Received: December 10, 2020. Revised: January 10, 2021. 314

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

SMS Spam Detection Based on Fuzzy Rules and Binary Particle Swarm Optimization

Sarab M. Hameed[1]* Zuhair Hussein Ali[2]

[1]*Department Computer Science, College of Science, University of Baghdad, Iraq*
[2]*Department of Computer Science, College of Education, Mustansiriyah University, Iraq*
* Corresponding author's Email: sarab.m@sc.uobaghdad.edu.iq

Abstract: Over the last decade, the usage of short message services (SMS) as one of the vital communication services on mobile devices has grown. The growth of using this service has correspondingly increased the number of attacks on mobile devices such as SMS Spam. SMS spam is a concern to telecommunications service providers as they annoy the subscribers and cause them to loose commercial. Most current researches have attempted to detect SMS spam using different classifiers. In this paper, we propose a new method that focuses on binary particle swarm optimization-based fuzzy rules selection for detecting SMS spam messages. First, we extract significant features from the SMS spam dataset. Then, a set of fuzzy rules based on the extracted features is generated. Finally, a binary particle swarm is suggested for picking the more powerful fuzzy rules that reduce the complexity and improve the model's performance. The SMS Spam benchmark dataset was used in the experiment. The attained results of the proposed model provide a recall of 98.8%, precision of 90.8%, F-measure of 94.6% and accuracy of 98.5% that indicate the proposed model can be a promising for detecting SMS spam.

Keywords: Fuzzy rule, Particle swarm optimization, SMS spam filtering, SMS spam detection.

1. Introduction

The development of mobile communication technologies and mobile phone extension led Short Message Service (SMS) to be one of the most common means of communication among millions of individuals, where the transmission of the messages must take place in compliance with conventional communication protocols. The number of SMS messages sent has increased incredibly according to a variety of reasons: Users read their SMS daily, while most of the emails message remain unread for more than one day. In addition, the cost of SMS is very cheap and possibly zero in some cases. Finally, the number of mobile phone users increased rapidly, it can be up to several million in some countries such as India and China. This gives it importance for use in many applications, especially in business [1, 2].
Unfortunately, mobile phones have been an instrument of what has been defined as SMS spam. SMS spam refers specifically to any unwanted messages sent through mobile networks [3]. In general, two main goals for spammers: advertising and stealing user information. Due to the widespread use of mobile phones, advertising through SMS is becoming easy and fast. SMS advertising is the promotion of products and services via mobile phones. Companies send short pieces of information to subscribers worldwide with the help of bulk SMS services to increase brand awareness and build trust with these audience [4]. The second goal of spammers is to steal important information about the user such as personal information, private photos, password, and credit card details. There are various strategies for information stealing. Phishing is the most popular strategy that is used usually to attack user information through email. Phishers send SMS messages through mobile phones due to their simplicity and impressive usage. The sent message may include a malicious Uniform Resource Locator (URL) that invites the user to visit that address, which is a trick to steal his information. In addition, it is

difficult to identify phishers by their phone numbers, as they can purchase more than one phone number. This provided an environment for researchers to discover and build new applications to filter SMS spam messages [5, 6]. To overcome the SMS spam issue that required to be identified when they are received, this paper proposed a model based on fuzzy rule and particle swarm for classifying a message into ham or spam with minimal set of rules. The paper's contribution is two-fold including:

- Analyzing the spam and ham messages in the adopted dataset and introducing a new feature set extracted from the message content. The extracted feature set should able to represent SMS messages with high discriminative capability.
- Generating fuzzy rules to alleviate the challenges of spam detection and developing a particle swarm optimization (PSO) based rule selection model that is able to remove the irrelevant and redundant rules to detect SMS effectively with high performance.

The rest of this paper is organized as follows. Section 2 presents the related work, then Section 3 elaborates on an introduction for preliminary concepts related to fuzzy logic. In Section 4, the proposed SMS method is introduced. The results and discussion of the proposed method are clarified in Section 5. Finally, Section 6 provides the conclusion and recommendations for future research.

