolfs

The application of the grey wolf algorithm begins with tracking and following the prey and continues with surrounding and attacking behaviors. Considering the hierarchy in the gray wolf pack as seen in Figure 3.1, the most appropriate solution is considered alpha (α). Beta (β) and delta (δ) are assumed as the second and third best solutions, respectively. Omega wolves (ω), which are at the bottom of the hierarchy, follow the other three wolves. Once the application is started, gray wolves track their prey, determine its location and begin to surround it (Mirjalili, Mirjalili, & Lewis, 2014). Algorithm 3.2 shows pseudo code of the GWO algorithm (Mirjalili, Mirjalili, & Lewis, 2014).

The GWO algorithm finds solutions to continuous optimization problems. A novel two binary versions named BGWO1 and BGWO2 which are converting continuous versions into binary form was proposed by Emary et al. for binary optimization problems in 2016. The new forms of GWO are using crossover position updating, converting the position into a binary vector with sigmoid function (Emary, Zawbaa, & Hassanien, 2016). In this study, BGWO2 is used for the detection of phishing websites.

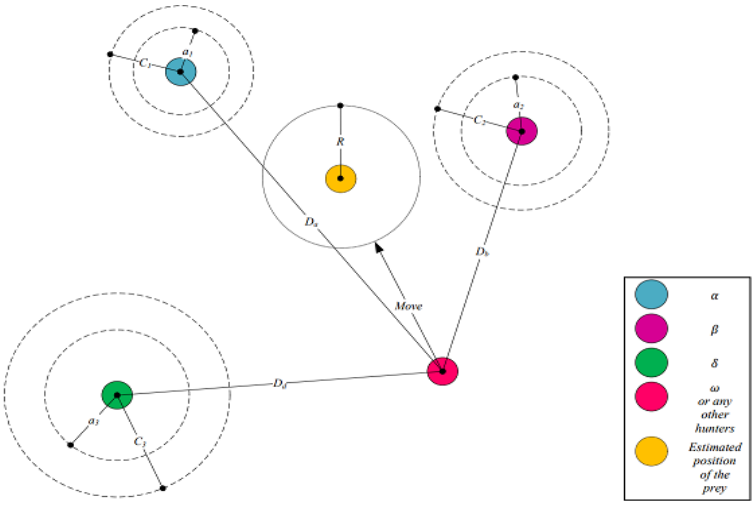

Figure 3.2. Position updating in GWO (Mirjalili, Mirjalili, & Lewis, 2014)

Initialize solutions for the grey wolf in the population X_i *(i = 1, 2, …, n)*
Initialization of C, a, and A
Calculate the fitness value of each wolf (search agent)
X_α *= the best search agent*
X_β *= the second best search agent*
X_δ *= the third best search agent*
while *(t < maxIter)*
    for *each search agent (*X_i*)*
        Update the position of X_i
    end for
    Update C, a, and A
    Calculate the fitness values for all wolves (search agents)
    Update positions of X_α *,* X_β *,* X_δ
    Increment t
end while
return X_α

Algorithm 3.2. Pseudo code of GWO search algorithm (Mirjalili, Mirjalili, & Lewis, 2014)

3.2.3. Harris Hawks Optimization

Harris Hawks Optimizer (HHO) is a new population based, nature-inspired optimization paradigm which is proposed in 2019 in (Heidari, et al., 2019). HHO mimics the hunting patterns, surprise attack and different attack strategies of Harris hawks in nature to adapt the algorithm. In this proposal, candidate solutions are represented by hawks. The best solution was considered as hunting. Harris hawks try to track down their prey using their powerful eyes. They catch the detected prey by making surprise attacks (AKTAŞ & KILIÇ, 2021).

