Palestine Journal of Mathematics

Vol. 11(Special Issue I)(2022) , 171–183

HYBRID APPROACHES FOR SPAM REVIEW DETECTION: A REVIEW

Sakshi Shringi and Harish Sharma

Communicated by Shimpi Singh Jadon

MSC 2010 Classifications: Primary 20M99, 13F10; Secondary 13A15, 13M05.

Keywords and phrases: Fake reviews, Spam detection methodologies, Machine learning techniques, Nature inspired algorithms, Social network.

Abstract Over the last few years, spam has infiltrated all modes of digital communication. With the rapid increase in the use of social media platforms such as Facebook, YouTube, and Twitter, etc., a huge amount of spam is generated, providing a new path for spammers to exploit these platforms. Through social media platforms customer's express their opinions in the form of online reviews which helps in making business decisions and product purchases. However, to attain profit, some of these reviews may be spam, resulting in high publicity of unworthy products. Hence, developing techniques that help to differentiate between spam and non-spam is a challenging task. In this paper, we have presented a study which focuses on the comprehensive analysis of recent developments in the field of spam detection. The methods illustrated in this study uses hybrid approach for detection of spams and are assessed based on the accuracy and results.

1 Introduction

A Spam is undesirable or unsolicited messages acquired electronically via email, messages, social networks, internet search with the intent of advertising, fraudulence, proliferating virus etc. The person involved in sending such messages is usually termed as "spammer". The spammers generate such messages for their personal profits or for any organization. Jindal et al. [19] categorized online reviews into the following:

(i) Untruthful reviews: The reviews which purposely deceive readers or review mining systems by writing unworthy positive reviews for a specific target objects for false promotions, also known as *hyper spam*, on the other handwriting negative reviews for some other specific objects to deteriorate their image, also known as *defaming spam*.

(ii) Non-reviews: Reviews that contain irrelevant content and commercials.

(iii) Review on brands: These reviews contain majorly focusses on promoting a brand rather than focusing on the product.

Initially, spams were only limited to e-mails, but with the progress of Web 2.0, spam has adequately breached all electronic platforms. The following media is majorly affected by spammers:

- Social Spam: Social networking platforms such as Twitter, Facebook, Foursquare, etc. suffers from different types of spams [18]. These spams can be in the form of fake or untruthful reviews, malicious links, personal data, fake friends, misbehaviour and hateful expressions.
- E-mail Spam: These spams are spontaneous commercial e-mails sent frequently in large amount along with some commercial cotent [6].
- Splog and Wiki Spam: The spams which occurs in blogs are splog spams [47]. These spams refers to the irrelevant comments on any topic of discussion, accompanied by the URL links to few commercial sites. The splogs may be written to promote a website such as verbose ads or they may consist of stolen original data from authentic websites. Attacks of similar nature are experienced by Wikis.

- Newsgroups and Forums Spams: The targets of such spams are Usenet newsgroups [9]. The newsgroups spams can be defined as excessive multiple posting. Publishig of ads irrelevant to the subject of discussion are named forum spams.
- Video Sites Spam: Video sites such as YouTube experiences spams in the form of comments and links to some irrelevant videos.
- Message spams in Online gaming: Regular requests to join a particular group, messages displaying breaching of copyright terms and conditions are considered as spam messages in online gaming.
- Instant messaging spam: The Instant Messengers (MIs) are used for spamming in instant messaging apps such as Yahoo Messenger, Skype tec. in the form of spontaneous messages from advertisements [24].
- Mobile Phone Spams: The mobile phone spams employ Short Messaging Services (SMS) as their tool to generate spam [2]. The user may get trapped in some kind of distorted subscriptions.
- Internet Telephony Spam: This spam is called as Spam over Internet Telephony (SPIT) [35]. For spamming, this uses Voice over Internet Telephony (VoIP).
- Spamdexing: It is a meticulous manipulation of indexes in search engines. also known as search engine spams [34]. This spam generally highlights pages which are less or of no importance.

Spam is inescapable in practically all types of online conversations today and is known to hamper the efficiency of the medium on which it shows up. Different measures have been taken to improve the durability of different online platforms against a variety of spam intrusions, known as anti-spamming approaches. Even though enough work has been done in the filed of spam detection, the current hypothesis of spam detection taction techniques is still not sufficient to identify spam. The continuously emerging, intractable graph of the social media such as web graph are majorly responsible generation of bulk spam [8]. One of the major reasons of spam creation is that the content generated by users on social media is very simplified and does not go through any restraint or control policy. This helps in excessive development of spam. The merchants use these platforms for their personal profits or for brand promotions, resulting in misguiding the users through fake reviews.

Some of the cases where online reviews play a prime role are:

(i) To buy something through an online retail website, both product and seller reviews are crucial.

(ii) To buy a software.

(iii) Making a decision on whether to watch a particular movie or not based on movie reviews.