2. Related work

There are several approaches for detecting SMS spam that have been classified into three categorize. First, the content-based approach includes the usage of word frequency. Second one is a non-content based approach that uses the characteristics of a particular message. A third category is a hybrid approach that merges features from the first and second categories for classification purposes [6]. The research method in this study was based on a content-based approach. Some of the previous work based on the content are: Jali (2016) carried out an analysis of the ability to control features, analyzing information, and affect circumstances in the classification of SMS spam messages [7]. Kaya and Ertuğrul (2016) have implemented a method based on local ternary patterns to extract two distinct features from SMS messages and many machine learning approaches have been applied to distinguish SMS spam [8]. An approach that can detect spam messages using 10 features and 5 machine learning algorithms namely: naive Bayes, logistic regression, J48, decision table, and random forest was proposed by Choudhary and Jain (2017). As an evaluation dataset, SMS Spam Corpus was used. The results demonstrated that the random forest gives the best detection rate equivalent to 96.1% [9]. Abdulhamid et al. (2017) introduced a study of potential pathways and obstacles for spam detection and the expected direction of investigation that could enable specialist researchers to consider areas for further improvement [1]. A generic model for classifying SMS spam messages based on many machine learning algorithms was presented by Kaliyar et al. (2018). Filtering messages from many languages including Singapore, American, and Indian English were possible with the presented SMS spam filtering model. The results revealed that when Indian English is used, the model attained a high precision [11]. A comparative study of the impact of feature selection methods on numerous classifiers was provided by Sharaff (2019). The results proved that the feature choice technique influences the classifier's efficiency and can improve specific classifiers [12].

Sjarif (2019) introduced a new method based on computing TF-IDF and many classifier algorithms. The best result was obtained by Random forest had 97.5% accuracy [13]. Kumar (2020) presented a method based on selecting eleven features from the dataset, submitted to a set of classification algorithms to classify them as spam or ham. Besides, another step that filters Smishing message from spam message was added. Random Forest classifier gives the best result with 94.9% accuracy [14]. Arnold and Andrew in 2020 introduced a hybrid model for SMS spam detection in which a genetic algorithm (GA) is utilized for feature selection, and the Bayesian network is adopted for classification. The results clarify that the performance of a blend of the GA and Bayesian network surpasses the performance of the Bayesian network alone [15].

Hameed suggested three models for detecting SMS spam using differential evolution (DE) in 2021. Moreover, a new set of features has been extracted from SMS messages, and three different DE algorithms are used as clustering SMS messages. The results reveal that the suggested models outperformed baseline methods [16].

The models in [1,13,14], which are adopted for comparison with the proposed model are based on the extraction of different features and applying many classification algorithms. The random forest algorithm gives the best results for all. As known, the main drawbacks of the classification algorithms are the large size of training data and the need for retraining for new data. Also, random forest suffers from a large number of trees that make the algorithm too slow and ineffective for real-time prediction. Therefore, this study proposed a model for

classifying SMS using fuzzy rule-based classification that does not involve retraining new data. Furthermore, this study includes the extraction of thematic of SMS spam feature that is not studied in the literature and affects the detection of SMS spam.

3. Preliminary concepts

In this section, the related backgrounds to fuzzy logic and particle swarm optimization are described.

3.1 Fuzzy logic

The development of Fuzzy sets theory produced by Zadeh in 1965 led to introducing the term "fuzzy logic." Fuzzy logic is a way to deal with the degree of truth instead of a standard "True or false"(0 or 1) that is commonly used in most computer systems. It is the generalization of the classical logic inference rules that can deal with approximate reasoning [17]. Each member of fuzzy logic has a degree of membership that is specified by the membership function. The membership function gives a degree ranging from zero to one for each fuzzy logic member. There are many advantages of fuzzy logic, including handling imprecise input, dealing with conflicting objectives, simply understanding the rules, and easily modifying or deleting existing rules [18].