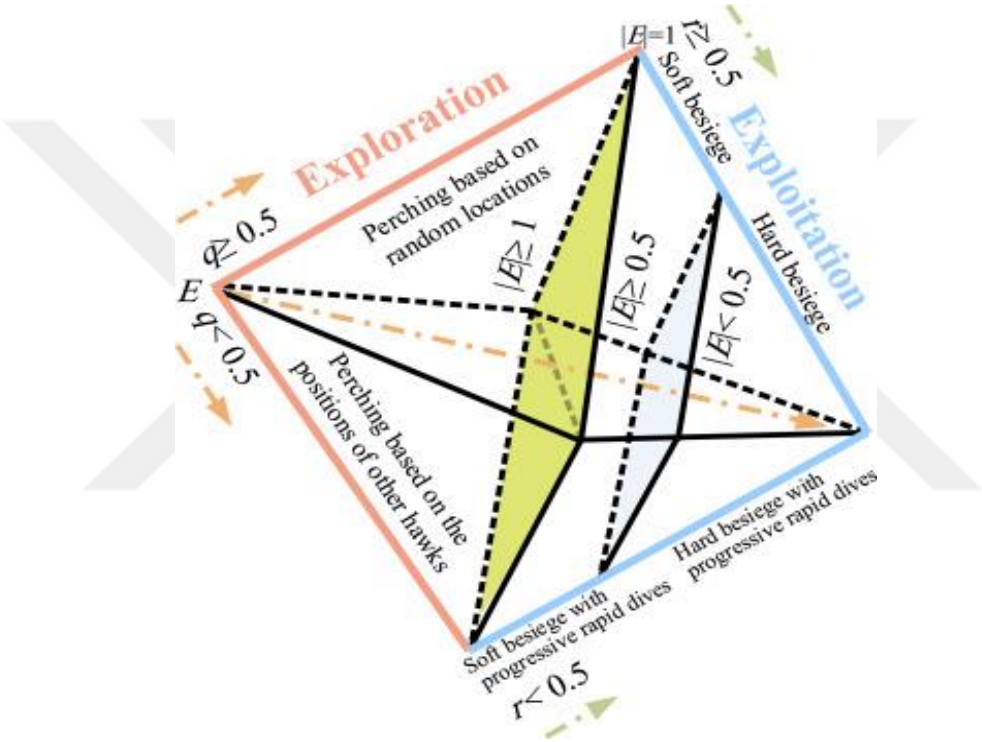

Figure 3.3. Different phases of Harris Hawks Algorithm (Heidari, et al., 2019)

A brief explanation of these phases is below:

1- Exploration phase: In this phase, it updates the hawk's locations via random location and other hawks' locations.

2- Transition from exploration to exploitation: It is the phase in which different exploitative behaviors are transitioned depending on the energy of the escaped prey.

3- Exploitation phase: It is the phase in which the escape behavior of the prey and the attack behavior of the Harris hawk are shaped. At this stage, the following 4 different attack models are applied according to the decisions made in the algorithm.

i. Soft besiege

ii. Hard besiege

iii. Soft besiege with progressive rapid dives

iv. Hard besiege with progressive rapid dives.

The pseudocode of the HHO algorithm is shown in Algorithm 3.3 (Heidari, et al., 2019).

Inputs(*The population size N and maximum iterations maxIter)*
Outputs(*The position of rabbit and its objective function value)*
Initialization of the random population, X_i (i=1, 2, …, N)
while *(t < maxIter)* ***do***
  Calculate the fitness values of hawks
  Set X_{rabbit} as the position of rabbit (best location)
  for *(each hawk (X_i))* ***do***
    Update the initial energy E_0 and jump strength J — ***E_0 =2*rand( )-1, J=2(1-rand( ))***
    Update the E
    if *($|E_n| \geq 1$)* ***then*** — ***Exploration condition***
      Update the location vector
    if *($|E_n| < 1$)* ***then*** — ***Exploitation condition***
      if *($r \geq 0.5$ & $|E_n| \geq 0.5$)* ***then*** — ***Soft besiege***
        Update the location vector
      else if *($r \geq 0.5$ & $|E_n| < 0.5$)* ***then*** — ***Hard besiege***
        Update the location vector
      else if *($r < 0.5$ & $|E_n| \geq 0.5$)* ***then*** — ***Soft besiege with progressive rapid dives***
        Update the location vector
      else if *($R < 0.5$ & $|E_n| < 0.5$)* ***then*** — ***Hard besiege with progressive rapid dives***
        Update the location vector
  Increment ***t***

Return *X_{rabbit}*

Algorithm 3.3. The pseudocode of the HHO algorithm (Heidari, et al., 2019)

In 2019, binary versions of HHO (BHHO and QBHHO) using an S-shaped or V-shaped transfer function to convert the continuous variable into a binary one are proposed by Too et al. QBHHO, which is the second of the algorithms presented in the same study, was recommended to show that BHHO does not solve the feature selection problem effectively and is still far from perfect (Too, Abdullah, & Mohd Saad, 2019). In this study, BHHO is used for the detection of phishing websites.

3.3. S-Shaped and Mirrored S-Shaped Transfer Function for Swarm Intelligence Optimization

The first versions of SI optimization algorithms are commonly introduced for continuous optimization problems. They are later adapted for binary and other optimization problems. S-shape transfer function is widely used to convert a binary dimension from continuous values (Kılıç, Kaya, & Yildirim, 2021). Eq. (3) shows S-Shape functions and Figure 3.4 displays the graph of S-Shape function between -10 and +10 (Beheshti, 2020).