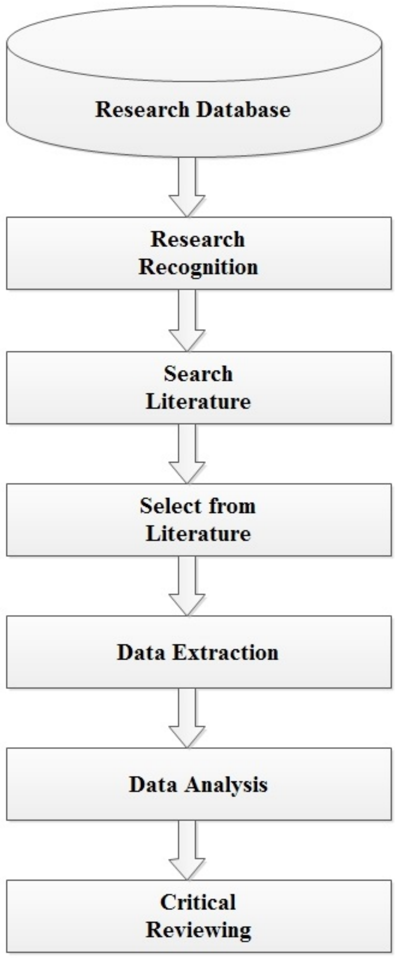
According to a survey, spam generation has increased by 355% in 2013 as many new users have joined social platforms with increasing time. As spam can highly effect the vale of any brand or product, the number of spams generated should be limited, if not eliminated to a certain extent. The social platforms have different characteristic features as compared to other search engines, spam detection becomes more challenging. Various contemporary approaches alongwith the existence one have been applied for spam detection. All these factors formed a basis for us to write this review.

2 Research Methodology

An organized search of relevant journal and conference papers was made inorder to classify the literature concerning spam detection. The search strategy comprises of the following steps:

(i) The search terms were established majorly comprising of "spam review detection methods". Different synonyms and keywords for spam review such as opinion spam, spam detection, fraud review, reviewer spam, and fake review, were used for searching. The keywords were recognized in relevant papers and articles.

Figure 1: Review Process



(ii) Various literature search resources were used for performing search such as Google Scholar, Science Direct, IEEE Explorer, ACM digital library etc.

(iii) The research papers collected were reviewd thoroughly to identify their relevance. Some more related papers were searched using the refernces of the selected papers.

(iv) Finally, all the collected papers were reviewed extensively. Figure 1 illustrates the steps involved in review process.

3 Categories of Social Spam

Due to recent growth in the Internet, the content generated by individuals on social platforms has curbed the content which is generated for professional purpose. This is because the social media provides a mutual platform to people for expressing their viewpoints and opinions. Prominent user specified content is majorly created via the social networking websites such as Twitter, Facebook, MySpace, Linkedin, etc. Other websites as Amazon, Flipkart, BookMyShow, etc. also play a vital role in online reviews. This captivates the vicious people to use such platforms for their personal benefits to promote a particular brand or product by generating fake reviews.

Based on the characteristics, properties and social media platforms used, social spam is of the following types:

(i) Fraud Reviews: The reveiwer writes false comment about a product claiming it to be good, without even using the product or defames a good product, are termed as Fraud Re-

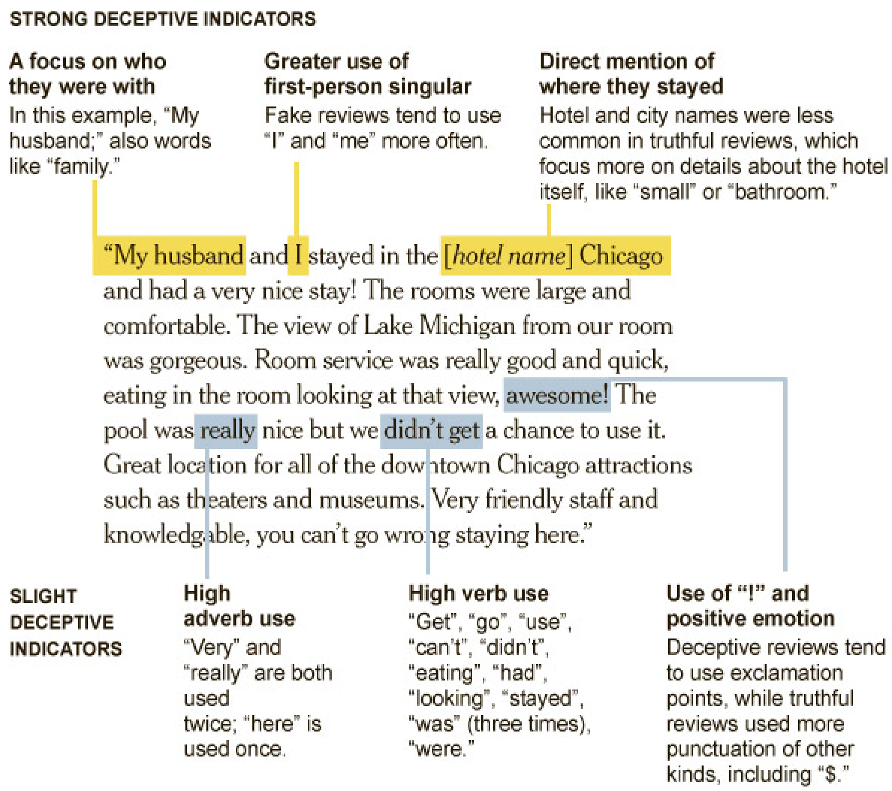


Figure 2: Example of fake review of a hotel

Source: [53]

views.Fraud reviews can be of a product, a hotel review, and a movie review. An exapmle of fake hotel review is illustrated in Figure 2.

