Journal of Theoretical and Applied Information Technology
30th June 2023. Vol.101. No 12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SWARM INTELLIGENCE ALGORITHMS FOR PHISHING WEBSITE DETECTION

SOHA ALHELALY
Saudi Electronic University, College of Computing and Informatics, Riyadh, Saudi Arabia 11673
E-mail: s.alhelaly@seu.edu.sa

ABSTRACT

Phishing websites are one of the biggest threats Internet users face today, and they require constantly updated techniques to combat the increasing number of such threats. So far, various methods have been proposed to increase the efficiency of phishing website detection. Swarm intelligence (SI) is one of the approaches that has garnered the interest of researchers working in the phishing website detection field. This article presents an up-to-date review of the SI techniques used for phishing website detection, which deserves wider investigation by researchers. Another contribution of this paper is to provide a comparison of the effectiveness of various SI-based phishing website detection techniques. Based on the survey result, we provide a clear overview of which approach is more suitable for each case and highlight the need for future research efforts devoted to the unique features of the SI for phishing website detection.

Keywords: *Phishing website detection, Swarm intelligence, Anti-phishing, Bat algorithm (BA), Particle swarm optimization (PSO), Artificial bee colony (ABC), Ant colony optimization (ACO), Gray wolf optimizer (GWO), Salp swarm algorithm (SSA), Firefly algorithm (FA),*

1. INTRODUCTION

The Internet has become so integral to daily life that its absence is unthinkable. Socializing, sharing knowledge, going shopping, and going to work are just a few of the many aspects of daily life that the Internet has revolutionized. Keeping data secure is a must for all computer networks. New forms of networks introduce new vulnerabilities and dangers, necessitating a wide range of recommended approaches to keeping computers secure against intrusion. Phishing is one type of attack, and it became a lot more common that year [1]. Social engineering combined with computer technologyis known as ”phishing” and it is used in network attacks to steal users’ private information. Using SMS, emails, or social media messages with misleading content, attackers try to trick users into clicking on fake links so that they can steal sensitive information (such as usernames, passwords, and credit card numbers) [2]. With the evolution of the Internet, phishing methods have evolved since they were first documented in a study in 1987 [3]. For instance, as the use of online payment systems grew, so did the number of phishing attempts targeting those systems. Phishing scams caused over 54 million USD in losses and accounted for nearly 30% of the cyber attack complaints received by the Internet Crime Complaint Center in 2020, as stated in the related Internet Crime Report [1]. Therefore, it is crucial for Internet users to be able to tell legitimate from fraudulent websites.

There are a number of methods used to thwart phishing scams [4][5]. Swarm intelligence (SI) approaches, which mimic animal swarm behavior for solving problems, have recently attracted a lot of attention for their potential usefulness in the detection of phishing websites. The strategies and tactics of swarms have been studied and modeled. Models including the bat algorithm (BA), particle swarm optimization (PSO), artificial bee colony (ABC), ant colony optimization (ACO), a gray wolf optimizer (GWO), the salp swarm algorithm (SSA), and the firefly algorithm (FA) were implemented for detecting phishing websites.

In light of the fact that SI has the potential to play a significant role in solving the phishing website problem, this paper offers a review of the work that has been proposed in the area of SI-based phishing websites. While previous publications have covered a variety of approaches, the focus here is on a single method—the detection of phishing websites using swarm intelligence. A further valuable contribution is the evaluation of different SI

phishing website detectors. This paper, therefore, makes an effort to define and organize the work done so far on the topic of SI-based phishing websites. To the best of my knowledge, this work is the first to provide a review paper of the SI techniques used for phishing website detection and provide a comparison of the effectiveness of various SI-based phishing website detection techniques. The rest of this article is organized as follows. In Section 2, we present the basics of phishing scams. Several similar works are discussed in Section 3. The SI methods are described in Section 4. In Section 5, we present SI methods for detecting phishing websites. Discussion, including an analysis of the effectiveness of these methods, is provided in Section 6. The paper is concluded in Section 7.

2. PHISHING ATTACK

Misleading Internet users in order to steal sensitive information is called ”phishing” [6][7]. Phishers are cyber criminals who engage in phishing assaults. In the mid-1990s, phishing became a major security concern when hackers began using it to gain login credentials [8][9]. The majority of phishing attacks are successful because of website spoofing [10][11] and email spoofing [12].

2.1 Phishing Attack Mechanism

The phishing attack mechanism was addressed by the authors of [13]. The first step is for an attacker to create a phishing website that is highly convincing in appearance. Attackers, however, spelled the URL incorrectly while trying to trick users. Even if the computer's browser may see the URL address, it is difficult for the non-expert user to identify it by sight and memory alone because it mimics authentic URLs. After that, they steal the page's content, including the design, logos, and text, by using programs to steal it from legal websites. Users are asked to enter personal details like login credentials and financial information into these phony websites. The second step is to force browsers to click the link by sending them an email. Links to malicious websites can be distributed using QR codes, short messages service (SMS), emails, voice messages, mobile applications, and social media [14]. After users click the link, they are taken to a phony website where the phishers can steal their information by tricking them into updating their accounts, making purchases, or resetting their passwords. Therefore, attackers obtain all information entered by users on the phony website. The next step is engaging in illegal activity, such as wire fraud or account takeover, using the victims' actual information.

2.2 Anti-Phishing Techniques

Before an attacker may successfully steal money from a user's account or use the information for other attacks, they must first complete a series of steps [13]. Therefore, a phishing assault can be stopped if the attack is stopped at a certain stage. This means that measures to combat phishing, such as web scraping, can be implemented at any time. This makes it harder for cyber criminals to use certain scripts to create crawlers that automatically obtain the content of legal web pages, intercept useful information, and copy it to phishing web pages. This is accomplished by employing obfuscation methods, displaying crucial data via sprites, and substituting text with graphics in order to thwart web scraping. Furthermore, using spam filters to identify unwanted emails prior to the user reading or clicking the link is another method used to combat phishing. Blacklists, whitelists, and empirical rules were the backbones of the first filters. In addition, some filters use machine-learning-based intelligent prediction models to detect spam that isn't on the predefined blacklist. Also, users are also unable to identify the phishing website by its URL alone; hence, several web browsers incorporate a security component to identify such sites. For phishing websites whose addresses are unknown, however, blacklist and whitelist-based methods are ineffective. Thankfully, new concepts and techniques for identifying phishing attempts have emerged thanks to the rapid growth of AI technology. Phishing links that aren't on the whitelist or blacklist can be detected using the machine learning-based predictive model. Furthermore, a second layer of authorization verification can be used to stop an attacker from exploiting stolen information to steal money, get access to a website, or otherwise misuse an account.

3. RELATED WORKS

Large-scale phishing detection methods have been presented in recent years, with some of the most successful implementations to date. Recent proposals are thought to be more cutting-edge. There are numerous literature reviews outlining and contrasting various methods for identifying phishing websites. The research team behind [15] conducted a survey of AI-based phishing detection methods. The authors studied statistical phishing records to determine the impact

and patterns of phishing attacks. Various phishing attack methods were compiled, along with the most common channels of communication and devices used as targets. The measures of the machine, deep, hybrid, and scenario-based learning are discussed in this paper as they pertain to anti- phishing efforts. In [15], the authors gave a brief history of phishing and notable phishing attack reports before reviewing machine learning-based phishing detection. Social engineering assaults and malware-based phishing are distinguished in the paper. They sorted rule-based characteristics into three distinct groups: features extracted from source code, features extracted from URLs, and features extracted from images. In [16], the authors presented a survey on major detection techniques and taxonomy for automatically detecting phishing solutions, categorizing them into web address-based methods, webpage content-based solutions, and hybrid approaches based on the input parameters. In each section, the authors detailed and provided an explanation of the most cutting-edge approaches. However, based on the methods used and the input parameters, the authors of [17] categorized phishing detection solutions into numerous groups. This research presents three distinct phishing methods and evaluates their relative accuracy. Moreover, in the comprehensive survey [18], Jain and Gupta examined phishing attack strategies, detection methods, and some current issues. The authors then presented and contrasted numerous countermeasures against phishing. Then, a number of significant obstacles, including choosing effective features, recognizing short URLs, and detecting smartphones, were provided.