3.2 Particle swarm optimization

Particle Swarm Optimization (PSO) is one of the metaheuristics algorithms introduced in 1995 [19]. The basic idea of PSO is inspired by simulating the behaviour of birds and fish. PSO has been widely used in various applications such as function optimization, pattern classification, and fuzzy system control [20]. The basic idea that the birds search for food either separately or together before finding it. The bird that is close to the food or who senses the smell of food will leak this information to the flock; in the same way, the PSO algorithm works. The birds' movement is equal to finding a new solution, useful information about the food source equal to the most optimal solution, and the food resources are equal to the most optimal solution through all iterations [21]. PSO is a population-based algorithm. The population (swarm) of PSO consists of a set of particles. Each particle's new location is determined by a velocity representing the particle's direction and distance at the t^{th} iteration. It depends on the previous best position found by the particle itself, while the position of each particle is used to measure its quality at iteration t [21].

4. Methodology for SMS spam detection

This section introduces the proposed method for encountering a SMS spam. Consider an SMS collection $\mathbb{M}$ of n text messages, i.e. $\mathbb{M} = \{m_1, \dots, m_n\}$. Each message, m_i, $1 \leq i \leq n$ contains words, numbers, special characters, etc. and its size is restricted to 160 characters. In the proposed method, the messages are pre-processed and then a set of features is extracted. After that, the feature values are converted to a fuzzy set, and several fuzzy rules are generated for prediction SMS spam. Finally, PSO is adopted to produce the best rules. The first endeavor to use binary PSO-based fuzzy rules selection to detect SMS spam to the best of authors' knowledge. The following subsections illustrate the details of each component.

4.1 Pre-processing and feature extraction

In the pre-processing, the tokenization, removing stop and stemming word processes are applied. After the pre-processing, the feature extraction process is performed. The feature extraction process plays a critical role because it affects the efficiency of classifiers for SMS spam detection. In this paper, a set of six features $F = \{f1, f2, \dots, f6\}$ is extracted from $\mathbb{M}$ as illustrated on what follows:

1. Message length ratio, as expressed in Eq. (1), is calculated as the number of characters in the message divided by the maximum length of the message in the SMS collection $\mathbb{M}$

$$f1 = \frac{l}{max_l} \tag{1}$$

Where
l represents the length of the message
max_l is the maximum length of the message in M

2. Number of words ratio, as expressed in Eq. (2) ,is calculated by dividing the number of words in the message by the largest number of words in the message in the SMS collection $\mathbb{M}$

$$f2 = \frac{w}{max_w} \tag{2}$$

Where
w represents the number of words in the message
max_w is the maximum number of words in the message in $\mathbb{M}$.

3. Number of words that have less than three characters ratio, as expressed in Eq. (3), is calculated as number of words within length less

than three over the largest number of words in the message in the SMS collection $\mathbb{M}$.

$$f3 = \frac{N_W}{max_w} \quad (3)$$

Where
N_W is the number of words of length less than three characters in the message.

4. Capital word ratio, as expressed in Eq. (4), is calculated by dividing the number of capitalized words in the message over the largest number of words in the message in the SMS collection $\mathbb{M}$

$$f4 = \frac{C_N}{max_w} \quad (4)$$

Where
C_N is the number of capital words in the message.

5. Alphanumeric characters ratio, as expressed in Eq. (5), is calculated as the number of alphanumeric characters in the message over the largest number of words in the message in the SMS collection $\mathbb{M}$

$$f5 = \frac{A_N}{max_l} \quad (5)$$

Where
A_N represents the number of alphanumeric characters in the message.