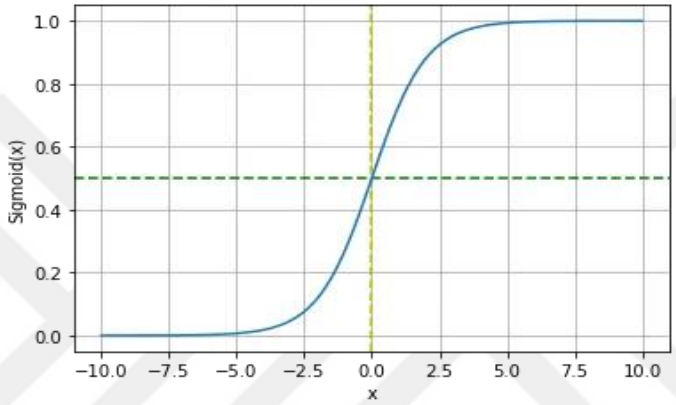

Figure 3.4. Graph of a sigmoid function

S-Shape function is also used in neural networks as an activation function. S-Shape function is adapted to Binary PSO, GWO, and HHO as Eqs. (5), (6), and (7) respectively.

$$TF = 1 / (1 + e^{(-V (i,d))}) \text{ for BPSO} \quad (5)$$

$$TF = 1 / (1 + e^{(-10 * (Xn - 0.5))}) \text{ for BGWO} \quad (6)$$

$$TF = 1 / (1 + e^{(-Xn)}) \text{ for BHHO} \quad (7)$$

In the literature, there are different transfer functions and versions of S-Shape function for binary conversion. In 2020, Beheshti proposed the time-varying mirrored S-shape (TVMS) for binary test problems and Kilic et. al. used TVMS transfer function for feature selection problems. The mirrored S-shape transfer function is designed to replicate the sigmoid activation function often used in neural networks. Its time-varying nature allows it to adapt and optimize its parameters as data patterns evolve over time, making it well-suited for dynamic datasets like those encountered in phishing detection.

The time-varying mirrored S-shape transfer function is a mathematical model that captures the dynamic behavior of complex systems. This function is particularly relevant to binary discrete problems like feature selection problems.

SI algorithms have iterated structures and some parts of them are repeated until maximum iteration (maxIter). Beheshti changed S-Shape according to iteration time (Beheshti, 2020). σ is a time-varying variable and computed using Eq. (8). In Eq. (8), σ_{min} and σ_{max} parameters minimum and maximum values of σ. It is initialized by σ_{max} and gradually decreased to σ_{min} in order to switch smoothly from exploration to exploitation as seen at Figure 3.5 (Beheshti, 2020).

$$\sigma = (\sigma_{max} - \sigma_{min}) \frac{\text{iter}}{\text{maxIter}} + \sigma_{min} \tag{8}$$

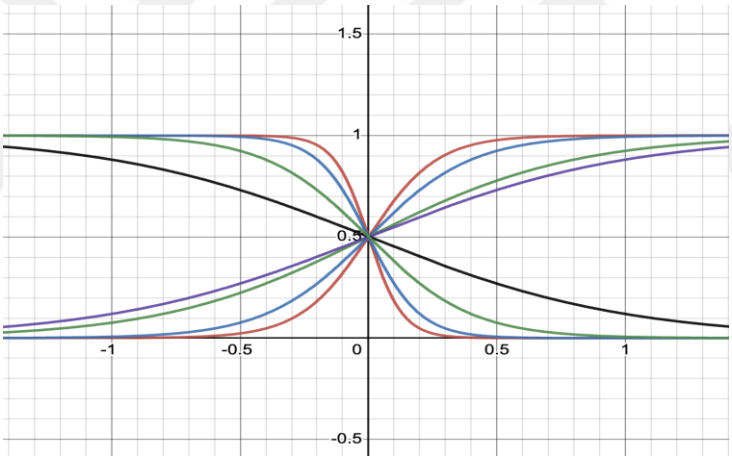

Figure 3.5. The proposed transfer functions with different values of control parameter σ (Beheshti, 2020)

Time-varying mirrored S-shape transfer function is adapted to Binary PSO, GWO, and HHO as Eqs. (9 and 10), (11 and 12), and (13 and 14) respectively.