(ii) Spurious Profiles: Fake online profiles are created by spammers, which appears authentic to non-spammers like a fake facebook profile, resulting in adding them as friends. Figure 3 illustrates an example of malicious link

(iii) Malevolent links: Figure 4 illustrates an example of malicious link. Such spam links sabotages the users or computers.

(iv) Submissions in bulk: This is also termed as spam bombing, in which mass spams are sent in the form of comments for the same context. An example of Google-Bombing is illustrated in Figure 5. the figure shows how the search query "miserable failure" was linked to George W. Bush and Michal Moore.

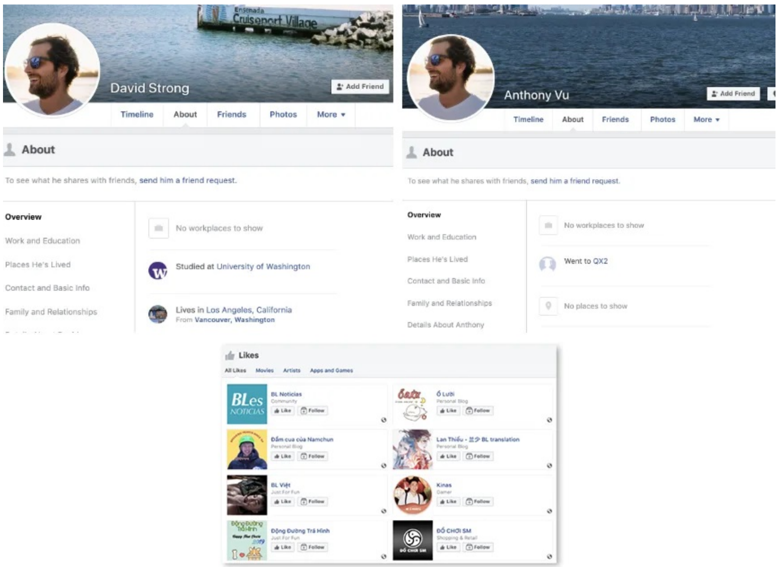
Many other forms of social spams also exist such as obscene words in statements and comments which involves use of some special characters, animosity speech, intimidation and abuse, etc. which are very hard to detect [28].

4 Spam Detection Approaches

The concept of spam is eminently abstract but we can affirm it as something which is undesirable for a valid user. The evolution of use of social networks and their inflexible security policy has lead to spammers in adjusting accordingly. Spam can be detected by suing various approaches such as Machine learning based, Network based, and Pattern minning based.

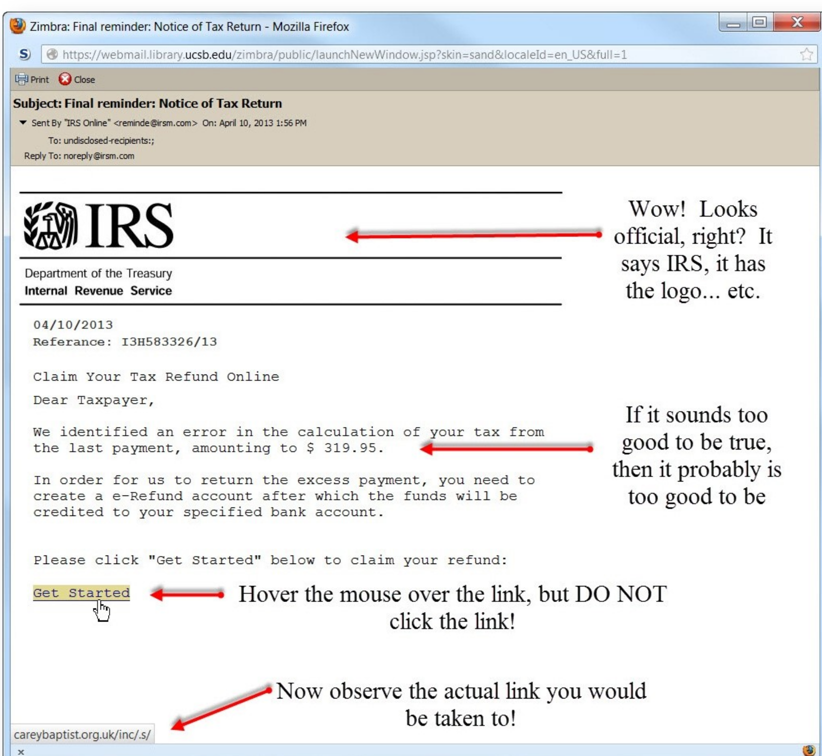
Oda et al. [31] used the Artificial immune system to detect email spams. They implemented their model in Perl due to its considerable adaptability for strings. They used simple text files to stock the lymphocytes and gene library. They attained 90% accuracy with 1000 lymphocytes. Oda et al. [32] extended their own model by using Artificial immune system for spam detection and compared the scoring-schemes, population size effect and the libraries that were used for

Figure 3: Example of a fake facebook profile



Source: [54]

Figure 4: Example of malicious link



Source: [52]

Figure 5: Example of Google bombing

Source: [55]

creation of detectors. They attained 93.6% accuracy with 700 heuristic lymphocytes. Lai et al. [25] used a hybrid approach in which Particle Swarm Optimization is used for feature selection and Support Vector Machine for classification. The proposed system comprised of two modules, training and testing. They acheived 92.7% accuracy with 63 features on spam-assassin corpus having 3002 emails out of which 501 were marked as spam and 2501 were ham. Abi et al. [12] proposed a model based on cross-regulation which was inspired by adaptive immune system. They tested their model on six e-mail datasets and acheived an average accuracy of 89% with the varying ratio of timestamped spam and non-spam emails. They compared their model with Naive Bayes and other classification models.