6. Thematic SMS spam words: this factor composite of many features. First feature special characters such as "+, =, etc.", the second feature is important words such as "call, money, mobile, phone" afterword the existence of URL. All of these features are computed for each message. Then the number of thematic SMS spam words divided by the largest number of thematic SMS in the SMS collection $\mathbb{M}$ is calculated to obtain thematic SMS spam words ratio as expressed in Eq. (6).

$$f6 = \frac{t_w}{max_{tw}} \quad (6)$$

Where
t_w represents the number of the thematic words in the message.
max_{tw} is the maximum number of thematic words in the message in $\mathbb{M}$

4.2 Fuzzy rule-based classification

The messages presented in the SMS collection are assigned by a feature score using the aforesaid six features. The computed feature score, F, is introduced as an input to the fuzzy logic. Then, the numerical value of each feature is transformed in terms of its linguistic variable by the fuzzier into three fuzzy sets: low (L), Medium (M) and High (H). The triangle membership function as defined in Eq. 7 is adopted in this paper to convert each feature score to a degree of membership between zero and one [22].

$$f(x: a, b, c) = \begin{cases} 0 & if\ x \le a \\ \frac{x-a}{b-a} & if\ a \le x \le b \\ \frac{c-x}{c-b} & if\ b \le x \le c \\ 0 & if\ x \ge c \end{cases} \quad (7)$$

When the inputs have been converted to the linguistic values, then, the inference engine which is regarded as an important part of fuzzy system generates if-then rules. The constructed rules consist of two parts antecedent and consequent. Antecedent is the independent input linguistic values, while the consequent is the inference of the rule that specifies the message as either spam or ham. The fuzzy rule is summarized in the following with $F = \{f_1, f_2, \dots, f_6\}$ as input whereas $O \in \{Spam, Ham\}$ as the output.
If f_1 is A and f_2 is A … and f_6 is A Then O
Where
A $is\ L, M, or\ H$

4.3 Fuzzy rules selection by binary PSO

The constructed rules are useful for SMS spam detection, but there are too many rules generated by the proposed fuzzy logic method that may impact the performance of the message classification, so in this paper, a new method based on binary PSO is proposed to select a subset of rules, R′, from the entire set of rules, R, that satisfy decreasing the number of generated rules and get the optimized rules for improving classification performance.

4.3.1. Particle representation and swarm initialization

Representation of the particle in the proposed PSO is binary, since there are six inputs each has three memberships, so each particle I is represented as a bit string of length 18. Every 3 bits corresponds to one feature. The value of 3 bits indicates the membership L, M or H. Fig. 1 shows the representation of the particle and how each membership is selected. In addition, each particle has

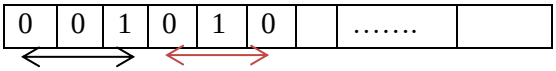

Figure. 1 Representation of the particle in PSO

a velocity V.Formally speaking I can be expressed as follows

$I = \{x_1, x_2, \dots x_{18}\}$

$x_i \in \{0,1\}$

$V = \{v_1, v_2, \dots v_{18}\}$

PSO is a population-based optimization algorithm. A swarm S of N particles is generated randomly, which can be formally expressed as follows

$S = \{I_1, I_2, \dots, I_N\}$

Also, the velocity of each particle in the initial swarm is generated according to Eq. (8).

$$V = V_{min} + [V_{min} + V_{max}] \times r \quad (8)$$

Where

r is a random variables between 0 and 1.

V_{min} and V_{max} are the minimum and maximum velocity the particle can reach.

4.3.2 Fitness function evaluation

In PSO, the fitness function assesses the quality of the particle I and gets optimal rules.

Considering the SMS spam detection problem, the fitness function is computed by the summation of the values of the features as in Eq. (9).

$$Maximize\ \ \Phi(I) = \sum_{i=1}^{6} f_i \quad (9)$$

Where

f_i is the value of the i^{th}feature.

4.3.3 Updating particle position and velocity

There is a specific velocity of every particle that can be modified according to it is own flying experience. The best obtained position is kept in the memory and sent out for all particles. The velocity of particle I is updated according to Eq. (10).

$\forall i \in \{1,2..18\}$

$$v_i^t = w \times v_i^{t-1} + C_1 \times r_1(L_i^{t-1} - I_i^{t-1}) + C_2 \times r_2(G_i^{t-1} - I_i^{t-1}) \quad (10)$$

Where

r_1 and r_2 are two random numbers in [0,1].

w is the inertia weight.

C_1 is cognitive parameter.

C_2 is social parameter.