3.1 Methodologies of Phishing Website Detection

As a social engineering problem, phishing attacks require solutions based on education, technical approaches, and legal oversight [14]. There are now three types of strategies for identifying phishing websites: list-based, heuristic, and machine- learning approaches [19]. Methods based on lists rely on systems to report and confirm either whitelists or blacklists. A whitelist is an approved list of websites or web addresses. The term ”blacklist” refers to a list of verified phishing domains. An automatically updated white list defense against phishing was proposed by the authors of [20]. The experimental results demonstrate that the method is effective, with an accuracy of 86.02% and a false-positive rate of less than 1.48%, and a response time that ensures real-time products and an environment. Also, phishing websites can be identified using heuristic algorithms, which compare a set of attributes retrieved from the textual content of the phony website with those of real websites. Machine learning approaches the subject of identifying potentially malicious websites as a classification problem using a set of features collected from the source code. More accurate performance and fewer false-positive rates are promised by machine learning approaches [14] for countering dynamic phishing assaults. Also, using identity keywords collected from the website’s textual content and a comparison of the domains of the legitimate and target websites is used by the authors of [21], which proposed a phishing detection method called PhishWHO. In [22], the authors checked the website’s legitimacy by looking at its logo. Thus, machine learning techniques are used to extract a logo from website photos, which is then used to detect phishing websites by creating a model to learn from a dataset containing structured features and making a prediction about whether or not the target website is real.

4. SWARM INTELLIGENCE (SI)

One source of motivation for humans in addressing difficult problems is the natural world. Recent years have seen the emergence of methods that draw inspiration from biology in fields as diverse as health, economics, engineering, the social sciences, and computer science. Many intrusion detection methods have also been presented that take biological inspiration. One such method is called swarm intelligence (SI). Swarm intelligence was originally used to describe a cellular robotics system in [23]. Many of the proposed algorithms and methodologies in these areas of study are based on swarm intelligence and cooperative problem-solving strategies of animals, insects, and birds. While a single person could never hope to undertake such a task; in reality, studies have shown that lone animals like insects, birds, and fish display very low levels of intelligence, and they are able to accomplish challenging tasks such as coordinating their movements and determining the quickest route to a food source when they interact socially with one another and their environment. Complex problems can often be solved with the help of computational intelligence methods like SI, such as some of NP-hard optimization issues such as the traveling salesman, routing, and scheduling. Similarly, SI methods have shown useful in phishing attack detection systems, either on their own as a self-

determining module or in conjunction with other predictive models.

5. SI APPROACHES IN PHISHING WEBSITE DETECTION

The swarm intelligence approaches that will be discussed in this section are particularly well-suited to the task of phishing website identification due to their ability to facilitate broad anomaly detection. In particular, swarm intelligence (SI) methods employ numerous agents to tackle difficult tasks. To find the best answer, every agent takes part and communicates with others in various ways. In this part, we take a close look at the various SI-based methods currently in use for detecting phishing websites. The primary method used to classify the systems shown here is the SI method, in which each section begins with a quick summary of the corresponding SI method that was used. We next demonstrate some application areas where SI is used, as well as some works that employ SI for phishing website identification.

5.1 Bat Algorithm (BA)

In 2010, [24] researchers created the meta-heuristic swarm method known as the Bat method (BA). The author was motivated to create this method by the echolocation mechanism employed by micro bats to help them locate prey and avoid obstacles during hunts. Bats use a method called echolocation to navigate their environments and find food when it's dark. This method is based on the echoes the bats make as they fly. In addition, echolocation is used to determine how far away potential sources of nourishment are. Bats increase the frequency of their calls and decrease the volume of the echo to better locate possible prey as they fly at different speeds and use different frequencies to communicate with one another as they search for food. The bats' position, speed, and frequency are adjusted, and its personal and world history are updated. Because all bats utilize echolocation to reach where they need to go, Yang based his algorithm on the rules of that biological system. The bats remember their current location and speed. The bats' position and speed are then modified according to the volume and variables [24][25] in order to determine the next search phase. The loud sound pulses that bats emit also assist them to estimate the distance between themselves and an obstruction. Multi-objective BA (MOBA) [26], Chaotic BA (CBA) [27], Differential operator and L'evy flights BA (DLBA) [28], K-means BA (KMBA) [29], Binary BA (BBA) [30], Fuzzy Logic BA (FLBA) [31], and Improved BA (IBA) [32] are just a few of the variants of the BA versions.

As a result, the BA is used for research in a wide range of disciplines. Discrete decision-making problems can be ad- dressed by the BA [25]. These BA arch mobility issues were proposed to be globally optimized using a chaotic method of the bat algorithm [33]. Hybridization of the BA has been investigated to boost the performance of the SVM algorithm [34], for example, by optimizing the SVM parameters for intrusion detection [35] with the help of the BA.

5.1.1 Bat Algorithm for Phishing Website Detection

An enhanced method called binary bat was proposed by the authors in [36]. The neural network is designed using the binary BA, and it uses this information to classify the URLs of the websites accessible via the network. The experimental results demonstrate that deep learning with the Adam optimizer achieves high classification accuracy at 94.8% when used with the SI approach for detecting phishing websites. Thus, to identify potential phishing sites, they suggested a deep learning model based on the SI-BBA algorithm. In [37], the authors developed an approach to tuning deep learning neural networks for phishing sites that makes use of a bat and hybrid-bat algorithm. When used to their detection, the new SI technique yields significant improvements over prior algorithms. The experimental results show promise, with increases in accuracy and performance when compared to existing classification algorithms for identifying phishing websites. Therefore, phishing website classification is an area where the suggested tuning deep learning utilizing a bat/hybrid-BA (TDLHBA/TDLHBA) approach excels.

5.2 Particle Swarm Optimization (PSO)

As a population-based computer technique, Particle Swarm Optimization (PSO) was initially created by Kennedy and Eberhart [38][39] to mimic the cooperative behavior of birds as they swarm to explore food. Using the PSO technique offers a number of benefits, including its ease of implementation (requiring fewer mathematical equations and parameters) [40][41]. In PSO, each potential solution is represented by a "particle" that travels across space at a specific speed within a "swarm." Then, each particle dynamically adjusts its position and velocity based on its own flying knowledge and the flying knowledge of its neighbors. Each

particle remembers its best position from the previous PSO iteration (pbest) and has access to the global best position that has been recorded. (gbest). Each particle, then, determines its own location and speed using pbest and gbest. Particles continually evaluate the viability of candidate solutions and remember the optimal environment. Each particle provides its neighbors with the best solution it has found so far, known as the particle best or local best [42]. Particles can then see where their neighbors have succeeded, guiding their own travels through the area; as a result, the population tends to converge at the end of an experiment. A particle's location is affected by both its own best location and the best location of its companion particles. The fitness function used to evaluate a particle's performance is different for each optimization issue.

5.2.1 Particle Swarm Optimization (PSO) for Phishing Web- site Detection

In [43], the authors offer a method for identifying fake online content. An effective model exists by combining the association and classification data techniques with the PSO algorithm for optimization. All the criteria and guidelines for categorizing the phishing website are characterized and identified using the proposed algorithms. An ACO algorithm was then used to optimize the outcomes of the categorization. However, this study has constraints, such as a lack of precision in estimating when the phishing categorization will converge and the possibility of sequences of random decisions. Because of this restriction, the authors turned to PSO to optimize a search space and predict social actions when they encounter phishing websites. When compared to other classification algorithms for detecting e-banking phishing websites, the associative classification algorithm using the PSO technique showed superior performance across two parameters (prediction accuracy and URL domain identity) and (security encryption). As such, PSO is regarded as a leading model in the field of network safety. In [44], it is suggested that feature weighting based on particle swarm optimization can improve phishing website identification. The proposed method proposes using PSO to weigh different website features efficiently to obtain better phishing website detection accuracy. Specifically, the suggested PSO-based website feature weighting distinguishes between the various website features based on their relevance in distinguishing phishing from legitimate websites. The experimental results showed that the proposed PSO- based feature weighting significantly improved the accuracy of machine learning models used to identify phishing websites, despite the fact that these models made use of fewer website features.