Yin et al. [49] used LDA and Ant colony algorithm [10] for detection of spam mails. They acheived 96.83% precision and 90.25% recall on Lingspam corpus consisting of 2893 emails out of which 481 were labeled as spam and 2412 were labeled as non-spam. Their experimented results outperforms other spam filtering methods. Ruan et al. [38] used Back Propagation Neural Network with two inputs for classification of emails. They generated the two inputs by using Concentration Based Feature Construction in which 'self' and 'non-self' concentrations are constructed through 'self' and 'non-self' gene libraries. Their model acheived 97% and 99% accuracy on PU1 and Ling corpus respectively by just using a two-element concentration feature vector. Mohammad et al. [27] deployed Artificial Immune System with Genetic Algorithm for optimization of spam detectors to find the time of culling and checking if self has changed, and used only Artificial Neural Network to detect spam. Their results showed 3.741% false negative with 600 lymphocytes in Artificial Immune System optimized with Genetic Algorithm and 3.668% false negative with 300 neurons in Artificial Neural Network on SpamAssassin corpus containining 5911 emails out of which 1764 were marked as spam and 4147 were marked as non-spam.

Salehi et al. [40] used a simple hybrid Artificial Immune System with Particle Swarm Optimization using mutation for optimization. They applied 20 runs on datasets for every threshold bout and acheived an accuracy of 88.33%. M Mahmoud et al. [26] used Artificial Immune System. They resulted an average accuracy of 91% on 1324 SMS messages out of which 1002 were non-spam messages collected from NUS SMS Corpus and Jon Stevenson Corpus, and 322 spam messages collected from Grumbletext mobile spam site. Natrajan et al. [29] used Enhanced Cuckoo Search to optimize bloom filter using total membership invalidation cost as the objective function and outperforms the Cuckoo Search Algorithm for all string sizes. He et al. [13] implemented local concentration for feature selection with firework algorithm with 10 cross validation for optimization and Support Vector Machine on selected features for classification. The experiment results determines that the model used improves the performs on the corpora and acheives 98.57% accuracy on 1099 emails out of which 481 were labeled as spam. Yevseyeva et al. [48] proposed a method to solve the problem of anti-spam filtering scores optimization. They optimized Grindstone 4SPAM, NSGA-II and SPEA2 anti-spam filters using Evolutionary Algorithm [42]. Idris et al. [14] proposed a method and attained 69.76% accuracy at 1000 generated detectors with threshold value of 0.4 by applying Differential Evaluation to optimize Negative Selection Algorithm by using local outlier factor as fitness function. As future work they proposed to develop a hybrid model which uses two evolutionary algorithms for parallel hybridization. Jain et al. [] used parallelly Support Vector Machine and Artificial Immune System for classification. They attained 98.3% accuracy on benchmark corpora PUA with 1142 messages, using both the classifiers. Zhang et al. [51] used wrapper based feature selection using Particle Swarm Optimization with mutation using cost derived from C4.5 Decision Tree as objective function and C4.5 Decision Tree as clasifier over selected features. The accuracy reported by this method was 94.27% accuracy on UCI database with 6000 samples. Rajamohana et al. [36] used an Adaptive Binary Flower pollination algorithm for feature selection using Naive Bayes classifier's accuracy as the objective function and k-nearest neighbors as the classifier using selected features. More than 85% accuracy was observed for 1600 reviews from the 20 most popular Chicago hotels.

Aswani et al. [5] used k-Means deploying LFA with chaos, LFA without chaos, FA with chaos, FA without chaos for tuning either the Absorption Coefficient(μ) or the Attractiveness Coefficient(α). They also implemented fuzzy C-Means to identify any overlapping among the two spam and fuzzy groups. 97.98% accuracy with k-Means with LFA with chaos for tuning was acheived. Ratnoo et al. [37] proposed a hybrid instance feature selection; HIFS-CHC method using heterogeneous recombination and cataclysmic mutation; CHC adaptive search genetic algorithm to solve the problem of dual selection. Singh et al. [45] used correlation based Feature Selection with Particle Swarm Optimization for feature selection with 5 classifiers namely Naive Bayes, J48, AdaBoost, Support Vector Machine, Multi Layer Perceptron. The proposed feature selection method improves the F-score of Support Vector Machine by 45.83%, AdaBoost by 33.02%,vMulti Layer Perceptron by 10.38%, J48 by 9.54%. Pandey et al. [33] adopted spiral Cuckoo search to optimize k-Means algorithm using sum squared error as the objective function. They tested their model on spam review, synthetic spam review, yelp hotel review, yelp resturant review, twitter spam dataset with 64.82%, 71.63%, 70.92%, 71.42% and 97.93% average accuracy respectively.