L^{t-1} is the best position being located locally by particle I_i

G^{t-1} is the best position among all particles in swarm S.

Then the piecewise function to restrict the velocity in range of $[V_{min}, V_{max}]$. After obtaining the velocity, the sigmoid function, as expressed in Eq. (11) is used to limit the result between 0 and 1.

$$s = \frac{1}{1+e^{-v_i^t}} \quad (11)$$

The particle I at iteration t can be updated according to Eq. (12).

$$\forall i \in \{1,2..18\}$$
$$x_i^t = \begin{cases} 1, rand < s \\ 0, other\ wise \end{cases} \quad (12)$$

Where

$rand$ is a random number between 0 and 1.

5. Experimental results and discussion

SMS collection UCI dataset from the National University of Singapore, and Caroline Tagg's Ph.D. thesis [23, 24] is used in this paper for performance evaluation. The SMS spam dataset consists of 5574 messages. Each message is categorized as either spam or ham. The number of ham and spam messages is summarized in Table 1.

The result of the proposed model was evaluated using well-known standard evaluation metrics for the classification algorithms namely, detection rate (recall) (R), precision (P) , F-measure (F), and accuracy (Acc) as calculated in Eqs. (13) - (16) [25-27]. Moreover, the Receiver Operating Characteristic (ROC) curve is used to check the performance of the proposed model cost against recall.

$$R = \frac{TP}{TP+FN} \quad (13)$$

$$P = \frac{TP}{TP+FP} \quad (14)$$

$$F - measure = \frac{2 \times P \times R}{P+R} \quad (15)$$

$$Acc = \frac{TN+TP}{TP+TN+FN+FP} \quad (16)$$

$$FNR = \frac{FN}{TP+FN} \quad (17)$$

Where

TP is a SMS spam that correctly classified as a spam.

Table 1. The number and percentage of ham and spam in SMS spam dataset

Message Type	Number of SMS	Percentage of SMS
Ham	4827	86.6%
Spam	747	13.4%

Table 2. Parameters setting.

Parameter	Value
c_1	1.495
c_2	1.495
w	0.728
V_{min}	-4
V_{max}	4
Maximum iteration	100

TN is a ham that correctly classified as a ham.

FN is a SMS spam that incorrectly classified as a ham.

Several parameters involving the characteristics of the proposed binary PSO based fuzzy rule selection model should be fixed to quantified values. The values of the parameters are reported in Table 2.

As mentioned earlier, the two objectives of the proposed model are the extraction of the important features which is a challenging task and the discrimination of the input message as spam or ham. Many steps were performed: initially, the feature vector was created by extracting six features from the dataset and a linguistic score is assigned to each feature vector. Finally, the generated linguistic vector is introduced to a set of fuzzy rules and a binary PSO is proposed for discovering the most useful rules. The following subsections clarify the results of the proposed model.

5.1 Feature result analysis

The SMS spam dataset contains a set of messages the categorized as ham and spam. The distribution of some common words appearing in ham and SMS spam is depicted in Fig. 2. As shown in the figure there is an interleaving between SMS spam and ham words, which makes the distinction between them difficult.

Moreover, to illustrate the significance of each feature, the weight of the selected features in ham and spam messages is determined. Concerning the imbalanced dataset, each function was measured as an average for all datasets. As shown in Fig. 3, thematic SMS spam feature has the highest weight among all feature sets. This feature is new since it is a combination of features that were used separately in previous researches. Combining these features together provide greater power to differentiate between SMS spam and ham. It has been observed to