5.3 Artificial Bee Colony (ABC)

Inspired by the foraging habits of bees, Dervis Karaboga in [45] devised an optimization method called the Artificial Bee Colony (ABC) algorithm. Several academics have looked into potential uses for the ABC algorithm. There are three types of bees in a colony: workers, observers, and scouts. Using waggle motions, these bees coordinate their efforts to search for, select, and navigate food sources, reproduce, and alert other agents to the location of the food source. Workers leave the hive to forage for food, while observers utilize the data gathered by the workers to decide where to forage. When one of the worker bees eats all of the available food, it becomes a scout bee and goes in quest of more. The first step is to pick a food source from the available options. After then, the worker bees will forage randomly for new food sources that have more nectar than the ones they were given. If the identified alternative is better than the current option, it is adopted as the new preference. The worker bees then communicate this information to the observer bee by performing waggle dances. In this way, the observing bees select and evaluate potential food sources, accepting the new option if its fitness value is higher than the current one and rejecting it otherwise. Therefore, ABC looks into which of numerous options is best.

The ABC algorithm is a strong one that may be implemented in a variety of ways [46]. The technique is also flexible enough to be used with others to solve a wide range of optimization issues [47]. The ABC method has been effectively applied by numerous scholars to problems in many different disciplines of study, including but not limited to mathematics, computer science, engineering, decision sciences, biochemistry, genetics, physics, environmental science, medicine, and neurology [48]. Examples of areas where ABC has been put to use include data clustering [49], software test suite optimization [50][51], and image processing [52]. The ABC method is also used in intrusion prevention and detection systems (IDS) [53][54]. Using a classifier based on ABC for cloud computing is one example of how the method is put to use in an IDS in conjunction with a classification learning algorithm [53]. In addition, Random neural networks (RNNs) can be trained with the help of the ABC method, which is what the authors of [54]

did to create their ABC-based random neural network intrusion detection system (RNN-ABC).

5.3.1 Artificial Bee Colony (ABC) for Phishing Website Detection

In [55], the authors suggest a method for safeguarding against phishing attempts by identifying fake web addresses. The researchers employ an ABC approach to determine if the claimed website is genuine. If the URL is legitimate, then a big chunk of the issue goes away. Half of the job is done once it's clear that the link needs to be clicked on is malicious. According to the data gathered during the experiments, the typical accuracy of the system is 89%.

5.4 Ant Colony Optimization (ACO)

The field of Ant Colony Optimization (ACO) analyzes artificial systems that take their cues from the behavior of ant colonies. Discrete optimization issues [56] have been approached with the ants' exploratory mindset and their ability to locate the most direct route from their nests to a food source. Ants follow a regimented behavioral pattern, even though some ants have little or restricted vision. At first, ants wander aimlessly in search of nourishment. After finding food, ants bring it back to their colony and leave a pheromone along the path they took to get there. Therefore, ants may determine, based on the concentration of pheromones placed along each possible route, which route to choose. The route that has a higher concentration of pheromones will be chosen more frequently. Because ants choose the shortest route and return to their nests more quickly, the pheromone concentration is higher along the shorter route than along the longer route. The experiments in [57] demonstrate that the ants always take the shortest route.

This has led to several research drawing parallels between ant behavior and other disciplines. For instance, the authors of [58] were motivated to find solutions to least cost path problems by the mentality of an ant colony. Also, ACO has been used to solve the shortest path problem in numerous telecommunication networks [59]. ACO has been used in a number of contexts, most notably as an algorithm that finds relevant information and produces high-quality solutions to a wide range of optimization problems. As a result, ACO algorithms are developed for a wide variety of NP-hard problems, including but not limited to: the traveling salesman problem [60][61][62][63][64], the sequential ordering problem [65][66], the quadratic assignment problem (QAP) [67], the multiple knapsacks [68][69], the k-cardinality trees [70], constraint satisfaction [71][72], classification rules [73], Bayesian networks [74][75], the set covering problem [76][77], open shop scheduling [78][79], maximum clique [80][81], protein-ligand docking [82][83], protein folding [84][85], scheduling problems, such as total weighted tardiness [86][87][88], course timetabling [89][90], project scheduling [91][92], graph coloring [93][94]. The IDS industry also makes use of ACO algorithms. Thus, several methods have been developed to construct IDS models in conjunction with ACO algorithms [95][96][97] to safeguard the system from intrusions. The K- harmonic means clustering algorithm [98] and the fuzzy c- means [99] are two examples of algorithms that are frequently combined with ACO in the proposed research. Furthermore, in [100], it is explored how combining ACO with kernel principal component analysis (KPCA) might enhance the quality of clustering. The advantages of SVM classification and the ACO clustering efficiencies were integrated in [101]. Additionally, ACO is used in conjunction with SVM as a feature selection technique, as shown in [102][103]. In their research, the authors of [104] proposed an approach that combines a genetic algorithm (GA) for feature selection with a modified binary coded ACO algorithm (MBACO). IP traceback problems (IPTBK) involve determining where an attack originated over the Internet, and ant colony techniques are often utilized for this task [105]. Combining ACO with PCA, a method for anomaly detection termed digital signature is suggested in [106]. ACO is a resilient algorithm [107] because it may be tweaked to solve a wide variety of optimization issues. Also, the performance of the model could be enhanced by combining ACO with other algorithms [95].

5.4.1 Ant Colony Optimization (ACO) for Phishing Website Detection

The authors of [108] employed ACO in conjunction with a crude set-based feature importance algorithm to get a subset of attributes from the dataset of phishing websites. The random forest (RF) classifier achieved 97.26% accuracy in this study.

5.5 Grey Wolf Optimizer (GWO)

The Grey Wolf Optimizer (GWO) was proposed by [109]; it is an intelligent swarm algorithm that mimics the social structure and hunting techniques of grey wolves. The average

pack size for gray wolves is between five and twelve individuals. Alpha, beta, delta, and omega are the four levels of the dominating hierarchy among gray wolves. Alphas are in charge, and they come in both sexes. They run things, and they have all the power. The beta wolves are the pack's second- highest ranking members; they assist the alphas with tasks including making decisions. The only wolves the beta wolf can't command are the alphas. And if an alpha gets too old or dies, the beta wolves (of any sex) are the next in line to take over as leader. Delta wolves are third in the wolf pack hierarchy, after alpha and beta but above the omega pack. The GWO algorithm's search method mimics the three stages of gray wolf hunting: locating the prey, enclosing it, and making an assault. Exploration is prioritized throughout the searching and surrounding phases, but the offensive phase is focused on exploitation. The GWO algorithm's value lies in its ability to cut down on the total number of search parameters required by various programs. The alpha solution is considered the best option when simulating the social hierarchy of gray wolves and is followed by the beta and delta solutions. The remaining options are all omegas.

Binary GWO and a neural network classifier were used to identify critical features for a network intrusion detection system [110], just one example of the many ways in which GWO has been put to use by researchers. To further enhance IDS functionality, a modified binary GWO (MBGWO) is proposed for feature selection [111]. Low intrusion detection efficiency, which may be due to attacks' dynamic alterations, and a lack of an adequate training set are two issues that cloud GWO (CGWO) proposes to address in [112]. The approach also combines the efficiency of the K-means algorithm with the precision of the one-class support vector machine. In order to improve the anomaly-based IDS model, the authors of [113] implemented a multi-objective GWO algorithm. In [114], GWO and Black forest (BF) classifiers are used to create an efficient data-integrity-based IDS (DI-EIDS). The BF classifier is used to find the most useful samples for the sampling ratio optimization performed with GWO. This means they are used over and over to pick the best characteristics. To effectively classify and identify the various forms of intrusion attacks, the authors of [115] devised a framework that combines a hybrid GWO cuckoo search optimization (HGWCSO) for optimal feature selection with an enhanced transductive SVM (ETSVM). Combining GWO with the cuckoo search algorithm improves search speed, system stability, and security against infiltration. The Laplaci and GWO clustering methods and the support vector machine (SVM) classification approach are also discussed and utilized to identify potential invaders [116].

5.5.1 Grey Wolf Optimization (GWO) for Phishing Website Detection

In [117], the authors attempt to apply a machine learning model to distinguish between phishing and authentic websites by analyzing several features of their URLs. The length of the IP address, the authenticity of the HTTPs request, the presence of pop-up windows for data entry, and the state of the server form handler are all examples of such characteristics. The legitimacy of a website was predicted using a support vector machine (SVM) binary classifier trained on an existing dataset. This was accomplished by locating the best hyperplane to divide the classes. Four optimization techniques are used to locate this optimal hyperplane: the Bat Algorithm (BA), the Firefly Algorithm (FA), the Grey Wolf Optimization (GWO), and the Whale Optimization Algorithm (WOA). The GWO method outperforms the FA among the four nature-inspired optimization techniques tested. This method has various real- world applications, such as being used in antivirus software and browser extensions to help consumers determine whether or not a website is trustworthy. By scanning websites and feeding the inputs into the model, this model might also be implemented as a classifier in real-time to safeguard users from phishing. In addition to improving spam filters, this technique might be used to check websites for potential security flaws. In this way, the model can aid online hosting firms in their search for accurate classifiers to identify phishing websites.