Ngo et al. [30] proposed a hybrid time series forecast model namely a moving-window firefly algorithm (FA)-based least squares support vector regression (MFA-LSSVR), which captures patterns of historical data and predicts future values of time series data while the FA is used to optimise the LSSVR's parameters to improve the predictive accuracy. Asha Kumari and Balkishan [22] proposed an ant colony optimisation based system for threatening account detection (ACOTAD). Kaur and Chahal [20] proposed a ANFIS-GA based forecasting model for the prediction of Cholera virus. They used non-dominated sorting genetic algorithm (NSGA) is used to tune hyper-parameters of ANFIS. Thepade et al. [46] used Thepade's Sorted Block Truncation Coding N-ary (TSBTC N-ary) for face feature extraction and further deploys machine learning classifiers to identify face as male or female. Sharma et al. [41] developed a local search strategy inspired by dung beetle orientation and foraging activity to intensify exploitation concept of ABC and amalgamated this strategy with ABC. Kumar Sunil et al. [23] presented a detail study of different text mining applications in the field of service and management. They have majorly focused on online reviews and social media data for their research. Kushwaha et al. [21] demonstrated a survey on strategies for data-driven decisions using the past 10 years papers.

Many other hybrid conventional and recent approaches were proposed to detect the sapm reviews [1, 3]. Table 1 illustrates some of the papers identified for spam detection in various categories such as E-mail Spam, Social Media Marketing Spam, SMS Spam, Spam Reviews, and Web Spamming. These methods show various hybrid approaches used for spam detection. May other machine learning and optimization techniques can also be used for identification of spam and increasing the accuracy of current available approaches.

Table 1: Spam Detection Approaches

Author (Year)	Methodology	Results
Oda et al.[31]	Artificial Immune System	90% accuracy with 1000 lymphocytes
Oda et al. [32]	Artificial Immune System	93.6% accuracy with 700 heuristic lymphocytes
Lai et al. [25]	Particle Swarm Optimization is used for feature selection and Support Vector Machine for classification	92.7% accuracy with 63 features on spam-assassin corpus having 3002 emails out of which 501 were marked as spam and 2501 were ham
Abi-Haidar et al. [12]	Immune cross-regulation model inspired by immune system	Average accuracy of 89% on six different datasets with varying ratio of timestamped ham and spam emails.
Yin et al.[49]	Linear Discriminant Analysis for feature reduction and Ant Colony Optimization algorithm with F1 value to calculate inverse of distance between cities which is in turn used for transaction probability to classify the emails in spam and ham	96.83% precision and 90.25% recall on Lingspam corpus which contain 2893 emails out of which 481 were labeled as spam and 2412 were labeled as ham
Ruan et al. [38]	Back Propagation Neural Network with two inputs was used to classify emails. These two inputs were generated by using Concentration Based Feature Construction in which 'self' and 'non-self' concentrations are constructed through 'self' and 'non-self' gene libraries.	97% and 99% accuracy on PU1 and Ling corpus respectively
Mohammad et al.[50]	Artificial Immune System with Genetic Algorithm to optimize spam detectors in finding out the time of culling and checking if self has changed, and using only Artificial Neural Network to detect spam	3.741% false negative with 600 lymphocytes in Artificial Immune System optimized with Genetic Algorithm and 3.668% false negative with 300 neurons in Artificial Neural Network on SpamAssassin corpus containining 5911 emails out of which 1764 were marked as spam and 4147 were marked as ham
Salehi et al. [40]	Hybrid Simple Artificial Immune System with Particle Swarm Optimization using mutation for optimization	88.33% accuracy
Natarajan et al. [29]	Enhanced Cuckoo Search to optimize bloom filter using total membership invalidation cost as the objective function	Comparing performance of Enhanced Cuckoo Search and Cuckoo Search with 10 nests, 50 iterations, pa = 0.3. ECS outperform CS for all string sizes
Mahmoud et al. [26]	Artificial Immune System	Average accuracy of 91% on 1324 SMS messages out of which 1002 were non-spam messages collected from NUS SMS Corpus and Jon Stevenson Corpus, and 322 spam messages collected from Grumbletext mobile spam site

Continued on next page

Table 1 – continued from previous page

Author	Methodology	Results
He et al. [13]	Local concentration for feature selection with firework algorithm with 10 cross validation for optimization and Support Vector Machine on selected features for classification.	98.57% accuracy on 1099 emails out of which 481 were labeled as spam
Yevseyeva et al. [48]	Optimized Grindstone 4SPAM, NSGA-II and SPEA2 anti-spam filters using Evolutionary Algorithm	99.36% accuracy from Grindstone 4SPAM, 99.45% accuracy from NSGA-II, and 99.41% accuracy from SPEA2 on SpamAssassin corpus containing 9349 samples out of which 2398 were labeled as spam and 6951 were labeled as ham
Idris et al. [15]	Differential Evaluation to optimize Negative Selection Algorithm by using local outlier factor as fitness function	69.76% accuracy at 1000 generated detectors with threshold value of 0.4
Idris et al. [14]	Particle Swarm Optimization to optimize Negative Selection Algorithm using local outlier factors as the fitness function	91.22% accuracy at 5000 generated detectors with threshold value of 0.4o
Jain et al. [16]	Support Vector Machine and Artificial Immune System are used parallelly for classification. The end result is calculated using both the classifier	98.3% accuracy on benchmark corpora PUA with 1142 messages
Zhang et al. [51]	Wrapper based feature selection using Particle Swarm Optimization with mutation using cost derived from C4.5 Decision Tree as objective function and C4.5 Decision Tree as clasifier over selected features	94.27% accuracy on UCI database with 6000 samples
Faris et al. [11]	Wrapper based method including Particle Swarm Optimization and RF for feature selection and then RF on selected features for classification	98.16% accuracy when features were selected using RMSE as the objective function for the wrapper based method
Zavvar et al. [50]	Artificial Neural Network with Particle Swarm Optimization for feature selection and Support Vector Machine for classification.	0.08733 RMSE value on UCI database having 4601 samples
Jantan et al. [17]	Enhanced Bat Algorithm to optimize Feed-Forward Neural Networks using learning error as fitness function	0.483 average Mean Squared Error over 10 runs with 11 neurons in hidden layer on UK 2011 WEBSPAM dataset
Rajamohana et al. [36]	Adaptive Binary Flower pollination algorithm for feature selection using Naive Bayes classifier's accuracy as the objective function and k-nearest neighbors as the classifier using selected features	More than 85% accuracy was observed for 1600 reviews from the 20 most popular Chicago hotels