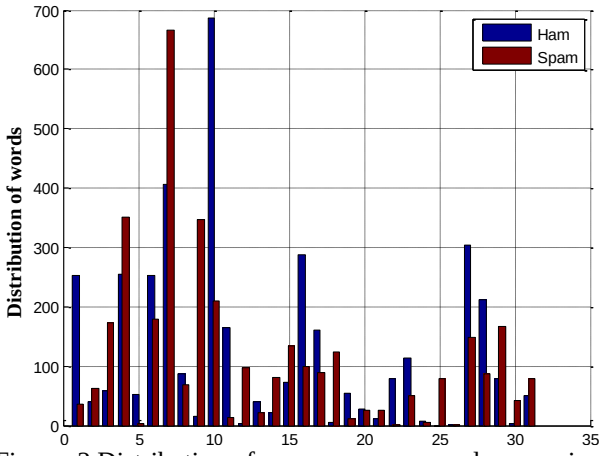

Figure. 2 Distribution of some common words appearing in ham and spam messages

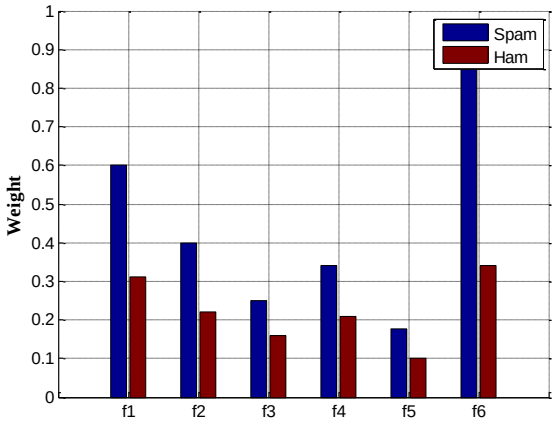

Figure. 3 The weights of features in spam and ham messages

be less important when it is alone as has been shown in some previous research. The importance of other features is as follows message length, word numbers, upper case word, word less than three and alphanumeric characters.

5.2 Performance comparison against other methods

Fig. 4 depicts the proposed model evaluation regarding ROC where Area Under Curve (AUC) equals 0.9863. The figure illustrates the ability of the proposed model to distinguish between spam and ham messages.

As shown in Table 3, the proposed model achieves the highest performance and outperforms the works presented in [1, 13-14] especially in terms of accuracy, which is the most important factor in measuring the performance of the classification. The superior performance of the proposed model is due to the inclusion of the new feature namely thematic SMS spam. The new feature has the highest weight among other features exist in previous work as shown

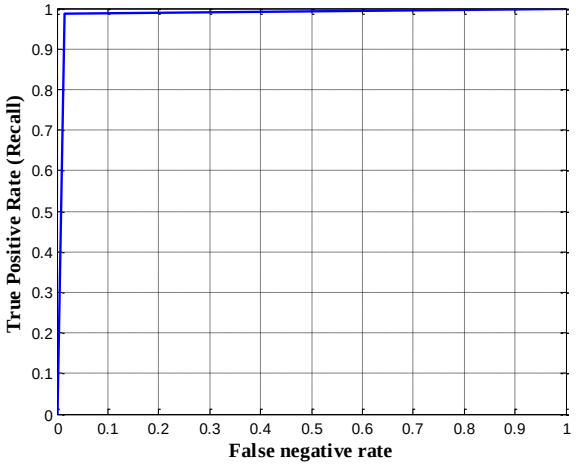

Figure. 4 ROC for the proposed model

Table 3. Comparisons among the proposed model and previous researches

Model	R	P	F-Measure	Acc
Proposed method	**98.8%**	90.8%	94.6%	**98.5%**
Ref. [1]	86%	96%	91%	97.7%
Ref. [13]	97%	98%	97%	97.5%
Ref. [14]	97.7%	95.9%	96.7%	94.9%

in Fig. 2. The proposed model is based on fuzzy rules for classification purposes that strengthen the extraction of the most relevant SMS spam features and correctly classify the SMS spam.

In addition, it can be observed the good performance of the proposed model based on fuzzy logic and PSO for spam detection problem. The obtained result is very efficient for the spam class and less efficient for the ham class based on two reasons: firstly, the fact that the selected dataset was imbalanced where the number of ham messages more than spam messages six times approximately. Secondly, the set of selected features and the fitness function both work towards a good definition of spam messages, thus the generated rules are more efficient in identifying spam messages more than ham messages.