$$TF = 1 / (1 + e^{(\sigma * (-V(i,d)))}) \text{ for BPSO} \tag{9}$$

$$TF' = 1 / (1 + e^{(\sigma * (V(i,d)))}) \text{ for BPSO} \tag{10}$$

$$TF = 1 / (1 + e^{(\sigma * (-10 * (Xn - 0.5)))}) \text{ for BGWO} \tag{11}$$

$$TF' = 1 / (1 + e^{(\sigma * (10 * (Xn - 0.5)))}) \text{ for BGWO} \quad (12)$$

$$TF = 1 / (1 + e^{(\sigma * (-Xn))}) \text{ for BHHO} \quad (13)$$

$$TF' = 1 / (1 + e^{(\sigma * (Xn))}) \text{ for BHHO} \quad (14)$$

Algorithm 3.4 shows the pseudocode of TMVS part of the proposed BPSO model (Beheshti, 2020). Sigmoid function f (x) is rotationally symmetric with respect to the point (0.5). The sigmoid function and its reflection are symmetric about the vertical axis. The reflection of the sigmoid function about the vertical axis is given by f (-x). We can also say that the sum of the sigmoid function f (x) and its reflection about the vertical axis f (-x) is 1 because the sigmoid function outputs the probability values and as we know that the sum of the probabilities in a probability distribution is always 1 (Shukla, 2023).

As a reference of properties of sigmoid functions and the propose of (Beheshti, 2020) about mirroring in BPSO which is seen in Figure 3.5, the proposed transfer functions are mirrored in BPSO, BGWO, and BHHO for all variables. The main goal is to scan the entire search space. Mirroring transfer functions is adapted in Eqs. (10, 12, and 14).

REPEAT
  For each particle i Do
    For each Dimension d Do
      Calculate v_i^d for global and local topology
      Compute TF with v_i^d (t+1) and TV(σ) as RSLT
      if *$rand_1$ < RSLT* ***then***
        $P_i^d(t+1) = 1$
      else
        $P_i^d(t+1) = 0$
      end if
      Compute TF′ with v_i^d (t+1) and TV(σ) as RSLT′
      if *$rand_2$ > RSLT′* ***then***
        $P'^d_i(t+1) = 1$
      else
        $P'^d_i(t+1) = 0$
      end if
    end for d
    Compute position $xb_i(t+1)$:
    if *Fitness ($P_i^d(t+1)$) better than Fitness ($P'^d_i(t+1)$)* ***then***
      $xb_i(t+1) = P_i^d(t+1)$

else

$xb_i(t+1) = P'^{d}_{i}(t+1)$

end if

end for i

Calculate g_{best}

Compute σ

UNTIL *certain stopping criteria is met*

Algorithm 3.4. The pseudocode of TMVS part of the proposed BPSO model (Beheshti, 2020)

4. RESULTS AND DISCUSSIONS

4.1. Results

In this thesis, a novel feature selection method using SI algorithms (BPSO, BGWO, BHHO) for phishing website detection is proposed. The performance of the proposed models is investigated using phishing websites dataset published by (Mohammad, Thabtah, & McCluskey, 2015). The methods and models are implemented in Matlab and are run on a computer with AMD Ryzen 5 5500u 2.1 GHz, and 16 GB RAM. We modified the SI algorithms by adding TVMS transfer function. The modified methods are called TVMS-BPSO, TVMS-BGWO, and TVMS-BHHO, when the standard methods are called BPSO, BGWO, and BHHO in this study. ReliefF is also used to reduce the number of features and to compare with the proposed algorithms. The experiments are performed 10 times for each algorithm to show the robustness of the algorithms. Table 4.1 shows the parameters of the proposed algorithms.

Table 4.1. The parameter values of the proposed algorithms

Parameter Name	Common parameter values
Max. Iteration	100
Swarm size	20
σ_{max}	1
σ_{min}	0,1
the number of runs	10
BPSO and TVMS-BPSO	
C1	2
C2	2
W_{max}	0.9
W_{min}	0.4

In Table 4.2, the values of the objective function with their average, minimum, maximum, and standard deviation are shown, and the table also shows the average values of Accuracy, Recall, Precision, F1 Score, for 10 runs using the proposed algorithms and ReliefF feature selection algorithm. The best values of the indicator are shown as bold. The feature rank scores are computed using ReliefF and 8 and 16 features are selected according to the highest rank values of them. The proposed SI algorithms achieve similar performances. However, the proposed SI algorithms with TVMS generate better statistical performance than the others considering

fitness and accuracy values. TVMS-BGWO has the best results without standard deviation. The accuracy and other indicators values of selected 8 and 16 features using ReliefF and all features are given in Table 4.2. The results of ReliefF models are not better than the prosed SI algorithms. To show the power of the proposed algorithms, the comparison of ReliefF is performed. It can be clearly seen in Table 4.2 that all proposed SI algorithms provide higher-quality results than the classical feature selection method.

Table 4.2. The performance of the proposed algorithms according to the statistical results of objective function, Recall, Precision, and F1 Score.

