







Enhancing Spam Detection Using Hybrid of Harris Hawks and Firefly Optimization Algorithms

Mosleh M. Abualhaj^{1,*} , Qusai Y. Shambour² , Adeeb Alsaaidah³ , Ahmad Abu-Shareha⁴ ,
Sumaya Al-Khatib⁵ , Mohammad O. Hiari⁶ 

^{1,2,3,4,5,6}Egypt Faculty of Information Technology, Al-Ahliyya Amman University, Amman 19111, Jordan

(Received: May 15, 2024; Revised: June 30, 2024; Accepted: July 4, 2024; Available online: July 16, 2024)

Abstract

The emergence of the modern Internet has presented numerous opportunities for attackers to profit illegally by distributing spam mail. Spam refers to irrelevant or inappropriate messages that are sent on the Internet to numerous recipients. Many researchers use many classification methods in machine learning to filter spam messages. However, more research is still needed to assess using metaheuristic optimization algorithms to classify spam emails in feature selection. In this paper, we endorse fighting spam emails by proposing a new feature selection method that employs a union of Firefly Optimization Algorithm (FOA) and Harris Hawks Optimization (HHO) algorithms to classify spam emails, along with one of the most well-known and efficient methods in this area, the Random Forest (RF) classifier. The proposed union feature selection method creates a more robust and comprehensive set of features by combining the selected features by FOA and HHO algorithms. This approach leverages the strengths of FOA and HHO algorithms to capture a wide range of important features that might be missed by using a single method. By integrating diverse methods, union feature selection enhances the model's ability to generalize to new data, reducing overfitting and improving overall accuracy. In this process, the experimental studies on the ISCX-URL2016 spam dataset yield promising results. For instance, the union of HHO and FOA, along with using an RF classifier, achieved an accuracy of 99.83% in detecting spam emails.

Keywords: Spam emails; Machine learning; Feature selection; Firefly Optimization Algorithm; Harris Hawks Optimization; and Random Forest.

1. Introduction

Information technology has facilitated the provision of diverse and productive services, such as e-mail. Although mobile messengers and chat apps have become increasingly popular, e-mail remains an essential part of our everyday internet activities [1]. Worldwide, the daily volume of e-mails sent and received in 2020 amounted to roughly 306 billion [2]. A significant number of internet users utilize e-mail addresses to register for websites and subscribe to newsletters, anticipating the subsequent inundation of unsolicited messages and promotional content. While most unsolicited e-mails may be bothersome but ultimately harmless, consumers should exercise caution regarding dangerous e-mails that can potentially damage their digital accounts and devices [3]. By 2020, around 50% of global e-mails were classified as spam [2].

Spam is a significant danger, leading to significant problems for users across several platforms, including email systems, online social networks, and consumer reviews. The proliferation of email services has transformed spam into a potent weapon for cybercriminals. They can now send deceptive content, including obfuscated URLs, to external sites that may harbor malware, phishing web pages, and other irrelevant content, such as advertisements. This poses a significant threat and can result in severe and damaging attacks. Spam risks persist beyond digital platforms and pose a risk to individuals in the real world. For example, numerous spammers specifically aim to obtain sensitive data like bank account details and credit card information [4], [5].

Currently, most spam filters implement machine learning (ML) techniques, which is part of artificial intelligence. Machine learning allows the machines to make predictions based on previous experience from the available data. Machine learning encompasses various approaches, including supervised learning, unsupervised learning, and reinforcement learning. Supervised learning involves training algorithms on labeled data to predict outcomes, whereas

*Corresponding author: Mosleh M. Abualhaj (1m.abualhaj@ammanu.edu.jo)

 DOI: <https://doi.org/10.47738/jads.v5i3.279>

This is an open access article under the CC-BY license (<https://creativecommons.org/licenses/by/4.0/>).

© Authors retain all copyrights

unsupervised learning involves training algorithms on unlabeled data. On the other hand, in reinforcement learning, the agents learn by interacting with their environment to maximize rewards [4], [5], [6].

Handling large amounts of data in ML is a key challenge in the performance of ML techniques. Using optimization algorithms can effectively handle and process large data, leading to more efficient and scalable machine-learning solutions. Metaheuristic optimization algorithms are powerful tools for handling large data and complex optimization problems in machine learning. These algorithms are designed to find near-optimal solutions within a reasonable time frame, especially when dealing with high-dimensional and non-convex optimization landscapes [7], [8].

In this work, we investigate the application ML model based on the FOA and HHO along with the Random Forest (RF) to improve the detection of Spam emails. HHO, FOA, and RF are widely used in the field of cybersecurity to detect different types of attacks [7], [8], [9].

The paper is organized as follows: Section 2 presents the literature review. Section 3 discusses the proposed methodology in detail, including the used ISCX-URL2016 dataset, the proposed union feature selection method, and the RF classifier and its hyperparameters. Section 4 shows and discusses the achieved results. Section 6 concludes the paper.