5.6 Salp Swarm Algorithm (SSA)

The Salp Swarm Algorithm (SSA) is a recently developed optimization method [118] that has the potential to address a wide range of optimization challenges. It mimics the actions of salps, which are barrel-shaped planktonic tunicates belonging to the family Salpidae in the animal kingdom. They move and have tissues like jellyfish, and a large proportion of their mass consists of water [119]. They change their postures by pushing water through their jelli-shaped bodies, which allows them to move around [120]. The salp chain is a swarm behavior used by salps in the ocean that aids in foraging and improves the efficiency with which the salps may move

[121][122]. In light of this behavior, the authors of [118] developed a mathematical model of salp chains and evaluated it in a number of optimization scenarios. At the outset of SSA, a population is divided into two parts: a leader, represented by the salp at the chain's top, and their followers. The location of the salps is calculated in the n-dimensional problem space represented by these archs. The salps' foraging behavior suggests that whatever they are attacking is a food source. The new location is then regularly updated. The SSA has been praised for its simplicity, power, adaptability, and ease of use in both serial and parallel configurations.

For optimization problems with many objectives, SSA has been a common choice as an algorithm [123][124][125][126]. To handle feature selection (FS) tasks, Faris et al. provided one of the most important papers on SSA in [124]. The best internal conductor was identified using simple SSA by the authors of [127] for the actual radial distribution system in use in Egypt. Gain and parameter optimization for a fractional- order proportional integral derivative controller using SSA is presented in [128]. However, authors in [129] proposed using SSA to optimize the size of a CMOS differential amplifier and the comparator circuit, which is another application of SSA in electrical engineering. Other examples of engineering challenges include designing a PID-fuzzy control for a seismic exited structural system [130], extracting the parameters of polarization curves of polymer exchange membrane fuel cells model [131], and optimizing load frequency control using SSA in managing the active power of an isolated renewable microgrid [132]. The authors of [133] applied SSA to optimize the least squares hyperparameters in an environmental application predicting emissions from Energy using Support Vector Machines (SVM).

5.6.1 Salp Swarm Algorithm (SSA) for Phishing Website Detection

A new phishing detection method based on the SSA was proposed in [134]. In the wrapper-based feature selection framework, the SSA algorithm is used as a search algorithm. The primary goal is to reduce the number of features used by the phishing system while increasing its classification performance. Three cutting-edge algorithms are used to compare the phishing system. In terms of user evaluation metrics, the results suggest that Binary SSA performed the best.

5.7 Firefly Algorithm (FA)

In late 2007 and early 2008, researchers at Cambridge University [135][136] drew inspiration from the flashing behavior of fireflies to create the Firefly Algorithm (FA). The FA basically employs several rules, such as the fact that fireflies are unisex and all of them attracted one other, and the fact that the FA controls the brightness of the firefly so that the FA's less light-flashing fireflies will migrate toward the FA's brighter ones. If there isn't even one really brilliant firefly, though, the behavior will be unpredictable at best [137].

Several researchers are interested in the FA, and it has found numerous uses. The authors of [138][139] showed that, when it came to compressing digital images, the firefly algorithm required the least amount of processing power. The firefly algorithm (FA) was utilized for feature selection by Banati and Bajaj. They demonstrated the FA's superior performance consistency over alternative algorithms [140]. Nonlinear and multimodal design issues were shown to be amenable to the FA in [141]. The FA was used by the authors of [142] to improve antenna design in [143]. In addition, extensive research has shown that FA is effective for a wide variety of test problems, such as multi-objective load dispatch issues [144] and traveling salesman problem (TSP) and scheduling problems [145]. FAs also show great performance in the fields of classification and clustering [146]. For instance, Senthilnathet al. presented a comprehensive performance analysis that compared FA to other algorithms and found that it performed well in clustering [146]. The FA typically achieves the best results compared to alternate algorithms. There is additional evidence that FAs can be used to teach NNs [147]. FAs can be very useful for optimization in uncertain settings, as shown in [148]. Discrete versions of the FA have been created with excel- lent performance [149], and they can be used for traveling- salesman problems, graph coloring, and other applications. These variants of the FA are useful for discrete problems and combinatorial optimization. The FA's potential for use in multi-objective optimization is also explored in [150]. Both chaos and hybridization of the FA with other algorithms can boost its efficiency [151][139].

5.7.1 Firefly Algorithm (FA) for Phishing Website Detection

In order to solve the parameter setting problem for a deep neural network, the authors of [152] introduced FA for phishing website detection. Three SI algorithms were employed in this study to the problem of phishing website classification: the bat algorithm, the hybrid-bat algorithm, and the firefly algorithm. In order to differentiate phishing from authentic websites, the authors used the presented approach versions to a classification problem. Experiments comparing the suggested method's performance to that of four different phishing datasets yielded encouraging results, demonstrating the method's potential. The suggested swarm intelligence- based solution vastly outperformed the manually configured deep neural network in terms of prediction performance. When testing various methods for identifying bogus websites, the proposed firefly approach performed the best.

Table 1: Performance Comparison of Several Swarm Intelligence-Based Phishing Website Detection Methods

Algorithm Used	Reference Paper	Number of Features	Data Set	Accuracy
Feature selection using Ant Colony Optimization (ACO)	[108]	23	Rough sets	97.259%
Deep learning based using Binary Bat Algorithm (BBA)	[36]	30	Kaggle	94.8%
Deep neural networks using Firefly Algorithm (FA)	[152]	30 9 111	by Mohammad by Abdelhamid by Vrbancic	96.65% 86.06% 94.39%
Tuning deep learning using Bat/Hybrid Bat Algorithm (TDLBA/TDLHBA)	[37]	31	UCI Machine Learning Repository	97%
Support vector machines and Gray Wolf Optimizer(GWO)	[117]	9	UCI Machine Learning Repository	90.38%
Associative classification algorithm with Particle Swarm Optimization (PSO)	[43]	27	Phishtank	91%
Particle Swarm Optimization trained Classification AssociationRule Mining (PSOCARM)	[153]	17	PhishTank	Email: 83% URL: 88%
Feature weighting using Particle Swarm Optimization(PSO)	[44]	30	UCI Machine Learning Repository	95.88%
Feature selection using Salp Swarm Optimization(SSO)	[134]	57	Phishtank	95%
Artificial Bee Colony Algorithm (ABC)	[55]	12	Created by the authors	89%

6. DISCUSSION

This study focused on the SI-based algorithms available for detecting phishing websites, which is different than other studies that presented the methods for phishing detection in general. In Section 4, we discussed the various SI methods that have been used to identify phishing websites. It is possible to draw some conclusions about the efficacy of such an algorithm, as shown in the table. Table 1 presents a comparison of several of the methods discussed thus far. As a result of this study, we need more future efforts dedicated to studying the techniques for detecting phishing websites using SI algorithms. As the complexity of phishing attempts grows daily, new detection techniques must be developed using the unique features dedicated to swarm intelligence algorithms.

7. CONCLUSION

Swarm intelligence (SI) refers to a set of techniques that draws inspiration from the collective intelligence displayed by swarms such as swarms of insects. Researchers were interested in phishing website identification once SI was successfully applied in other fields. In this survey, an analysis of the methods proposed for detecting phishing websites has been performed. However, the paper focused on SI-based algorithms available for detecting phishing websites. Furthermore, the SI is used to categorize the works and show how well the corresponding algorithm performs. The survey presented the algorithms along with highlighting their efficacies. The result of the study is identified that the SI-based phishing website detection domain is limited which still needs to be explored. As the sophistication of phishing attempts grows daily, new methods of detection must be developed. Thus, this paper can be considered a reference for researchers working in the domain of phishing website detection and swarm intelligence.

REFERENCES:

[1] Fbi: Internet crime report 2020," *Computer Fraud Security*, vol. 2021, no. 4, p. 4, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1361372321000385

[2] A. Aleroud and L. Zhou, "Phishing environments, techniques, and countermeasures: A survey," *Computers & Security*, vol. 68, pp. 160– 196, 2017.