	Fitness Values				**Acc**	**Recall**	**Precision**	**F1 Score**
	Avg.	**Min.**	**Max.**	**Stn.**	**Avg.**			
BPSO	0,0466	0,0443	0,0502	0,0020	0,9534	0,9722	0,9457	0,9587
TVMS-BPSO	0,0462	0,0439	0,0475	**0,0012**	0,9538	0,9729	0,9457	0,9591
BGWO	0,0452	0,0421	0,0502	0,0028	0,9548	0,9723	0,9480	0,9600
TVMS-BGWO	**0,0429**	**0,0412**	0,0462	0,0016	**0,9571**	**0,9761**	**0,9484**	**0,9621**
BHHO	0,0477	0,0452	0,0507	0,0018	0,9523	0,9721	0,9438	0,9578
TVMS-BHHO	0,0474	0,0448	0,0498	0,0018	0,9526	0,9716	0,9448	0,9580
ReliefF	Top 8 features				0,9154	0,9432	0,9085	0,9255
	Top 16 Features				0,9376	0,9554	0,9341	0,9446
ALL FEATURES	30 Features				0,9417	0,9586	0,9380	0,9482

In Table 4.3 and Figure 4.1, the number of selected features for each run of the proposed algorithms is presented. The statistical information of selected feature numbers is not considered to decrease since the number of selected features is not used in the fitness function. Therefore, the algorithms only aim to increase the accuracy. However, BGWO generates the solution having the least feature number according to minimum and average of the selected feature number. The accuracy of the solution having the least features is 0.9520. The best solution of TVMS-BGWO according to accuracy has 19 features and its accuracy is 0.9588.

Table 4.3. The number of selected features for each run of the proposed algorithms

NUMBER OF SELECTED FEATURES PER RUN						
RUN	**BPSO**	**TVMS-BPSO**	**BGWO**	**TVMS-BGWO**	**BHHO**	**TVMS-BHHO**
1	18	19	20	19	24	25
2	20	20	18	21	21	23
3	19	18	20	18	25	16
4	19	22	21	22	21	22
5	22	17	20	22	25	22
6	20	19	18	19	24	22
7	22	23	18	22	20	21
8	17	21	17	19	24	22
9	22	20	20	22	22	22
10	21	19	10	20	21	22
AVERAGE	**20**	**19,8**	**18,2**	**20,4**	**22,7**	**21,7**
Min.	**17**	**17**	**10**	**18**	**20**	**16**
Max.	**22**	**23**	**21**	**22**	**25**	**25**