2. Literature Review

Extensive research has been conducted on email spam filtering. Xu and Yu [10] present a spam filtering framework that employs a modified back propagation (RBP) neural network and autonomous thesaurus generation. The conventional backpropagation (BP) neural network has moderate learning speed and is prone to getting stuck in local minima, resulting in low performance and efficiency. The authors demonstrate that the RBP neural network overcomes the limitations of the conventional BP neural network. A well-designed thesaurus is considered a valuable tool in effectively organizing content. It can also overcome the limitations of keyword-based spam filters that fail to recognize the semantic relationships between words.

The work in [11] presented a hybrid approach that combines the Negative Selection Algorithm (NSA) with the differential evolution (DE). DE has implemented the random generation phase detector distance NSA, which aims to maximize the results and reduce the overlapping of the detector. DE is utilized to enhance the creation of detectors during the NSA stage, with the local outlier factor (LOF) serving as the fitness function. The method was evaluated on a spam dataset and achieved an accuracy of approximately 83.06%. An innovative approach suggested by [12] was adopted to enhance the random generation of a detector in the NSA. This approach utilizes stochastic distribution to represent the data point and employs particle swarm optimization (PSO) for optimization. The LOF is utilized as the fitness function to ascertain the local best (Pbest) of the candidate detector that yields the optimal solution. The dataset utilized is the Spambase dataset supplied by the UCI machine learning library. The proposed NSA-PSO method was tested on a spam dataset and achieved a precision level of 91.22%.

Using the n-gram method, Congfu Xu et al. [13] have developed a feature extraction method based on the images' Base64 encoding. Through the training of a Support Vector Machine (SVM), these features demonstrate their usefulness and efficiency in distinguishing spam photos from real images. The results of the experiments indicate that it has remarkable performance in classifying spam photos in terms of accuracy, precision, and recall. Holly Esquivel et al.'s [14] primary focus was a pre-acceptance modifying process of IP reputation. First, they divided SMTP senders into three primary types: genuine servers, end hosts, and spam gangs. Next, they conducted an empirical investigation to determine the limitations of the effectiveness of IP reputation filtering for each of these categories.

Considering the works mentioned above, the main concern with the methods is dealing with Spam emails. Some of the methods use typical feature selection algorithms to reduce the size of high-dimensional data. More specifically, these methods do not employ a combination of optimization algorithms for feature selection, nor do they use any optimization algorithms at all. Therefore, in this work, we will investigate and combine two of the most common optimization algorithms (HHO and FOA algorithms) in order to provide a design of an ML model that deals with the high dimensionality of the data in Spam emails.

3. Research Methodology

3.1. ISCX-URL2016 Dataset

This paper will use the ISCX-URL2016 dataset. The ISCX-URL2016 dataset is a comprehensive collection of URLs designed for cybersecurity research, specifically focusing on the detection of malicious URLs. The dataset was created by researchers at the Canadian Institute for Cybersecurity (CIC), which is part of the University of New Brunswick (UNB). It is part of a broader effort by the institute to provide high-quality datasets for the research community to develop and evaluate security solutions. The dataset was built using 79 features extracted from URLs. The dataset labels five URL classes: benign, defacement, malware, phishing, and spam. This paper is only interested in a subset of the ISCX-URL2016 dataset that contains spam samples, called ISCX-URL2016-spam. The ISCX-URL2016-spam dataset contains 14479 samples [15], [16].

Data pre-processing is crucial for ensuring the consistency and cleanliness of datasets. The ISCX-URL2016-spam dataset contains balanced samples of benign and spam (6,698 Spam and 7,780 benign samples). In addition, all the values in the dataset are numerical. However, the dataset contains many Null values. Many methods can be used to handle the Null values, including dropping the entire feature column. All the features with Null values were removed from the dataset by removing their column. Removing features with null values is a justified approach when it improves data quality, reduces complexity, and maintains or enhances model performance. Seven features with Null values were removed, leaving only 72 features in the dataset. Table 1 shows the features of the ISCX-URL2016-spam dataset [15], [16].