[3] F. Jerry and H. Chris, "System security: A hacker's perspective," in *Proceedings of the 1987 North American conference of Hewlett-Packard business computer users, Las Vegas, NV, USA*, 1987, pp. 20–25.

[4] J. Kumar, A. Santhanavijayan, B. Janet, B. Rajendran, and B. Bindhu- madhava, "Phishing website classification and detection using machine learning," in *2020 International Conference on Computer Communi- cation and Informatics (ICCCI)*, 2020, pp. 1–6.

[5] G. Varshney, M. Misra, and P. K. Atrey, "A survey and classification of web phishing detection schemes," *Security and Communication Networks*, vol. 9, no. 18, pp. 6266–6284, 2016.

[6] E. E. Lastdrager, "Achieving a consensual definition of phishing based on a systematic review of the literature," *Crime Science*, vol. 3, no. 1, pp. 1–10, 2014.

[7] R. M. Mohammad, F. Thabtah, and L. McCluskey, "Tutorial and critical analysis of phishing websites methods," *Computer Science Review*, vol. 17, pp. 1–24, 2015.

[8] S. Garera, N. Provos, M. Chew, and A. D. Rubin, "A framework for detection and measurement of phishing attacks," in *Proceedings of the 2007 ACM workshop on Recurring malcode*, 2007, pp. 1–8.

[9] M. Khonji, Y. Iraqi, and A. Jones, "Phishing detection: a literature survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 2091–2121, 2013.

[10] G. Varshney, A. Sardana, and R. C. Joshi, "Secret information display based authentication technique towards preventing phishing attacks," in *Proceedings of the International Conference on Advances in Com- puting, Communications and Informatics*, 2012, pp. 602–608.

[11] J. Hong, "The state of phishing attacks," *Communications of the ACM*, vol. 55, no. 1, pp. 74–81, 2012.

[12] K. Pandove, A. Jindal, and R. Kumar, "Email spoofing," *International Journal of Computer Applications*, vol. 5, no. 1, pp. 27–30, 2010.

[13] L. Tang and Q. H. Mahmoud, "A survey of machine learning-based solutions for phishing website detection," *Machine Learning and Knowledge Extraction*, vol. 3, no. 3, pp. 672–694, 2021.

[14] Y. A. Alsariera, V. E. Adeyemo, A. O. Balogun, and A. K. Alaz- zawi, "Ai meta-learners and extra-trees algorithm for the

detection of phishing websites," *IEEE Access*, vol. 8, pp. 142 532–142 542, 2020.

[15] A. Basit, M. Zafar, X. Liu, A. R. Javed, Z. Jalil, and K. Kifayat, "A comprehensive survey of ai-enabled phishing attacks detection techniques," *Telecommunication Systems*, vol. 76, no. 1, pp. 139–154, 2021.

[16] M. Vijayalakshmi, S. Mercy Shalinie, M. H. Yang, and R. M. U,"Web phishing detection techniques: a survey on the state-of-the-art, taxonomy and future directions," *Iet Networks*, vol. 9, no. 5, pp. 235– 246, 2020.

[17] P. Kalaharsha and B. M. Mehtre, "Detecting phishing sites–an overview," *arXiv preprint arXiv:2103.12739*, 2021.

[18] A. K. Jain and B. Gupta, "A survey of phishing attack techniques, defence mechanisms and open research challenges," *Enterprise Information Systems*, vol. 16, no. 4, pp. 527–565, 2022.

[19] M. Zabihimayvan and D. Doran, "Fuzzy rough set feature selection to enhance phishing attack detection," in *2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 2019, pp. 1–6.

[20] A. K. Jain and B. B. Gupta, "A novel approach to protect against phishing attacks at client side using auto-updated white-list," *EURASIP Journal on Information Security*, vol. 2016, no. 1, pp. 1–11, 2016.

[21] C. L. Tan, K. L. Chiew, K. Wong *et al.*, "Phishwho: Phishing webpage detection via identity keywords extraction and target domain name finder," *Decision Support Systems*, vol. 88, pp. 18–27, 2016.

[22] K. L. Chiew, E. H. Chang, W. K. Tiong *et al.*, "Utilisation of website logo for phishing detection," *Computers & Security*, vol. 54, pp. 16– 26, 2015.

[23] G. Beni and J. Wang, "Swarm intelligence in cellular robotic systems," in *Robots and biological systems: towards a new bionics?* Springer, 1993, pp. 703–712.

[24] X.-S. Yang, "A new metaheuristic bat-inspired algorithm," *Nature inspired cooperative strategies for optimization (NICSO 2010)*, pp. 65–74, 2010.

[25] A.-C. Enache and V. Sgarciu, "An improved bat algorithm driven by support vector machines for intrusion detection," in *International Joint Conference: CISIS'15 and ICEUTE'15*. Springer, 2015, pp. 41–51.

[26] X.-S. Yang, "Bat algorithm for multi-objective optimisation," *Interna- tional Journal of Bio-Inspired Computation*, vol. 3, no. 5, pp. 267–274,2011.

[27] J.-H. Lin, C.-W. Chou, C.-H. Yang, H.-L. Tsai *et al.*, "A chaotic levy flight bat algorithm for parameter estimation in nonlinear dynamic biological systems," *CIT*, 2012.

[28] J. Xie, Y. Zhou, and H. Chen, "A novel bat algorithm based on differ- ential operator and lévy flights trajectory," *Computational intelligence and neuroscience*, vol. 2013, 2013.

[29] G. Komarasamy and A. Wahi, "An optimized k-means clustering tech- nique using bat algorithm," *European Journal of Scientific Research*, vol. 84, no. 2, pp. 263–273, 2012.

[30] R. Y. M. Nakamura, L. A. M. Pereira, D. Rodrigues, K. A. P. Costa,
J. P. Papa, and X.-S. Yang, "Binary bat algorithm for feature selection," in *Swarm intelligence and bio-inspired computation*. Elsevier, 2013, pp. 225–237.

[31] K. Khan, A. Nikov, and A. Sahai, "A fuzzy bat clustering method for ergonomic screening of office workplaces," in *Third international conference on software, services and semantic technologies S3T 2011*. Springer, 2011, pp. 59–66.

[32] M. Jamil, H. Zepernic, and X. Yang, "Improved bat algorithm for global optimization," *Applied Soft Computing*, 2013.

[33] A. H. Gandomi and X.-S. Yang, "Chaotic bat algorithm," *Journal of computational science*, vol. 5, no. 2, pp. 224–232, 2014.

[34] M. A. Laamari and N. Kamel, "A hybrid bat based feature selection approach for intrusion detection," in *Bio-Inspired Computing-Theories and Applications: 9th International Conference, BIC-TA 2014, Wuhan, China, October 16-19, 2014. Proceedings*. Springer, 2014, pp. 230– 238.

[35] A.-C. Enache and V. Sgârciu, "Anomaly intrusions detection based on support vector machines with an improved bat algorithm," in *2015 20th international conference on control systems and computer science*. IEEE, 2015, pp. 317–321.

[36] P. P. Kumar, T. Jaya, and V. Rajendran, "Si-bba–a novel phishing website detection based on swarm intelligence with deep learning," *Materials Today: Proceedings*, 2021.

[37] G. Vrbančič, I. Fister Jr, and V. Podgorelec, "Swarm intelligence approaches

for parameter setting of deep learning neural network: case study on phishing websites classification," in *Proceedings of the 8th international conference on web intelligence, mining and semantics*, 2018, pp. 1–8.

[38] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, vol. 4, 1995, pp. 1942–1948 vol.4.

[39] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *MHS'95. Proceedings of the sixth international symposium on micro machine and human science*. Ieee, 1995, pp. 39–43.

[40] A. Kawamura and B. Chakraborty, "A hybrid approach for optimal feature subset selection with evolutionary algorithms," in *2017 IEEE 8th International Conference on Awareness Science and Technology (iCAST)*, 2017, pp. 564–568.

[41] J. Wei, Z. Jian-qi, and Z. Xiang, "Face recognition method based on support vector machine and particle swarm optimization," *Expert Systems with Applications*, vol. 38, no. 4, pp. 4390–4393, 2011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417410010651

[42] J. F. Schutte, "The particle swarm optimization algorithm," *Structural Optimization*, 2005.