Continued on next page

Table 1 – continued from previous page

Author	Methodology	Results
Aswani et al. [5]	k-Means with LFA with chaos, LFA without chaos, FA with chaos, FA without chaos for tuning either the Absorption Coefficient(μ) or the Attractiveness Coefficient(α)	97.98% accuracy with k-Means with LFA with chaos for tuning μ
Singh et al. [45]	Correlation based Feature Selection with Particle Swarm Optimization for feature selection with 5 classifiers namely Naive Bayes, J48, AdaBoost, Support Vector Machine, Multi Layer Perceptron	Proposed feature selection method improves the F-score of Support Vector Machine by 45.83%, AdaBoost by 33.02%,vMulti Layer Perceptron by 10.38%, J48 by 9.54%
Chikh et al. [7]	Combined clustered negative selection algorithm and fruitfly optimization	93.88% accuracy on 4601 emails out of which 39% were labeled spam and 61% were labeled non spam
Assaf and Jassam [4]	Chaotic Binary PSO for feature selection using classification accuracy of SVM as objective function. SVM is also used as a classifier.	95% accuracy with 21 features
Saleh et al. [39]	Negative Selection Algorithm	98.5% accuracy on six Enron email datasets containing a total of 33,792 emails out of which 17,184 were spams and 16,608 were non-spam
Shuaib et al. [43]	Whale Optimization Algorithm for feature selection and rotation forest for classification.	99.89% accuracy with 20 fold cross validation on spambase corpus containing 4601 emails out of which 1813 were spams and 2788 were non-spams
Singh et al. [44]	Intelligent Water Drop for feature selection and Naive Bayes over selected features for classification	94% accuracy on UCI repository containing 4601 emails out of which 1813 were labeled as spam and 2788 were labeled as ham
Pandey et al. [33]	Spiral Cuckoo search to optimize k-Means algorithm using sum squared error as the objective function	Tested on spam review, synthetic spam review, yelp hotel review, yelp resturant review, twitter spam dataset with 64.82%, 71.63%, 70.92%, 71.42% and 97.93% average accuracy respectively

5 Conclusion

In the existing web-based platforms, spamming is unavoidable. With the different levels of progress, spam filtering techniques have been analysed across different platforms. This review focuses on the recent developments in spam detection methods. The overview of the conventional approaches is covered along with the emerging trends in detection of spam. The different paltforms where spam is generated such as e-mails, social networking websites like Facebook, Twitter, LinkedIn, etc., microblogging sites, blogs and forums are critically analysed for spam detection techniques. The identified methods vary broadly in deterministic, graph-based, probabilistic and optimization-based categories. A deliberate problem in the filed of review spam detection has been identified as not enough work is done this area. From the literature, it is apparent that the features in social networks vary form those in Web pages and documents, making social networks more prone to spamming. The posts on social platforms are eminently private, full of opinions and consist of a lot of local implications, inclusive of various languages and sarcasm. This makes it very difficult for a system to efficiently identify spam. Hence, to identify all the attributes in social media content and marking them with an equitable amount of accuracy is not a trivial task and forms a promising direction of research. In this review, we attempt to accumulate a compilation of different spam detection techniques and how they have been used.

References

[1] Kriti Agarwal and Tarun Kumar. Email spam detection using integrated approach of naïve bayes and particle swarm optimization. In *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, pages 685–690. IEEE, (2018).

[2] Tiago A Almeida, José María G Hidalgo, and Akebo Yamakami. Contributions to the study of sms spam filtering: new collection and results. In *Proceedings of the 11th ACM symposium on Document engineering*, pages 259–262, (2011).

[3] Muhammad Zubair Asghar, Asmat Ullah, Shakeel Ahmad, and Aurangzeb Khan. Opinion spam detection framework using hybrid classification scheme. *Soft computing*, 24(5):3475–3498, (2020).

[4] Omer Y Assaf and Noor M Jassam. An enhanced particle swarm optimization algorithm for e-mail spam filtering (2019).

[5] Reema Aswani, Arpan Kumar Kar, and P Vigneswara Ilavarasan. Detection of spammers in twitter marketing: a hybrid approach using social media analytics and bio inspired computing. *Information Systems Frontiers*, 20(3):515–530, (2018).

[6] Enrico Blanzieri and Anton Bryl. A survey of learning-based techniques of email spam filtering. *Artificial Intelligence Review*, 29(1):63–92, (2008).