Table 1. Features of the ISCX-URL2016-spam dataset

#	Feature	#	Feature	#	Feature
1	Querylength	25	ArgLen	49	Query_LetterCount
2	domain_token_count	26	pathurlRatio	50	LongestPathTokenLength
3	path_token_count	27	ArgUrlRatio	51	Domain_LongestWordLength
4	avgdomaintokenlen	28	argDomanRatio	52	Path_LongestWordLength
5	longdomaintokenlen	29	domainUrlRatio	53	sub-Directory_LongestWordLength
6	tld	30	pathDomainRatio	54	Arguments_LongestWordLength
7	charcompvowels	31	argPathRatio	55	URL_sensitiveWord
8	charcompaace	32	executable	56	URLQueries_variable
9	ldl_url	33	isPortEighty	57	spcharUrl
10	ldl_domain	34	NumberofDotsinURL	58	delimiter_Domain
11	ldl_path	35	ISIpAddressInDomainName	59	delimiter_path
12	ldl_filename	36	CharacterContinuityRate	60	delimiter_Count
13	ldl_getArg	37	LongestVariableValue	61	NumberRate_URL
14	dld_url	38	URL_DigitCount	62	NumberRate_Domain
15	dld_domain	39	host_DigitCount	63	NumberRate_DirectoryName
16	dld_path	40	Directory_DigitCount	64	NumberRate_FileName
17	dld_filename	41	File_name_DigitCount	65	SymbolCount_URL
18	dld_getArg	42	Extension_DigitCount	66	SymbolCount_Domain
19	urlLen	43	Query_DigitCount	67	SymbolCount_Directoryname
20	domainlength	44	URL_Letter_Count	68	SymbolCount_FileName

21	pathLength	45	host_letter_count	69	SymbolCount_Extension
22	subDirLen	46	Directory_LetterCount	70	SymbolCount_Afterpath
23	fileNameLen	47	Filename_LetterCount	71	Entropy_URL
24	this.fileExtLen	48	Extension_LetterCount	72	Entropy_Domain

In addition, the dataset contains values with different scales. Data normalization is a widely utilized pre-processing technique that involves transforming data to a standardized scale. It enhances both accuracy and learning speed. Traditionally, the Min-max scaling normalization has been extensively employed in most machine learning applications as a data normalization technique. The Min-max scaling technique scales the features of a dataset to a fixed range, usually between 0 and 1 [16]. Table 2 shows a sample of the ISCX-URL2016-spam dataset before and after normalization, using the Min-max scaling technique.

Table 2. Sample of the ISCX-URL2016-spam dataset before and after normalization

Before Normalization	After Normalization
0, 2, 5.5, 2, 7	0, 0, 0.318182, 0, 0.049296
0, 3, 5, 3, 8	0, 0.333333, 0.272727, 0.333333, 0.056338
2, 2, 4, 2, 11	0.001444, 0, 0.181818, 0, 0.077465
0, 2, 4.5, 2, 10	0, 0, 0.227273, 0, 0.070423
19, 2, 6, 2, 5	0.013718, 0, 0.363636, 0, 0.035211

3.2. Proposed Union Feature Selection

We analyze the features of the ISCX-URL-2016-spam dataset to gain a deeper understanding of how to create effective methods for detecting and categorizing spam emails. The dataset has a total of 72 features. These features inherently differ in their ability to predict outcomes and, as a result, their utility for a machine learning system. Therefore, employing a suitable feature selection algorithm has several benefits. First, the irrelevant features are effectively reduced in the ISCX-URL-2016 spam dataset, minimizing the computational resources needed to make accurate predictions for real-world samples. Furthermore, less time and computational resources are allocated to producing irrelevant or redundant features during the spam identification process. Moreover, the features with highest correlation to the classification of a malicious URL sample will be selected [18], [19], [20].

In this paper, the FOA and HHO will be used for feature selection. HHO is a nature-inspired optimization algorithm modeled after the cooperative hunting strategies of Harris hawks. It emulates the dynamic and collaborative behaviors of these birds as they engage in surprise attacks and other complex hunting tactics to capture prey [8]. On the other hand, FOA is a nature-inspired optimization technique based on the flashing behavior of fireflies. FOA mimics the way fireflies attract each other using bioluminescent signals [7]. In feature selection for spam detection HHO and FOA offer complementary strengths. HHO dynamically explores the search space with strategies like soft and hard besiege, ensuring thorough coverage and effective avoidance of local optima. It maintains a global perspective by balancing between exploration and exploitation phases, thereby enhancing the diversity of solutions explored. On the other hand, FA utilizes an attraction mechanism based on brightness, guiding fireflies towards better solutions and promoting convergence. Its randomized movement fosters exploration across the search space, preventing premature convergence and enabling the algorithm to explore multiple potential solutions simultaneously. By integrating these strengths, the combined approach in the Spam detection achieves comprehensive feature selection, improving accuracy and robustness in spam detection applications [7], [8].