[43] M. Radha Damodaram and M. Valarmathi, "Phishing website detection and optimization using particle swarm optimization technique," *Inter- national Journal of Computer Science and Security (IJCSS)*, vol. 5, no. 5, p. 477, 2011.

[44] W. Ali and S. Malebary, "Particle swarm optimization-based feature weighting for improving intelligent phishing website detection," *IEEE Access*, vol. 8, pp. 116 766–116 780, 2020.

[45] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," Technical report-tr06, Erciyes university, engineering faculty, computer . . . , Tech. Rep., 2005.

[46] B. Akay and D. Karaboga, "A modified artificial bee colony algorithm for real-parameter optimization," *Information sciences*, vol. 192, pp. 120–142, 2012.

[47] W. Gao, S. Liu, and L. Huang, "A global best artificial bee colony algorithm for global optimization," *Journal of Computational and Applied Mathematics*, vol. 236, no. 11, pp. 2741–2753, 2012.

[48] J. C. Bansal, H. Sharma, and S. S. Jadon, "Artificial bee colony algorithm: a survey," *International Journal of Advanced Intelligence Paradigms*, vol. 5, no. 1-2, pp. 123–159, 2013.

[49] X. Lei, X. Huang, and A. Zhang, "Improved artificial bee colony algorithm and its application in data clustering," in *2010 IEEE fifth international conference on bio-inspired computing: theories and applications (BIC-TA)*. IEEE, 2010, pp. 514–521.

[50] D. J. Mala, V. Mohan, and M. Kamalapriya, "Automated software test optimisation framework–an artificial bee colony optimisation-based approach," *IET software*, vol. 4, no. 5, pp. 334–348, 2010.

[51] S. S. Dahiya, J. K. Chhabra, and S. Kumar, "Application of artificial bee colony algorithm to software testing," in *2010 21st Australian software engineering conference*. IEEE, 2010, pp. 149–154.

[52] C. Chidambaram and H. S. Lopes, "A new approach for template matching in digital images using an artificial bee colony algorithm," in *2009 World Congress on Nature & Biologically Inspired Computing (NaBIC)*. IEEE, 2009, pp. 146–151.

[53] S. Kalaivani, A. Vikram, and G. Gopinath, "An effective swarm optimization based intrusion detection classifier system for cloud com-puting," in *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*. IEEE, 2019, pp. 185–188.

[54] H. Larijani, A. Javed, N. Mtetwa, J. Ahmad *et al.*, "Intrusion detection using swarm intelligence," in *2019 UK/China Emerging Technologies (UCET)*. IEEE, 2019, pp. 1–5.

[55] T. Bhardwaj, T. K. Sharma, and M. R. Pandit, "Social engineering prevention by detecting malicious urls using artificial bee colony algorithm," in *Proceedings of the Third International Conference on Soft Computing for Problem Solving: SocProS 2013, Volume 1*. Springer, 2014, pp. 355–363.

[56] M. D. L. M. Gambardella, M. B. A. Martinoli, and R. P. T. Stützle, "Ant colony optimization and swarm intelligence," 2004.

[57] S. Goss, S. Aron, J.-L. Deneubourg, and J. M. Pasteels, "Self-organized shortcuts in the argentine ant," *Naturwissenschaften*, vol. 76, no. 12, pp. 579–581, 1989.

[58] M. Dorigo, M. Birattari, and T. Stutzle, "Ant colony optimization," *IEEE computational intelligence magazine*, vol. 1, no. 4, pp. 28–39, 2006.

[59] R. Schoonderwoerd, O. E. Holland, J. L. Bruten, and L. J. Rothkrantz, "Ant-based load balancing in telecommunications networks," *Adaptive behavior*, vol. 5, no. 2, pp. 169–207, 1997.

[60] M. Dorigo, V. Maniezzo, and A. Colorni, "Ant system: optimization by a colony of cooperating agents," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 26, no. 1, pp. 29–41, 1996.

[61] T. Stutzle and H. Hoos, "Max-min ant system and local search for the traveling salesman problem," in *Proceedings of 1997 IEEE international conference on evolutionary computation (ICEC'97)*. IEEE, 1997, pp. 309–314.

[62] A. F. Tuani, E. Keedwell, and M. Collett, "Heterogenous adaptive ant colony optimization with 3-opt local search for the travelling salesman problem," *Applied Soft Computing*, p. 106720, 2020.

[63] X. Yang and J.-s. Wang, "Application of improved ant colony opti- mization algorithm on traveling salesman problem," in *2016 Chinese Control and Decision Conference (CCDC)*. IEEE, 2016, pp. 2156– 2160.

[64] K. Yang, X. You, S. Liu, and H. Pan, "A novel ant colony optimization based on game for traveling salesman problem," *Applied Intelligence*, pp. 1–14, 2020.

[65] R. Skinderowicz, "An improved ant colony system for the sequential ordering problem," *Computers & Operations Research*, vol. 86, pp. 1–17, 2017.

[66] L. M. Gambardella, R. Montemanni, and D. Weyland, "An enhanced ant colony system for the sequential ordering problem," in *Operations Research Proceedings 2011*. Springer, 2012, pp. 355–360.

[67] S. Oliveira, M. S. Hussin, A. Roli, M. Dorigo, and T. Stützle, "Analysis of the population-based ant colony optimization algorithm for the tsp and the qap," in *2017 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2017, pp. 1734–1741.

[68] I. B. Mansour, I. Alaya, and M. Tagina, "A gradual weight-based ant colony approach for solving the multiobjective multidimensional knapsack problem," *Evolutionary Intelligence*, vol. 12, no. 2, pp. 253–272, 2019.

[69] W. Zouari, I. Alaya, and M. Tagina, "A hybrid ant colony algorithm with a local search for the strongly correlated knapsack problem," in *2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)*. IEEE, 2017, pp. 527–533.

[70] C. Blum and M. J. Blesa, "New metaheuristic approaches for the edge-weighted k-cardinality tree problem," *Computers & Operations Research*, vol. 32, no. 6, pp. 1355–1377, 2005.

[71] C. Solnon, "Ants can solve constraint satisfaction problems," *IEEE transactions on evolutionary computation*, vol. 6, no. 4, pp. 347–357, 2002.

[72] T. Toya, K. Mizuno, and T. Masukane, "Dynamic adjustment of control parameters in aco for constraint satisfaction problems," in *2020 3rd International Conference on Intelligent Autonomous Systems (ICoIAS)*. IEEE, 2020, pp. 128–132.

[73] R. S. Parpinelli, H. S. Lopes, and A. A. Freitas, "Data mining with an ant colony optimization algorithm," *IEEE transactions on evolutionary computation*, vol. 6, no. 4, pp. 321–332, 2002.

[74] L. M. De Campos, J. M. Fernandez-Luna, J. A. Gámez, and J. M. Puerta, "Ant colony optimization for learning bayesian networks," *International Journal of Approximate Reasoning*, vol. 31, no. 3, pp. 291–311, 2002.

[75] J. Jun-Zhong, H.-X. ZHANG, H. Ren-Bing, and L. Chun-Nian, "A bayesian network learning algorithm based on independence test and ant colony optimization," *Acta Automatica Sinica*, vol. 35, no. 3, pp. 281–288, 2009.

[76] L. Lessing, I. Dumitrescu, and T. Stützle, "A comparison between aco algorithms for the set covering problem," in *International Workshop on Ant Colony Optimization and Swarm Intelligence*. Springer, 2004, pp. 1–12.

[77] Z.-G. Ren, Z.-R. Feng, L.-J. Ke, and Z.-J. Zhang, "New ideas for applying ant colony optimization to the set covering problem," *Computers & Industrial Engineering*, vol. 58, no. 4, pp. 774–784, 2010.

[78] C. Blum, "Beam-aco—hybridizing ant colony optimization with beam search: An application to open shop scheduling," *Computers & Oper- ations Research*, vol. 32, no. 6, pp. 1565–1591, 2005.

[79] G. Campos Ciro, F. Dugardin, F. Yalaoui,

and R. Kelly, "Open shop scheduling problem with a multi-skills resource constraint: a genetic algorithm and an ant colony optimisation approach," *International Journal of Production Research*, vol. 54, no. 16, pp. 4854–4881, 2016.

[80] S. Fenet and C. Solnon, "Searching for maximum cliques with ant colony optimization," in *Workshops on Applications of Evolutionary Computation*. Springer, 2003, pp. 236–245.