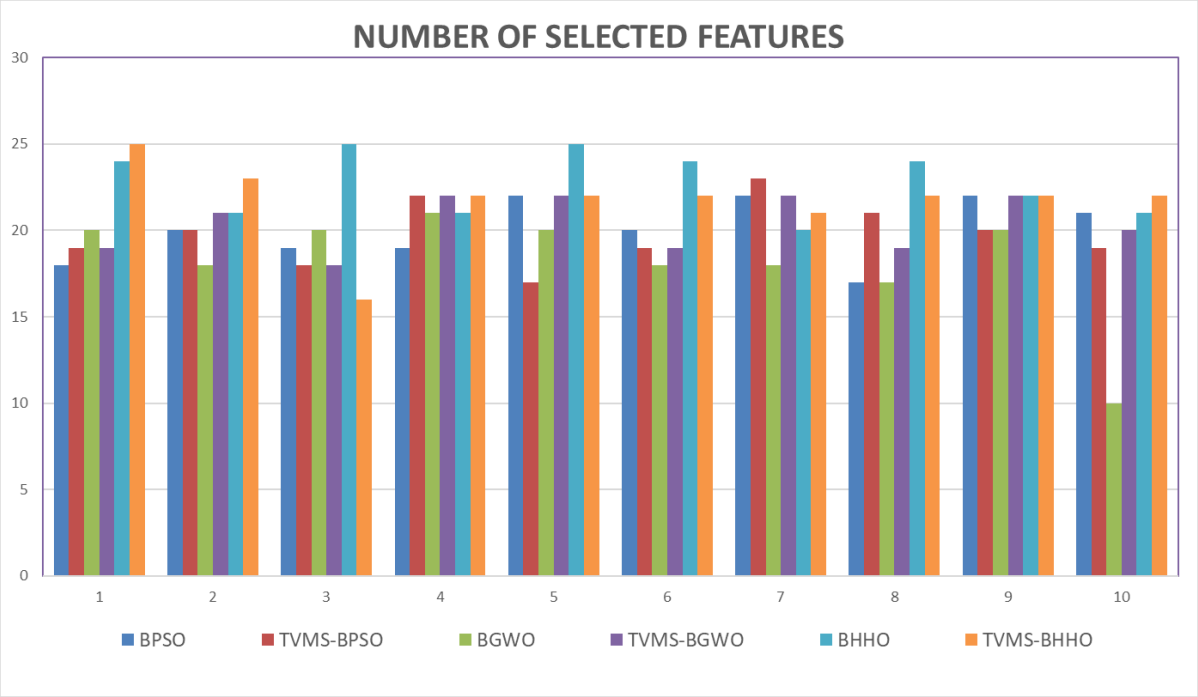

Figure 4.1. The number of selected features for each run

In Figure 4.2, the average convergence rates of BPSO, TVMS-BPSO, BGWO, and TVMS-BGWO, BHHO, and TVMS-BHHO algorithms on the dataset for 10 runs are presented. It can be seen from the figure that the TVMS-BGWO algorithm can improve the convergence ability compared to that of other proposed methods. Furthermore, the proposed algorithms using TVMS have better convergence compared with their standard algorithms.