Typically, feature selection uses a single method, such as the HHO algorithm, to identify features with the best predictive capabilities, thereby improving the performance of a machine learning model. However, this paper proposes a novel feature selection approach that combines two robust and well-known algorithms: FOA and HHO. First, the HHO algorithm is applied to the ISCX-URL-2016-spam dataset to select the best features based on its operational

behavior. Next, the FOA algorithm is used on the same dataset to identify the most relevant features according to its mechanisms. Finally, the two subsets of features selected by the HHO and FOA algorithms are combined to form a single subset. This union subset contains features deemed optimal by both algorithms, thus leveraging the strengths of both FOA and HHO. The union subset will be evaluated using the RF classifier to determine its effectiveness in improving classification performance. Metrics such as accuracy, precision, recall, and F1-score are used for evaluation. The performance of the union subset will be compared against the subsets selected by HHO and FOA individually. This comparison helps in validating whether the combined subset offers any significant improvement over the individual subsets (See Section 6). Table 3 lists the union of features from FOA and HHO algorithms. Figure 1 illustrates the union feature selection method. By using this approach, the resulting subset of features is optimized based on the complementary strengths of the FOA and HHO algorithms, leading to improved predictive performance of the machine learning model.

Table 3. Selected feature by different methods

Method	Selected features (feature #)	Total selected
FOA	0,1,3,5,7,9,11,12,19,21,22,25,28,30,32,35,38,39,41,42,44,46,47,49,51,53,59,67,68,70,71	31
HHO	5,17,25,26,29,33,35,46,61,62	10
Union of FOA & HHO	0,1,3,5,7,9,11,12,17,19,21,22,25,26,28,29,30,32,33,35,38,39,41,42,44,46,47,49,51,53,59,61,62,67,68,70,71	38

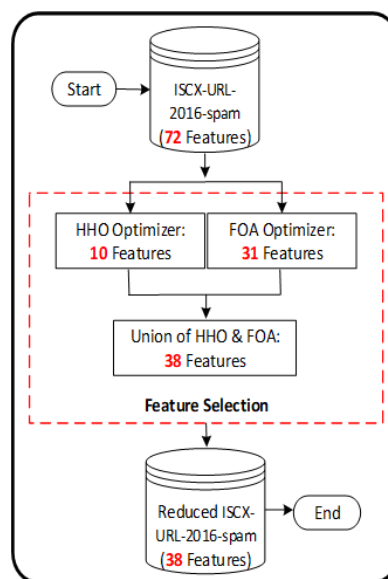


Figure 1. Proposed Union feature selection method

3.3. Spam Classification using RF

In the previous steps (Section 3 and 4), the data was processed and prepared for classification, distinguishing spam from benign e-mails. In this paper, the classification task will utilize the RF classifier. RF is an ML classifier specifically developed to address classification tasks. RF utilizes ensemble learning, which involves the integration of several classifiers to tackle intricate problems. The RF classifier consists of many decision trees (DT), as shown in Figure 2. The predictions of the RF classifier are obtained by combining the projected outcomes of these DTs. RF improves accuracy by calculating the average or mean of the outputs generated by multiple DTs. The RF classifier is crucial in addressing the constraints of individual DTs, particularly in avoiding overfitting and enhancing overall accuracy. Increasing the quantity of trees in the ensemble further amplifies accuracy [21], [22].

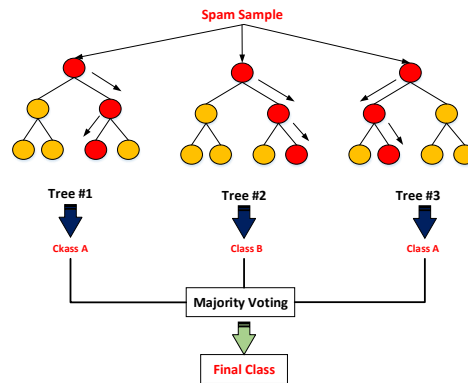


Figure 2. RF classifier

The RF classifier algorithm contains several hyperparameters. These hyperparameters are utilized to either enhance the performance and predictive power of systems or to make the systems faster. Hyperparameters utilized to increase the predictive power are ‘n_estimators’, ‘max_features’, ‘mini_sample_leaf’, ‘criterion’, and ‘max_leaf_nodes’. Hyperparameters utilized to increase the speed are ‘n_jobs’, ‘random_state’, and ‘oob_score’. In the suggested HHO-FOA-RF model, the RF hyperparameters are assigned certain values to improve the system's performance in malware detection. Table 4 shows the values of the RF hyperparameters utilized by the HHO-FOA-RF model [21], [22], [23]. The value of these hyperparameters is chosen using grid search. Grid search is a method for systematically exploring the hyperparameter space and identifying the most impactful hyperparameters for a classifier to optimize ML model performance.