6. Conclusion

SMS spam messages are unwanted messages that must be detected before the user gets them. This paper focuses on introducing a new model for detecting SMS spam messages based on the message content, and it is the first research taking advantage of the fuzzy rule to detect SMS spam. In the proposed model, a new set of relevant features are extracted from the SMS spam dataset. Then, fuzzy rules are generated for discriminating SMS spam from ham. Finally, a binary swarm optimization is introduced to pick the most relevant fuzzy rules. The SMS spam dataset is used as an evaluation dataset. The proposed model achieved 98.8% recall, 90.8% precision, 94.6% F-measure, 98.5% accuracy and 0.9863 AUC. The works presented in [1, 13-14] used commonly method for classification. However, the proposed model adopted a fuzzy based model for generating the rules for classification and introduced PSO for selecting the relevant fuzzy rules for improving the classification performance. The results of the evaluation reveal that the extracted features have a significant relation with the message class. Furthermore, the proposed model efficiently distinguishes the class of messages with a high detection rate and accuracy compared against [1, 13-14]. For future work, the proposed model can be used to detect SMS spam messages written in other languages and can be applied to other datasets containing more messages.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization, Sarab M. Hameed and Zuhair Hussein Ali; methodology and implementation, Sarab M. Hameed and Zuhair Hussein Ali; writing original draft preparation, Sarab M. Hameed and Zuhair Hussein Ali; writing, review and editing, Sarab M. Hameed and Zuhair Hussein Ali; supervision and funding acquisition, Sarab M. Hameed and Zuhair Hussein Ali.

References

[1] D. Suleiman, and G. Al-Naymat, “SMS Spam Detection Using H2O Framework”, *Procedia Computer Science*, Vol. 113, pp. 154–161, 2017.

[2] P. Roy, J. Singh, and S. Banerjee, “Deep Learning to Filter SMS Spam”, *Future Generation Computer Systems*, Vol. 2, pp. 524–533, 2020.

[3] T. Xia and X. Chen, “A Discrete Hidden Markov Model for SMS Spam Detection”, *Applied Science*, Vol. 10, 2020.

[4] D. Drossos, G. Giaglis, and G. Lekakos, “Determinants of Effective SMS Advertising: An experimental Study”, *Journal of Interactive Advertising*, Vol. 7, No. 2, pp. 16-27, 2007.

[5] S. Sheikhi, M. T. Kheirabadi, and A. Bazzazi, “An Effective Model for SMS Spam Detection Using Content-based Features and Averaged Neural Network”, *International Journal of Engineering*, Vol. 33, No. 2, pp. 221-228, 2020.

[6] O. Abayomi-Alli, S. Misra, A. Abayomi-Alli, and M. Odusami, "A Review of Soft Techniques for SMS Spam Classification: Methods, Approaches and Applications Engineering", *Applications of Artificial Intelligence*, Vol. 86, pp. 197–212, 2019.

[7] K. Zainal and M. Z. Jali, "A Review of Feature Extraction Optimization in SMS Spam Messages Classification", *Berry M., Hj. Mohamed A., Yap B. (Eds) Soft Computing in Data Science. SCDS. Communications in Computer and Information Science*, Vol. 652, pp. 158-170 2016.

[8] Y. Kaya and O. F. Ertuğrul, "A Novel Feature Extraction Approach in SMS Spam filtering for Mobile Communication: One-Dimensional Ternary Patterns", *Security and Communication Networks*, 2016

[9] N. Choudhary and A. K. Jain, "Towards Filtering of SMS Spam Messages Using Machine Learning Based Technique", In: Singh D., Raman B., Luhach A., Lingras P. (eds) Advanced Informatics for Computing Research. *Communications in Computer and Information Science*, Vol. 712, pp. 18-30, 2017.

[10] S. M. Abdulhamid, M. S. hafie Abd Latiff, H. Chiroma, O. Osho, G. Abdul-Salaam, A. I. Abubakar, and T. Herawan, "A Review on Mobile SMS Spam Filtering Techniques", *IEEE Access*, Vol. 5, pp. 15650-15666, 2017.