[7] Ramdane Chikh and Salim Chikhi. Clustered negative selection algorithm and fruit fly optimization for email spam detection. *Journal of Ambient Intelligence and Humanized Computing*, 10(1):143–152, (2019).

[8] Colin Cooper and Alan Frieze. A general model of web graphs. *Random Structures & Algorithms*, 22(3):311–335, (2003).

[9] Lorrie Faith Cranor and Brian A LaMacchia. Spam! *Communications of the ACM*, 41(8):74–83, (1998).

[10] Marco Dorigo, Mauro Birattari, and Thomas Stutzle. Ant colony optimization. *IEEE computational intelligence magazine*, 1(4):28–39, (2006).

[11] Hossam Faris, Ibrahim Aljarah, and Bashar Al-Shboul. A hybrid approach based on particle swarm optimization and random forests for e-mail spam filtering. In *International conference on computational collective intelligence*, pages 498–508. Springer, (2016).

[12] Alaa Abi-Haidar and Luis Mateus Rocha. Adaptive spam detection inspired by the immune system. In *ALIFE*, pages 1–8, (2008).

[13] Wenrui He, Guyue Mi, and Ying Tan. Parameter optimization of local-concentration model for spam detection by using fireworks algorithm. In *International Conference in Swarm Intelligence*, pages 439–450. Springer, (2013).

[14] Ismaila Idris and Ali Selamat. Improved email spam detection model with negative selection algorithm and particle swarm optimization. *Applied Soft Computing*, 22:11–27, (2014).

[15] Ismaila Idris, Ali Selamat, and Sigeru Omatu. Hybrid email spam detection model with negative selection algorithm and differential evolution. *Engineering Applications of Artificial Intelligence*, 28:97–110, (2014).

[16] Kunal Jain and Sanjay Agrawal. A hybrid approach for spam filtering using support vector machine and artificial immune system. In *2014 First International Conference on Networks & Soft Computing (ICNSC2014)*, pages 5–9. IEEE, (2014).

[17] AMAN JANTAN, WAHEED AHM GHANEM, and SANAA AA GHALEB. Using modified bat algorithm to train neural networks for spam detection. *Journal of Theoretical & Applied Information Technology*, 95(24), (2017).

[18] Xin Jin, Cindy Xide Lin, Jiebo Luo, and Jiawei Han. Socialspamguard: A data mining-based spam detection system for social media networks. *Proceedings of the VLDB Endowment*, 4(12):1458–1461, 2011.

[19] Nitin Jindal and Bing Liu. Opinion spam and analysis. In *Proceedings of the 2008 international conference on web search and data mining*, pages 219–230, (2008).

[20] Sandeep Kaur and Kuljit Kaur Chahal. Hybrid anfis-genetic algorithm based forecasting model for predicting cholera-waterborne disease. *International Journal of Intelligent Engineering Informatics*, 8(4):374–393, (2020).

[21] Amit Kumar Kushwaha, Arpan Kumar Kar, and Yogesh K Dwivedi. Applications of big data in emerging management disciplines: A literature review using text mining. *International Journal of Information Management Data Insights*, 1(2):100017, (2021).

[22] Asha Kumari and Balkishan. Detection of threatening user accounts on twitter social media database. *International Journal of Intelligent Engineering Informatics*, 7(5):457–489, (2019).

[23] Sunil Kumar, Arpan Kumar Kar, and P Vigneswara Ilavarasan. Applications of text mining in services management: A systematic literature review. *International Journal of Information Management Data Insights*, 1(1):100008, (2021).

[24] Zhijun Liu, Weili Lin, Na Li, and David Lee. Detecting and filtering instant messaging spam-a global and personalized approach. In *1st IEEE ICNP Workshop on Secure Network Protocols, 2005.(NPSec).*, pages 19–24. IEEE, (2005).

[25] Chih-Chin Lai and Chih-Hung Wu. Particle swarm optimization-aided feature selection for spam email classification. In *Second International Conference on Innovative Computing, Informatio and Control (ICICIC 2007)*, pages 165–165. IEEE, (2007).

[26] Tarek M Mahmoud and Ahmed M Mahfouz. Sms spam filtering technique based on artificial immune system. *International Journal of Computer Science Issues (IJCSI)*, 9(2):589, (2012).

[27] Adel Hamdan Mohammad and Raed Abu Zitar. Application of genetic optimized artificial immune system and neural networks in spam detection. *Applied Soft Computing*, 11(4):3827–3845, (2011).

[28] Hamdy Mubarak, Kareem Darwish, and Walid Magdy. Abusive language detection on arabic social media. In *Proceedings of the first workshop on abusive language online*, pages 52–56, (2017).

[29] Arulanand Natarajan and Premalatha K Subramanian. An enhanced cuckoo search for optimization of bloom filter in spam filtering. *Global Journal of Computer Science and Technology*, 2012.

[30] Ngoc-Tri Ngo and Thi Thu Ha Truong. Forecasting time series data using moving-window swarm intelligence-optimised machine learning regression. *International Journal of Intelligent Engineering Informatics*, 7(5):422–440, (2019).

[31] Terri Oda and Tony White. Developing an immunity to spam. In *Genetic and Evolutionary Computation Conference*, pages 231–242. Springer, (2003).

[32] Terri Oda and Tony White. Immunity from spam: An analysis of an artificial immune system for junk email detection. In *International conference on artificial immune systems*, pages 276–289. Springer, (2005).