Table 4. RF Hyperparameters

Hyperparameters	Assigned value	Description and Impact
n_estimators	100	Number of trees in the forest; balances performance and computational cost
max_features	sqrt	Number of features to consider for best split; balances between randomness and accuracy
mini_sample_leaf	1	Minimum samples required at a leaf node; allows deep growth to capture patterns
criterion	gini	Function to measure split quality; 'gini' is computationally efficient
max_leaf_nodes	None	Maximum number of leaf nodes; allows trees to grow based on other stopping criteria
n_jobs	None	Number of jobs to run in parallel; '-1' utilizes all available CPU cores
random_state	None	Controls randomness for reproducibility; any fixed integer ensures consistent results
oob_score	False	Use out-of-bag samples to estimate accuracy; alternative to cross-validation

At this level, the suggested ML model is complete to detect spam e-mails. The suggested model uses a combination of the HHO and FOA algorithms for feature selection. In addition, it uses the RF classifier with specific parameters to perform the classification task. Therefore, the suggested ML model is called HHO-FOA-RF. Figure 3 shows the suggested HHO-FOA-RF ML model. The following section presents the performance of the HHO-FOA-RF model.

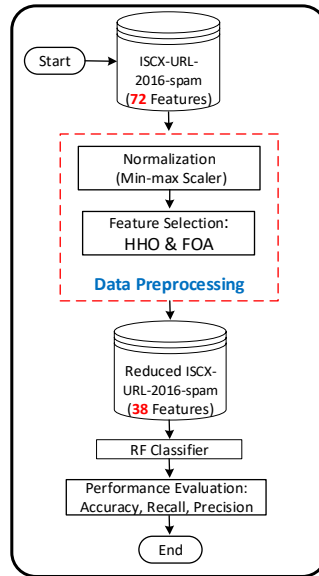


Figure 3. The suggested HHO-FOA-RF ML model

4. Results and Discussion

In this section we evaluated the proposed method in classifying spam emails using the ISCX-URL2016-spam dataset. Several tools of Python were used in the evaluation of the proposed model including Numpy, Pandas, sklearn, ensemble, RandomForestClassifier, mealpy, swarm_based, HHO.OriginalHHO, and FOA.DevFOA.

4.1. Evaluation metrics

The suggested HHO-FOA-RF model is assessed using accuracy, precision, and recall. Taking into account all three measures guarantees the reliability of the findings. In figure 4, the confusion matrix is provided to calculate the evaluation measures [24]. In the proposed HHO-FOA-RF spam detection model, the spam class is designated as the positive class, whereas the benign class is designated as the negative class. Each evaluation metric is described based on this assumption, and they are as follows. Accuracy is the proportion of correctly classified spam and benign emails divided by the total number classified emails, (1) [24]. Precision is the proportion of the correctly classified spam emails divided by the total number of predicted spam emails, (2) [24]. Recall is the number of correctly predicted spam emails divided by the total number of spam emails, (3) [24]. Matthews Correlation Coefficients (MCC) is a measure of the quality of classification with two classes, (4) [24]. F1-Score is the harmonic mean of Precision and Recall, (5) [24].

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

Figure 4. Confusion matrix

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (1)$$

$$Recall = \frac{TP}{(TP + FN)} \quad (2)$$

$$Precision = \frac{TP}{(TP + FP)} \quad (3)$$

$$MCC = \frac{((TP * TN) - (FP * FN))}{\sqrt{(TP + FP) * (TP + FN) * (TN + FP) * (TN + FN)}} \quad (4)$$

$$F1 - score = 2 \times \frac{Pre \times Rec}{Pre + Rec} \quad (5)$$

4.2. Experiments

The proposed HHO-FOA-RF ML model is evaluated using the accuracy, precision, and recall metrics on the ISCX-URL-2016-spam dataset. Five separate experiments are conducted for each metric, with 80% of the dataset used for training the HHO-FOA-RF model and the remaining 20% for testing. The final value of each metric is determined by averaging the results from the five experiments. To ensure reliable and comprehensive results, the training and testing sets are rotated through all the dataset samples. The values of FP, TN, FN, and TP are 0, 1522, 6, and 1368, respectively, with the HHO algorithm. The values of FP, TN, FN, and TP are 2, 1520, 6, and 1368, respectively, with the FOA algorithm. The values of FP, TN, FN, and TP are 3, 1519, 6, and 1368, respectively, with HHO union FOA algorithms.