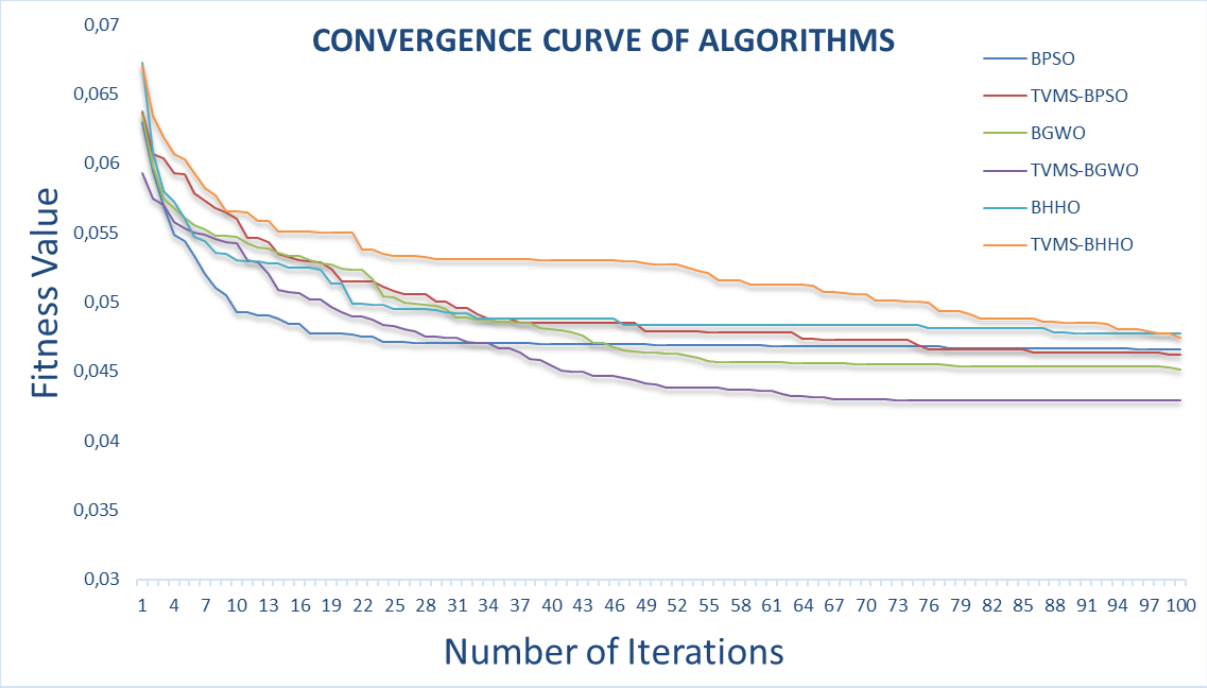

Figure 4.2. 10-run average convergence rates of the proposed models

5. CONCLUSION

Due to the increasing use of internet, researchers widely study to prevent cybercrimes. One of the cybercrimes is phishing websites. In this study, a novel phishing website detection method is proposed for feature selection in Machine Learning Models. TVMS transfer function is applied on BPSO to solve well known binary problems by (Beheshti, 2020) and the model has promising results. Furthermore, Kilic et al. presented good results using BPSO and Multi population PSO with TVMS for feature selection (Kılıç, Kaya, & Yildirim, 2021). The transfer functions of these studies are not investigated for phishing web site and BHHO, BGWO algorithms. The main contribution of this model is that the TVMS function is implemented BHHO, BGWO, and BPSO on the phishing websites dataset. In the experimental results, the proposed models have promising results. Although it is seen in this study that there was no direct connection between the number of features and the accuracy obtained, it was seen that it was important to choose the features that gave high information gain at the optimum level. Average convergence values are lower when TVMS are used for each algorithm. It was observed that these values decreased from 0.0452 to 0.0429 for BGWO, from 0.0466 to 0.0462 for BPSO, and from 0.0477 to 0.0474 for BHHO. Another the result in this study is that the TVMS-BGWO algorithm reached an average accuracy value of 0.9571 with 20.40 selected mean attributes, and the same algorithm reached an accuracy value of 0.9588 with 19 attributes in an individual run. The small differences obtained in the modifications of the algorithms are more valuable than the obvious differences in the results with the high 0.9417 obtained according to the selections made with Weka's ReliefF Attribute Evaluator. The models have acceptable results in detecting phishing websites. In the feature works, the model will be extended using other latest SI algorithms such as firefly, artificial fish swarm, felines, and bacterial foraging swarm algorithms.

REFERENCES

AKTAŞ, M., & KILIÇ, F. (2021). İkili Gri Kurt ve İkili Harris Şahin Optimizasyonları ile Web Haber Sayfalarının Sınıflandırılması. *Avrupa Bilim ve Teknoloji Dergisi*, 234–241.

Ali, W., & Ahmed, A. A. (2019). Hybrid intelligent phishing website prediction using deep neural networks with genetic algorithm-based feature selection and weighting. *IET Information Security, 13*, 659-669. doi:https://doi.org/10.1049/iet-ifs.2019.0006

Almseidin, M., Zuraiq, A. A., Al-Kasassbeh, M., & Alnidami, N. (2019). Phishing detection based on machine learning and feature selection methods.

Alzaqebah, A., Aljarah, I., & Al-Kadi, O. (2023). A hierarchical intrusion detection system based on extreme learning machine and nature-inspired optimization. *Computers & Security, 124*, 102957.

Alzaqebah, A., Aljarah, I., Al-Kadi, O., & Damaševičius, R. (2022). A modified grey wolf optimization algorithm for an intrusion detection system. *Mathematics, 10*, 999.

APWG. (2023). Phishing Activity Trends Report4th Quarter 2022. *Phishing Activity Trends Report4th Quarter 2022*. https://docs.apwg.org/reports/apwg_trends_report_q4_2020.pdf. adresinden alındı

Beheshti, Z. (2020). A time-varying mirrored S-shaped transfer function for binary particle swarm optimization. *Information Sciences, 512*, 1503-1542. doi:https://doi.org/10.1016/j.ins.2019.10.029

Bonyadi, M. R., & Michalewicz, Z. (2017). Particle Swarm Optimization for Single Objective Continuous Space Problems: A Review. *Evolutionary Computation, 25*, 1-54. doi:10.1162/EVCO_r_00180

Brezočnik, L., Fister Jr, I., & Podgorelec, V. (2018). Swarm intelligence algorithms for feature selection: a review. *Applied Sciences, 8*, 1521.

Chantar, H., Thaher, T., Turabieh, H., Mafarja, M., & Sheta, A. (2021). BHHO-TVS: A binary harris hawks optimizer with time-varying scheme for solving data classification problems. *Applied Sciences, 11*, 6516.

Chaudhuri, A., & Sahu, T. P. (2021). Feature selection using Binary Crow Search Algorithm with time varying flight length. *Expert Systems with Applications, 168*, 114288.

Davahli, A., Shamsi, M., & Abaei, G. (2020). Hybridizing genetic algorithm and grey wolf optimizer to advance an intelligent and lightweight intrusion detection system for IoT wireless networks. *Journal of Ambient Intelligence and Humanized Computing, 11*, 5581–5609.

Dhamija, R., Tygar, J. D., & Hearst, M. (2006). Why phishing works. *Proceedings of the SIGCHI conference on Human Factors in computing systems*, (s. 581–590).

Do, N. Q., Selamat, A., Krejcar, O., Herrera-Viedma, E., & Fujita, H. (2022). Deep Learning for Phishing Detection: Taxonomy, Current Challenges and Future Directions. *IEEE Access, 10*, 36429-36463. doi:10.1109/ACCESS.2022.3151903

Emary, E., Zawbaa, H. M., & Hassanien, A. E. (2016). Binary grey wolf optimization approaches for feature selection. *Neurocomputing, 172*, 371-381. doi:https://doi.org/10.1016/j.neucom.2015.06.083

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

García, V., Sánchez, J. S., & Mollineda, R. A. (2012). On the effectiveness of preprocessing methods when dealing with different levels of class imbalance. *Knowledge-Based Systems, 25*, 13-21. doi:https://doi.org/10.1016/j.knosys.2011.06.013

Heidari, A. A., Mirjalili, S., Faris, H., Aljarah, I., Mafarja, M., & Chen, H. (2019). Harris hawks optimization: Algorithm and applications. *Future generation computer systems, 97*, 849–872.