[33] Avinash Chandra Pandey and Dharmveer Singh Rajpoot. Spam review detection using spiral cuckoo search clustering method. *Evolutionary Intelligence*, 12(2):147–164, (2019).

[34] Paul Przemyslaw Polanski. Spam, spamdexing and regulation of internet advertising. *International Journal of Intellectual Property Management*, 2(2):139–152, (2008).

[35] Juergen Quittek, Saverio Niccolini, Sandra Tartarelli, and Roman Schlegel. On spam over internet telephony (spit) prevention. *IEEE Communications Magazine*, 46(8):80–86, (2008).

[36] S P Rajamohana, K Umamaheswari, and B Abirami. Adaptive binary flower pollination algorithm for feature selection in review spam detection. In *2017 International Conference on Innovations in Green Energy and Healthcare Technologies (IGEHT)*, pages 1–4. IEEE, (2017).

[37] Saroj Ratnoo, Seema Rathee, and Jyoti Ahuja. A clustering-based hybrid approach for dual data reduction. *International Journal of Intelligent Engineering Informatics*, 6(5):468–490, (2018).

[38] Guangchen Ruan and Ying Tan. A three-layer back-propagation neural network for spam detection using artificial immune concentration. *Soft computing*, 14(2):139–150, (2010).

[39] Abdul Jabbar Saleh, Asif Karim, Bharanidharan Shanmugam, Sami Azam, Krishnan Kannoorpatti, Mirjam Jonkman, and Friso De Boer. An intelligent spam detection model based on artificial immune system. *Information*, 10(6):209, (2019).

[40] Saber Salehi and Ali Selamat. Hybrid simple artificial immune system (sais) and particle swarm optimization (pso) for spam detection. In *2011 Malaysian Conference in Software Engineering*, pages 124–129. IEEE, (2011).

[41] Nirmala Sharma, Harish Sharma, Ajay Sharma, and Jagdish Chand Bansal. Dung beetle inspired local search in artificial bee colony algorithm for unconstrained and constrained numerical optimisation. *International Journal of Intelligent Engineering Informatics*, 8(4):268–304, (2020).

[42] Sourabh Sharma, Harish Sharma, and Janki Ballabh Sharma. Artificial intelligence based watermarking in hybrid dds domain for security of colour images. *International Journal of Intelligent Engineering Informatics*, 8(4):331–345, (2020).

[43] Maryam Shuaib, Shafi'i Muhammad Abdulhamid, Olawale Surajudeen Adebayo, Oluwafemi Osho, Ismaila Idris, John K Alhassan, and Nadim Rana. Whale optimization algorithm-based email spam feature selection method using rotation forest algorithm for classification. *SN Applied Sciences*, 1(5):1–17, (2019).

[44] Maneet Singh. Classification of spam email using intelligent water drops algorithm with naive bayes classifier. In *Progress in Advanced Computing and Intelligent Engineering*, pages 133–138. Springer, (2019).

[45] Surender Singh and Ashutosh Kumar Singh. Web-spam features selection using cfs-pso. *Procedia computer science*, 125:568–575, (2018).

[46] Sudeep D Thepade, Deepa Abin, Rik Das, and Tanuja Sarode. Human face gender identification using thepade's sorted n-ary block truncation coding and machine learning classifiers. *International Journal of Intelligent Engineering Informatics*, 8(2):77–94, (2020).

[47] Adam Thomason. Blog spam: A review. In *CEAS*, (2007).

[48] Iryna Yevseyeva, Vitor Basto-Fernandes, David Ruano-Ordás, and José R Méndez. Optimising anti-spam filters with evolutionary algorithms. *Expert systems with applications*, 40(10):4010–4021, (2013).

[49] Hui Yin, Fengjuan Cheng, and Dexian Zhang. Using lda and ant colony algorithm for spam mail filtering. In *2009 Second International Symposium on Information Science and Engineering*, pages 368–371. IEEE, (2009).

[50] Mohammad Zavvar, Meysam Rezaei, and Shole Garavand. Email spam detection using combination of particle swarm optimization and artificial neural network and support vector machine. *International Journal of Modern Education and Computer Science*, 8(7):68, (2016).

[51] Yudong Zhang, Shuihua Wang, Preetha Phillips, and Genlin Ji. Binary pso with mutation operator for feature selection using decision tree applied to spam detection. *Knowledge-Based Systems*, 64:22–31, (2014).

[52] CGS Blogs. 'how a pro-trump network is building a fake empire on facebook and getting away with it'. *https://www.cgsinc.com/blog/how-to-identify-a-malicious-email-6-tips*, (2016) (accessed 30 April 2021).

[53] The New York Times. 'is that review a fake?'. *https://www.nytimes.com/2011/08/20/technology/finding-fake-reviews-online.html* (2011) (accessed 30 April 2021).

[54] Snopes. 'is that review a fake?'. *https://www.snopes.com/news/2019/11/12/bl-fake-profiles/*, (2019) (accessed 30 April 2021).

[55] Wikipedia. 'google bombing'. *https://en.wikipedia.org/wiki/Google-bombing*, (2006) (accessed 30 April 2021).

Author information

Sakshi Shringi and Harish Sharma, Department of Computer Science Engineering, Rajasthan Technical University, Kota, Rajasthan 324010, INDIA.
E-mail: `sakshi.shringi@gmail.com`