Figure 5 illustrates the accuracy of the proposed HHO-FOA-RF model, which combines the HHO and FOA algorithms for feature selection, compared to the typical ML model using HHO or FOA algorithms individually. The proposed HHO-FOA-RF model achieved an accuracy of 99.83%, whereas the typical model attained accuracies of 99.69% and 99.72% when using HHO and FOA, respectively. Figure 6 illustrates the precision of the proposed HHO-FOA-RF model, compared to the typical ML model using HHO or FOA algorithms individually. The proposed HHO-FOA-RF model achieved a precision of 99.72%, whereas the typical model attained precisions of 99.69% and 99.69% when using HHO and FOA, respectively. Figure 7 illustrates the recall of the proposed HHO-FOA-RF model, compared to the typical ML model using HHO or FOA algorithms individually. The proposed HHO-FOA-RF model achieved a recall of 99.74%, whereas the typical model attained recalls of 99.71% and 99.69% when using HHO and FOA, respectively. Figure 8 illustrates the MCC of the proposed HHO-FOA-RF model, compared to the typical ML model using HHO or FOA algorithms individually. The proposed HHO-FOA-RF model achieved a MCC of 99.45%, whereas the typical model attained MCCs of 99.38% and 99.37% when using HHO and FOA, respectively. Figure 9 illustrates the F1-Score of the proposed HHO-FOA-RF model, compared to the typical ML model using HHO or FOA algorithms individually. The proposed HHO-FOA-RF model achieved F1-Score of 99.72%, whereas the typical model attained MCCs of 99.69% and 99.69% when using HHO and FOA, respectively. As shown, the proposed HHO-FOA-RF model outperforms the typical models employing HHO or FOA separately in accuracy, precision, recall, MCC, and F1-Score metrics. This comprehensive presentation of performance indicators offers a nuanced assessment of the model's efficacy across diverse metrics. The proposed union HHO-FOA-RF model contributes to advancing systems for mitigating the impact of spam, underscoring the effectiveness of employing feature unions through HHO and FOA to enhance feature selection in the context of spam classification.