[81] D. El Baz, M. Hifi, L. Wu, and X. Shi, "A parallel ant colony optimization for the maximum-weight clique problem," in *2016 IEEE International Parallel and Distributed Processing Symposium Work- shops (IPDPSW)*. IEEE, 2016, pp. 796–800.

[82] O. Korb, T. Stützle, and T. E. Exner, "Plants: Application of ant colony optimization to structure-based drug design," in *International work- shop on ant colony optimization and swarm intelligence*. Springer, 2006, pp. 247–258.

[83] ——, "An ant colony optimization approach to flexible protein–ligand docking," *Swarm Intelligence*, vol. 1, no. 2, pp. 115–134, 2007.

[84] A. Shmygelska and H. H. Hoos, "An ant colony optimisation algorithm for the 2d and 3d hydrophobic polar protein folding problem," *BMC bioinformatics*, vol. 6, no. 1, p. 30, 2005.

[85] A. Llanes, C. Vélez, A. M. Sánchez, H. Pérez-Sánchez, and J. M. Cecilia, "Parallel ant colony optimization for the hp protein folding problem," in *International Conference on Bioinformatics and Biomed- ical Engineering*. Springer, 2016, pp. 615–626.

[86] M. Den Besten, T. Stützle, and M. Dorigo, "Ant colony optimization for the total weighted tardiness problem," in *International Conference on Parallel Problem Solving from Nature*. Springer, 2000, pp. 611– 620.

[87] L. Li, F. Qiao, and Q. Wu, "Aco-based scheduling of parallel batch processing machines with incompatible job families to minimize total weighted tardiness," in *International Conference on Ant Colony Optimization and Swarm Intelligence*. Springer, 2008, pp. 219–226.

[88] D. Merkle and M. Middendorf, "Ant colony optimization with global pheromone evaluation for scheduling a single machine," *Applied Intelligence*, vol. 18, no. 1, pp. 105–111, 2003.

[89] K. Socha, M. Sampels, and M. Manfrin, "Ant algorithms for the university course timetabling problem with regard to the state-of- the-art," in *Workshops on Applications of Evolutionary Computation*. Springer, 2003, pp. 334–345.

[90] V. D. Matijaš, G. Molnar, M. Č upić, D. Jakobović, and B. D. Bašić, "University course timetabling using aco: a case study on laboratory exercises," in *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*. Springer, 2010, pp. 100–110.

[91] D. Merkle, M. Middendorf, and H. Schmeck, "Ant colony optimization for resource-constrained project scheduling," *IEEE transactions on evolutionary computation*, vol. 6, no. 4, pp. 333–346, 2002.

[92] D. Thiruvady, C. Blum, and A. T. Ernst, "Maximising the net present value of project schedules using cmsa and parallel aco," in *Interna- tional Workshop on Hybrid Metaheuristics*. Springer, 2019, pp. 16–30.

[93] D. Costa and A. Hertz, "Ants can colour graphs," *Journal of the operational research society*, vol. 48, no. 3, pp. 295–305, 1997.

[94] L. Lv, C. Gao, J. Chen, L. Luo, and Z. Zhang, "Physarum-based ant colony optimization for graph coloring problem," in *International Conference on Swarm Intelligence*. Springer, 2019, pp. 210–219.

[95] L. P. Rajeswari, A. Kannan, and R. Baskaran, "An escalated approach to ant colony clustering algorithm for intrusion detection system," in *International Conference on Distributed Computing and Networking*. Springer, 2008, pp. 393–400.

[96] N. Sreelaja and G. V. Pai, "Ant colony optimization based approach for efficient packet filtering in firewall," *Applied Soft Computing*, vol. 10, no. 4, pp. 1222–1236, 2010.

[97] Y. Li, J. Xia, S. Zhang, J. Yan, X. Ai, and K. Dai, "An efficient intrusion detection system based on support vector machines and gradually feature removal method," *Expert Systems with Applications*, vol. 39, no. 1, pp. 424–430, 2012.

[98] H. Jiang, S. Yi, J. Li, F. Yang, and X. Hu, "Ant clustering algorithm with k-harmonic means clustering," *Expert Systems with Applications*, vol. 37, no. 12, pp. 8679–8684, 2010.

[99] P. M. Kanade and L. O. Hall, "Fuzzy ants and clustering," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 37, no. 5, pp. 758–769, 2007.

[100] L. Zhang and Q. Cao, "A novel ant-based clustering algorithm using the kernel method," *Information Sciences*, vol. 181, no. 20, pp. 4658– 4672, 2011.

[101] W. Feng, Q. Zhang, G. Hu, and J. X. Huang, "Mining network data for intrusion detection through combining svms with ant colony networks," *Future Generation Computer Systems*, vol. 37, pp. 127– 140, 2014.

[102] T. Mehmood and H. B. M. Rais, "Svm for network anomaly detection using aco feature subset," in *2015 International symposium on mathe- matical sciences and computing research (iSMSC)*. IEEE, 2015, pp. 121–126.

[103] P. R. K. Varma, V. V. Kumari, and S. S. Kumar, "Feature selection using relative fuzzy entropy and ant colony optimization applied to real-time intrusion detection system," *Procedia computer science*, vol. 85, pp. 503–510, 2016.

[104] Y. Wan, M. Wang, Z. Ye, and X. Lai, "A feature selection method based on modified binary coded ant colony optimization algorithm," *Applied Soft Computing*, vol. 49, pp. 248–258, 2016.

[105] P. Wang, H.-T. Lin, and T.-S. Wang, "An improved ant colony system algorithm for solving the ip traceback problem," *Information Sciences*,vol. 326, pp. 172–187, 2016.

[106] G. Fernandes Jr, L. F. Carvalho, J. J. Rodrigues, and M. L. Proença Jr, "Network anomaly detection using ip flows with principal component analysis and ant colony optimization," *Journal of Network and Com-puter Applications*, vol. 64, pp. 1–11, 2016.

[107] M. A. J. Ghasab, S. Khamis, F. Mohammad, and H. J. Fariman, "Feature decision-making ant colony optimization system for an auto- mated recognition of plant species," *Expert Systems with Applications*, vol. 42, no. 5, pp. 2361–2370, 2015.

[108] R. K. V. Penmatsa and P. Kakarlapudi, "Web phishing detection: feature selection using rough sets and ant colony optimisation," *International Journal of Intelligent Systems Design and Computing*, vol. 2, no. 2, pp. 102–113, 2018.

[109] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer,"
Advances in engineering software, vol. 69, pp. 46–61, 2014.

[110] J. K. Seth and S. Chandra, "Intrusion detection based on key feature selection using binary gwo," in *2016 3rd international conference on computing for sustainable global development (INDIACom)*. IEEE, 2016, pp. 3735–3740.

[111] Q. M. Alzubi, M. Anbar, Z. N. Alqattan, M. A. Al-Betar, and
R. Abdullah, "Intrusion detection system based on a modified binary grey wolf optimisation," *Neural Computing and Applications*, pp. 1–13, 2019.

[112] H. Yang and Z. Zhou, "A novel intrusion detection scheme using cloud grey wolf optimizer," in *2018 37th Chinese Control Conference (CCC)*. IEEE, 2018, pp. 8297–8302.

[113] T. A. Alamiedy, M. Anbar, Z. N. Alqattan, and Q. M. Alzubi, "Anomaly-based intrusion detection system using multi-objective grey wolf optimisation algorithm," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–22, 2019.

[114] R. Benisha and S. R. Ratna, "Detection of data integrity attacks by constructing an effective intrusion detection system," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–12, 2020.

[115] E. Roopa Devi and R. Suganthe, "Enhanced transductive support vector machine classification with grey wolf optimizer cuckoo search optimization for intrusion detection system," *Concurrency and Com- putation: Practice and Experience*, vol. 32, no. 4, p. e4999, 2020.

[116] P. Anitha and B. Kaarthick, "Oppositional based laplacian grey wolf optimization algorithm with svm for data mining in intrusion detectionsystem," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–12, 2019.

[117] S. Anupam and A. K. Kar, "Phishing website detection using sup- port vector machines and nature-inspired optimization algorithms," *Telecommunication Systems*, vol. 76, no. 1, pp. 17–32, 2021.

[118] S. Mirjalili, A. H. Gandomi, S. Z. Mirjalili, S. Saremi, H. Faris, and S. M. Mirjalili, "Salp swarm algorithm: A bio-inspired optimizer for engineering design problems," Advances in engineering software, vol. 114, pp. 163–191, 2017.