Kaytan, M., & Hanbay, D. (2017). Effective Classification of Phishing Web Pages Based on New Rules by Using Extreme Learning Machines. https://api.semanticscholar.org/CorpusID:59441964 adresinden alındı

Kennedy, J. (2010). Particle Swarm Optimization. C. Sammut, & G. I. Webb (Dü) içinde, *Encyclopedia of Machine Learning* (s. 760–766). Boston, MA: Springer US. doi:10.1007/978-0-387-30164-8_630

Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. *Proceedings of ICNN'95-international conference on neural networks, 4*, pp. 1942–1948.

Kennedy, J., & Eberhart, R. C. (1997). A discrete binary version of the particle swarm algorithm. *1997 IEEE International conference on systems, man, and cybernetics. Computational cybernetics and simulation, 5*, s. 4104–4108.

Khonji, M., Iraqi, Y., & Jones, A. (2013). Phishing Detection: A Literature Survey. *IEEE Communications Surveys & Tutorials, 15*, 2091-2121. doi:10.1109/SURV.2013.032213.00009

Kılıç, F., Kaya, Y., & Yildirim, S. (2021). A novel multi population based particle swarm optimization for feature selection. *Knowledge-Based Systems, 219*, 106894.

Kunhare, N., Tiwari, R., & Dhar, J. (2020). Particle swarm optimization and feature selection for intrusion detection system. *Sādhanā, 45*, 1–14.

Liu, H., & Yu, L. (2005). Toward integrating feature selection algorithms for classification and clustering. *IEEE Transactions on Knowledge and Data Engineering, 17*, 491-502. doi:10.1109/TKDE.2005.66

Mafarja, M., Aljarah, I., Heidari, A. A., Faris, H., Fournier-Viger, P., Li, X., & Mirjalili, S. (2018). Binary dragonfly optimization for feature selection using time-varying transfer functions. *Knowledge-Based Systems, 161*, 185–204.

Minocha, S., & Singh, B. (2022). A novel phishing detection system using binary modified equilibrium optimizer for feature selection. *Computers & Electrical Engineering, 98*, 107689.

Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey wolf optimizer. *Advances in engineering software, 69*, 46–61.

Mohammad, R. M., Thabtah, F., & McCluskey, L. (2014). Predicting phishing websites based on self-structuring neural network. *Neural Computing and Applications, 25*, 443–458.

Mohammad, R. M., Thabtah, F., & McCluskey, L. (2015). Phishing websites features. *School of Computing and Engineering, University of Huddersfield*.

Mohammad, R., & McCluskey, L. (2015). Phishing Websites. *Phishing Websites*.

Mohd Yusof, N., Muda, A. K., Pratama, S. F., & Abraham, A. (2023). A novel nonlinear time-varying sigmoid transfer function in binary whale optimization algorithm for descriptors selection in drug classification. *Molecular diversity, 27*, 71–80.

Moizuddin, M. D., & Jose, M. V. (2022). A bio-inspired hybrid deep learning model for network intrusion detection. *Knowledge-Based Systems, 238*, 107894.

Shukla, P. (2023). *https://pub.towardsai.net/the-sigmoid-function-a-key-building-block-in-neural-networks-c417a8748190*. https://pub.towardsai.net/. adresinden alındı

Singh, T., Kumar, M., & Kumar, S. (2023). Enhancing Phishing Website Detection Using Particle Swarm Optimization and Feature Selection Techniques. *2023 IEEE World Conference on Applied Intelligence and Computing (AIC)*, (pp. 977–982).

Thabtah, F., Mohammad, R. M., & McCluskey, L. (2016). A dynamic self-structuring neural network model to combat phishing. *2016 international joint conference on neural networks (ijcnn)*, (s. 4221–4226).

Too, J., Abdullah, A. R., & Mohd Saad, N. (2019). A New Quadratic Binary Harris Hawk Optimization for Feature Selection. *Electronics, 8*. doi:10.3390/electronics8101130

Torres-Treviño, L. (2021). A 2020 taxonomy of algorithms inspired on living beings behavior. *arXiv preprint arXiv:2106.04775*.

Zhang, Y., Xiao, Y., Ghaboosi, K., Zhang, J., & Deng, H. (2012, April). A survey of cyber crimes. *Security and Communication Networks, 5*. doi:10.1002/sec.331