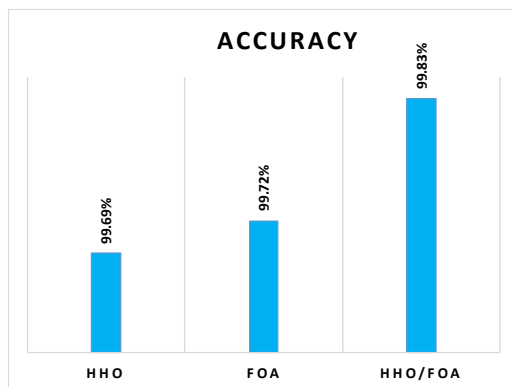


Figure 5. Accuracy of the proposed HHO-FOA-RF model

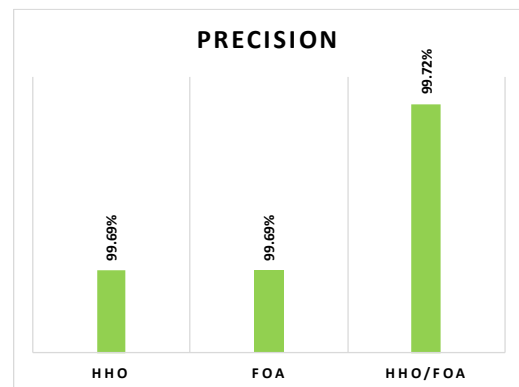


Figure 6. Precision of the proposed HHO-FOA-RF model

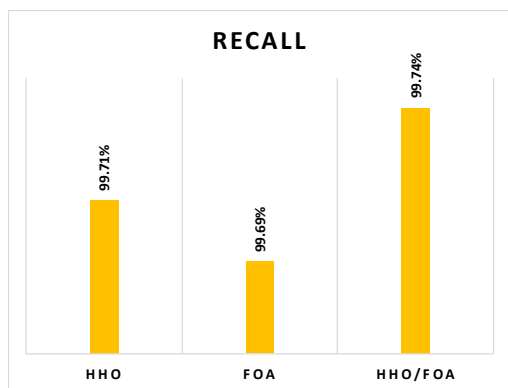


Figure 7. Recall of the proposed HHO-FOA-RF model

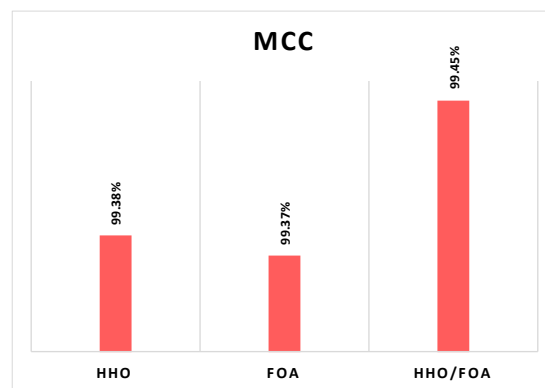


Figure 8. MCC of the proposed HHO-FOA-RF model

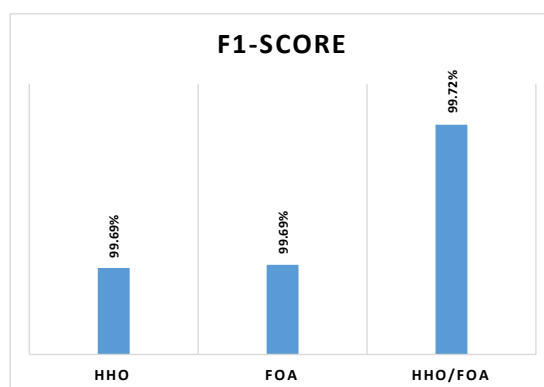


Figure 9. F1-Score of the proposed HHO-FOA-RF model

5. Conclusion

An application of a union of two well-known feature selection algorithms, the HHO and FOA algorithms, was developed in this study. This application aimed to identify the most significant features of the learning process and help detect spam. The ISCX-URL2016, which is available publicly, served as the source for the features extracted from the datasets. Two phases of the experiments were carried out. The initial experiment was conducted using the RF classifier to independently examine the performance of HHO or FOA algorithms. After that, we explored the effect of the union of HHO and FOA algorithms with the RF classifier. RF's performance in spam identification was improved by the combination of HHO and FOA algorithms, as demonstrated by the study's findings. Regarding accuracy, precision, and recall measures, the suggested HHO-FOA-RF model outperforms the commonly used models that employ either HHO or FOA independently. The HHO-FOA-RF ML model can be integrated into many real-world applications spanning various domains, including email service providers, enterprise security, social media platforms, and mobile and web applications. Future research could focus on exploring additional datasets to validate the Spam-FA-HHO model's generalizability and effectiveness across diverse spam types. Investigating other combinations of optimization algorithms, beyond FOA and HHO, as well as testing other classifiers.

6. Declarations

6.1. Author Contributions

Conceptualization: M.M.A. and A.A.S.; Methodology: M.M.A. and A.A.S.; Software: S.A. and M.O.H.; Validation: A.A. and Q.Y.S.; Formal Analysis: M.M.A.; Investigation: M.M.A.; Resources: S.A. and M.O.H.; Data Curation: S.A.; Writing Original Draft Preparation: M.M.A. and Q.Y.S.; Writing Review and Editing: Q.Y.S.; Visualization: A.A.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] J. Wei, X. Chen, J. Wang, X. Hu and J. Ma, "Enabling (End-to-End) Encrypted Cloud Emails With Practical Forward Secrecy," in *IEEE Transactions on Dependable and Secure Computing*, vol. 19, no. 4, pp. 2318-2332, 1 July-Aug. 2022, doi: 10.1109/TDSC.2021.3055495.
- [2] A. Rastogi and M. Mehrotra, "Impact of Behavioral and Textual Features on Opinion Spam Detection," *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India*, vol. 2, no. 1, pp. 852-857, 2018. doi: 10.1109/ICCONS.2018.8662912.
- [3] R. Li, Z. Zhang, J. Shao, R. Lu, X. Jia and G. Wei, "The Potential Harm of Email Delivery: Investigating the HTTPS Configurations of Webmail Services," in *IEEE Transactions on Dependable and Secure Computing*, vol. 21, no. 1, pp. 125-138, Jan.-Feb. 2024, doi: 10.1109/TDSC.2023.3246600.
- [4] G. Kambourakis, G. D. Gil and I. Sanchez, "What Email Servers Can Tell to Johnny: An Empirical Study of Provider-to-Provider Email Security," in *IEEE Access*, vol. 8, no. 1, pp. 130066-130081, 2020, doi: 10.1109/ACCESS.2020.3009122.
- [5] A. AlMahmoud, E. Damiani, H. Otrouk and Y. Al-Hammadi, "Spamdoop: A Privacy-Preserving Big Data Platform for Collaborative Spam Detection," in *IEEE Transactions on Big Data*, vol. 5, no. 3, pp. 293-304, 1 Sept. 2019, doi: 10.1109/TBDATA.2017.2716409.
- [6] M. M. Abualhaj, Ahmad Adel Abu-Shareha, Qusai Shambour, Adeeb Alsaaidah, S. N. Al-Khatib, and M. Anbar, "Customized K-nearest neighbors' algorithm for malware detection," *International journal of data and network science*, vol. 8, no. 1, pp. 431-438, Jan. 2024, doi: 10.5267/j.ijdns.2023.9.012.
- [7] A. Al Saaaidah, M. M. Abualhaj, Q. Y. Shambour, "Enhancing malware detection performance: leveraging K-Nearest Neighbors with Firefly Optimization Algorithm," *Multimed Tools Appl*, vol. 1, no. 1, pp. 1-12, 2024, doi:10.1007/s11042-024-18914-5.
- [8] K. Dev, P. K. R. Maddikunta, T. R. Gadekallu, S. Bhattacharya, P. Hegde and S. Singh, "Energy Optimization for Green Communication in IoT Using Harris Hawks Optimization," in *IEEE Transactions on Green Communications and Networking*, vol. 6, no. 2, pp. 685-694, June 2022, doi: 10.1109/TGCN.2022.3143991.
- [9] M. M. Abualhaj, Ahmad Adel Abu-Shareha, M. O. Hiari, Yousef Alrabanah, Mahran Al-Zyoud, and M. A. Alsharaiah, "A Paradigm for DoS Attack Disclosure using Machine Learning Techniques," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 3, pp. 192-200, Jan. 2022, doi: https://doi.org/10.14569/ijacsa.2022.0130325.
- [10] H. Xu and B. Yu, "Automatic thesaurus construction for spam filtering using revised back propagation neural network," *Expert Syst. Appl.*, vol. 37, no. 1, pp. 18-23, Jan. 2010.
- [11] I. Idris, A. Selamat, and S. Omatu, "Hybrid email spam detection model with negative selection algorithm and differential evolution," *Eng. Appl. Artif. Intell.*, vol. 28, no. 1, pp. 97-110, Feb. 2014.
- [12] I. Idris and A. Selamat, "Improved email spam detection model with negative selection algorithm and particle swarm optimization," *Appl. Soft Comput.*, vol. 22, no. 1, pp. 11-27, 2014.
- [13] C. Xu, Y. Chen and K. Chiew, "An Approach to Image Spam Filtering Based on Base64 Encoding and N-Gram Feature Extraction," 2010 22nd IEEE International Conference on Tools with Artificial Intelligence, Arras, France, 2010, vol. 1, no.