[11] R. K. Kaliyar, P. Narang, and A. Goswami, "SMS Spam Filtering on Multiple Background Datasets Using Machine Learning Techniques: A Novel Approach", In: *Proc. of International Conf. On Advance Computing*, Greater Noida, India, pp. 59-65, 2018.

[12] A. Sharaff, "Spam Detection in SMS Based on Feature Selection Techniques", *Abraham A., Dutta P., Mandal J., Bhattacharya A., Dutta S. (eds) Emerging Technologies in Data Mining and Information Security. Advances in Intelligent Systems and Computing*, Vol. 813, pp. 555-563, Springer, 2019.

[13] N. N. Sjarif, N. F. Mohd Azmi, S. Chuprat, H. Sarkan, Y. Yahya, S. M. Sam Razak, "SMS Spam Message Detection using Term Frequency-Inverse Document Frequency and Random Forest Algorithm", *Procedia Computer Science*, Vol. 161, pp. 509–515, 2019.

[14] A. Kumar, "A Novel Approach to Detect Spam and Smishing SMS Using Machine Learning Techniques", *International Journal of E-Services and Mobile Applications*, Vol. 12, No. 1, 2020.

[15] A. A. Ojugo and A. O. Eboka, "Memetic Algorithm for Short Messaging Service Spam Filter Using Text Normalization and Semantic Approach", *International Journal of Informatics and Communication Technology*, Vol. 9, No.1, pp. 9-18, 2020,

[16] S. M. Hameed, "Differential Evolution Detection Models for SMS Spam", *International Journal of Electrical and Computer Engineering*, Vol. 11, No. 1, pp. 596-601, 2021,

[17] B. Barnabas, "Mathematics of fuzzy sets and fuzzy logic", 2013.

[18] S. Waris and Z. Ahmad, "Application of Fuzzy Logic", In: *Proc. of International Conf. on Recent Advances in Statistics*, Lahore, Pakistan, pp. 367-376, 2011.

[19] J Kennedy, Eberhart R., "Particle Swarm Optimization", In: *Proc. of IEEE International Conf. on Neural Network*, Perth, Australia, pp. 1942-1948, 1995.

[20] M. R. Bonyadi and Z. Michalewicz, and X. Li, "An Analysis of the Velocity Updating Rule of the Particle Swarm Optimization Algorithm", *Journal of Heuristics*, Vol. 20, pp. 417–452, 2014.

[21] Q. Bai, "Analysis of Particle Swarm Optimization Algorithm", *Computer and Information Science*, Vol. 3, No. 1, 2010.

[22] T. Hong, and C. Lee. "Induction of Fuzzy Rules and Membership Functions from Training Examples", *Fuzzy Sets and Systems*, Vol. 84, No. 1, pp. 33-47, 1996.

[23] T. A. Almeida, G. Hidalgo and J. M., Yamakami, "Contributions to the Study of SMS Spam filtering: New Collection and Results. In: *Proc. of ACM Symposium on Document Engineering*, pp. 259-262, 2011.

[24] T. A., Almeida, J. M., Gómez Hidalgo, T. P. Silva, "Towards SMS Spam Filtering: Results under a New Dataset", *International Journal of Information Security Science*, Vol. 2, No. 1, pp. 1-18, 2013.

[25] K. S. Reddy and E. S. Reddy, "Integrated Approach to Detect Spam in Social Media Networks Using Hybrid Features", *International Journal of Electrical and Computer Engineering*, Vol. 9, No. 1, pp. 562-569, 2019.

[26] A. Boukhalfa, A. Abdellaoui, N. Hmina, H. Chaoui, "LSTM Deep Learning Method for Network Intrusion Detection System", *International Journal of Electrical and Computer Engineering*, Vol. 10, No. 3, pp. 3315-3322, 2020.

[27] S. M. Hameed, M. B. Mohammed, B. A. Attea, "Fuzzy Based Spam Filtering", *Iraqi Journal of Science*, Vol. 56, No. 1B, pp. 506-519, 2015.