[119] N. Henschke, J. D. Everett, A. J. Richardson, and I. M. Suthers, "Rethinking the role of salps in the ocean," *Trends in*

Ecology & Evolution, vol. 31, no. 9, pp. 720–733, 2016.

[120] L. P. Madin, "Aspects of jet propulsion in salps," *Canadian Journal of Zoology*, vol. 68, no. 4, pp. 765–777, 1990.

[121] P. Anderson and Q. Bone, "Communication between individuals in salp chains. ii. physiology," *Proceedings of the Royal Society of London. Series B. Biological Sciences*, vol. 210, no. 1181, pp. 559–574, 1980.

[122] K. R. Sutherland and D. Weihs, "Hydrodynamic advantages of swim- ming by salp chains," *Journal of The Royal Society Interface*, vol. 14, no. 133, p. 20170298, 2017.

[123] L. Abualigah, M. Shehab, M. Alshinwan, and H. Alabool, "Salp swarm algorithm: a comprehensive survey," *Neural Computing and Applications*, vol. 32, pp. 11 195–11 215, 2020.

[124] H. Faris, M. M. Mafarja, A. A. Heidari, I. Aljarah, A.-Z. Ala'm,
S. Mirjalili, and H. Fujita, "An efficient binary salp swarm algorithm with crossover scheme for feature selection problems," *Knowledge-Based Systems*, vol. 154, pp. 43–67, 2018.

[125] I. Aljarah, M. Mafarja, A. A. Heidari, H. Faris, Y. Zhang, and
S. Mirjalili, "Asynchronous accelerating multi-leader salp chains for feature selection," *Applied Soft Computing*, vol. 71, pp. 964–979, 2018.

[126] V. Kansal and J. S. Dhillon, "Emended salp swarm algorithm for mul- tiobjective electric power dispatch problem," *Applied Soft Computing*, vol. 90, p. 106172, 2020.

[127] S. M. Ismael, S. H. A. Aleem, A. Y. Abdelaziz, and A. F. Zobaa, "Practical considerations for optimal conductor reinforcement and hosting capacity enhancement in radial distribution systems," *IEEE Access*, vol. 6, pp. 27 268–27 277, 2018.

[128] T. K. Mohapatra and B. K. Sahu, "Design and implementation of ssa based fractional order pid controller for automatic generation control of a multi-area, multi-source interconnected power system," in *2018 Technologies for Smart-City Energy Security and Power (ICSESP)*. IEEE, 2018, pp. 1–6.

[129] S. Asaithambi and M. Rajappa, "Swarm intelligence-based approach for optimal design of cmos differential amplifier and comparator circuit using a hybrid salp swarm algorithm," *Review of Scientific Instruments*, vol. 89, no. 5, p. 054702, 2018.

[130] A. K. Barik and D. C. Das, "Active power management of isolated renewable microgrid generating power from rooftop solar arrays, sewage waters and solid urban wastes of a smart city using salp swarm algorithm," in *2018 Technologies for Smart-City Energy Security and Power (ICSESP)*. IEEE, 2018, pp. 1–6.

[131] A. A. El-Fergany, "Extracting optimal parameters of pem fuel cells using salp swarm optimizer," *Renewable Energy*, vol. 119, pp. 641– 648, 2018.

[132] S. M. H. Baygi, A. Karsaz, and A. Elahi, "A hybrid optimal pid-fuzzy control design for seismic exited structural system against earthquake: A salp swarm algorithm," in *2018 6th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)*. IEEE, 2018, pp. 220–225.

[133] H. Zhao, G. Huang, and N. Yan, "Forecasting energy-related co2 emissions employing a novel ssa-lssvm model: considering structural factors in china," *Energies*, vol. 11, no. 4, p. 781, 2018.

[134] R. A. Khurma, K. E. Sabri, P. A. Castillo, and I. Aljarah, "Salp swarm optimization search based feature selection for enhanced phishing websites detection," in *Applications of Evolutionary Computation: 24th International Conference, EvoApplications 2021, Held as Part of EvoStar 2021, Virtual Event, April 7–9, 2021, Proceedings 24*. Springer, 2021, pp. 146–161.

[135] X.-S. Yang, "Firefly algorithms for multimodal optimization," in *Stochastic Algorithms: Foundations and Applications: 5th Interna- tional Symposium, SAGA 2009, Sapporo, Japan, October 26-28, 2009. Proceedings 5*. Springer, 2009, pp. 169–178.

[136] ——, *Nature-inspired metaheuristic algorithms*. Luniver press, 2010.

[137] X.-S. Yang and X. He, "Firefly algorithm: recent advances and applications," *International journal of swarm intelligence*, vol. 1, no. 1, pp. 36–50, 2013.

[138] M.-H. Horng, Y.-X. Lee, M.-C. Lee, and R.-J. Liou, "Firefly meta- heuristic algorithm for training the radial basis function network for data classification and disease diagnosis," *Theory and new applications of swarm intelligence*, vol. 4, no. 7, pp. 115–132, 2012.

[139] M.-H. Horng, "Vector quantization

using the firefly algorithm for image compression," Expert Systems with Applications, vol. 39, no. 1, pp. 1078–1091, 2012.

[140] H. Banati and M. Bajaj, "Fire fly based feature selection approach," International Journal of Computer Science Issues (IJCSI), vol. 8, no. 4, p. 473, 2011.

[141] A. H. Gandomi, X.-S. Yang, and A. H. Alavi, "Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems," Engineering with computers, vol. 29, pp. 17–35, 2013.

[142] B. Basu and G. Mahanti, "Fire fly and artificial bees colony algorithm for synthesis of scanned and broadside linear array antenna," Progress In Electromagnetics Research B, vol. 32, pp. 169–190, 2011.

[143] A. Chatterjee, G. Mahanti, and A. Chatterjee, "Design of a fully digital controlled reconfigurable switched beam concentric ring array antenna using firefly and particle swarm optimization algorithm," Progress In Electromagnetics Research B, vol. 36, pp. 113–131, 2012.

[144] X.-S. Yang, "Swarm-based metaheuristic algorithms and no-free-lunch theorems," Theory and new applications of swarm intelligence, vol. 9, pp. 1–16, 2012.

[145] G. K. Jati, "Evolutionary discrete firefly algorithm for travelling salesman problem," in Adaptive and Intelligent Systems: Second International Conference, ICAIS 2011, Klagenfurt, Austria, September 6-8, 2011. Proceedings. Springer, 2011, pp. 393–403.

[146] J. Senthilnath, S. Omkar, and V. Mani, "Clustering using firefly algorithm: performance study," Swarm and Evolutionary Computation, vol. 1, no. 3, pp. 164–171, 2011.

[147] S. Nandy, P. P. Sarkar, and A. Das, "Analysis of a nature inspired firefly algorithm based back-propagation neural network training," arXiv preprint arXiv:1206.5360, 2012.

[148] S. M. Farahani, A. A. Abshouri, B. Nasiri, and M. Meybodi, "A gaussian firefly algorithm," International Journal of Machine Learning and Computing, vol. 1, no. 5, p. 448, 2011.

[149] M. Sayadi, R. Ramezanian, and N. Ghaffari-Nasab, "A discrete firefly meta-heuristic with local search for makespan minimization in per- mutation flow shop scheduling problems," International Journal of Industrial Engineering Computations, vol. 1, no. 1, pp. 1–10, 2010.

[150] T. Apostolopoulos and A. Vlachos, "Application of the firefly algo- rithm for solving the economic emissions load dispatch problem," International journal of combinatorics, vol. 2011, 2011.

[151] L. dos Santos Coelho, D. L. de Andrade Bernert, and V. C. Mariani, "A chaotic firefly algorithm applied to reliability-redundancy optimiza- tion," in 2011 IEEE congress of evolutionary computation (CEC). Ieee, 2011, pp. 517–521.

[152] G. Vrbancˇicˇ, I. Fister Jr, and V. Podgorelec, "Parameter setting for deep neural networks using swarm intelligence on phishing websites classification," International Journal on Artificial Intelligence Tools, vol. 28, no. 06, p. 1960008, 2019.

[153] K. Tayal and V. Ravi, "Particle swarm optimization trained class asso- ciation rule mining: Application to phishing detection," in Proceedings of the International Conference on Informatics and Analytics, 2016, pp. 1–8.