- 1, pp. 171-177, doi: 10.1109/ICTAI.2010.31.
- [14] Holly Esquivel and Aditya Akella, "On the effectiveness of IP reputation for spam filtering", IEEE international Conference on Communication Systems and Networks, vol. 2010, no. 1, pp. 1-10, 2010.
- [15] R. Aloufi and A. R. Alharbi, "K-means and Principal Components Analysis Approach For Clustering Malicious URLs," 2023 3rd International Conference on Computing and Information Technology (ICCIT), Tabuk, Saudi Arabia, vol. 3, no. 1-12, pp. 359-364, 2023, doi: 10.1109/ICCIT58132.2023.10273923.
- [16] M. S. I. Mamun, M. A. Rathore, A. H. Lashkari, N. Stakhanova, and A. A. Ghorbani, "Detecting malicious urls using lexical analysis." In *Network and System Security: 10th International Conference, NSS 2016, Taipei, Taiwan Springer International Publishing*, vol. 10, no. sept, pp. 467-482, 2016, doi: 10.1007/978-3-319-46298-1_30.
- [17] Al-Mimi, H.M. et al., "Improved intrusion detection system to alleviate attacks on DNS service," *Journal of Computer Science*, vol. 19, no. 12, pp. 1549–1560, 2023. doi:10.3844/jcssp.2023.1549.1560.
- [18] A. Slowik and K. Cpalka, "Guest Editorial: Hybrid Approaches to Nature-Inspired Population-Based Intelligent Optimization for Industrial Applications," in *IEEE Transactions on Industrial Informatics*, vol. 18, no. 1, pp. 542-545, Jan. 2022, doi: 10.1109/TII.2021.3091137.
- [19] M. Kolhar, F. Al-Turjman, A. Alameen and M. M. Abualhaj, "A Three Layered Decentralized IoT Biometric Architecture for City Lockdown During COVID-19 Outbreak," in *IEEE Access*, vol. 8, no. 1, pp. 163608-163617, 2020, doi: 10.1109/ACCESS.2020.3021983.
- [20] F. Liu et al., "Nonconvex Compressed Sensing by Nature-Inspired Optimization Algorithms," in *IEEE Transactions on Cybernetics*, vol. 45, no. 5, pp. 1042-1053, May 2015, doi: 10.1109/TCYB.2014.2343618.
- [21] Y. Ren, X. Zhu, K. Bai, and R. Zhang, "A new random forest ensemble of intuitionistic fuzzy decision trees," *IEEE Transactions on Fuzzy Systems*, vol. 31, no. 5, pp. 1729–1741, May 2023, doi: 10.1109/tfuzz.2022.3215725.
- [22] H. Al-Mimi, N. A. Hamad, and M. M. Abualhaj, "A Model for the Disclosure of Probe Attacks Based on the Utilization of Machine Learning Algorithms," 2023 10th International Conference on Electrical and Electronics Engineering (ICEEE), vol. 10, no. May, pp. 241-247, May 2023, doi: 10.1109/iceee59925.2023.00051.
- [23] N. Pavitha and S. R. Sugave, "Ensemble Approach with Hyperparameter Tuning for Credit Worthiness Prediction," 2022 IEEE 3rd Global Conference for Advancement in Technology (GCAT), vol. 3, no. oct, pp. 1-5, Oct. 2022, doi: 10.1109/gcat55367.2022.9971879.
- [24] M. M. Abualhaj and S. N. Al-Khatib, "Using decision tree classifier to detect Trojan Horse based on memory data," *Telkomnika*, vol. 22, no. 2, pp. 393–393, Apr. 2024.