Fake News Identification Using Artificial Immune Systems

Kazi Shah Nawaz Ripon
Dept. of Computer Science and Communication
Østfold University College
Halden, Norway
ksripon@hiof.no

Md. Tazel Hossan
Dept. of Computer Science and Engineering
Jahangirnagar University
Savar, Dhaka, Bangladesh
tazel.stu2017@juniv.edu

Musfika Ikfat Munia
Dept. of Computer Science and Engineering
Jahangirnagar University
Savar, Dhaka, Bangladesh
munia.stu20161@juniv.edu

Mohammad Shorif Uddin
Dept. of Computer Science and Engineering
Jahangirnagar University
Savar, Dhaka, Bangladesh
shorifuddin@juniv.edu

***Abstract*—The dissemination of fake news through unreliable media can have severe negative consequences on national security and society, making it a significant issue in modern society. Although there has been an increase in awareness of fake news since the 2016 US presidential election, effectively detecting and preventing its rapid spread online remains a challenge. This work investigates the feasibility of identifying fake news on social media by combining Artificial Immune Systems and Genetic Algorithms. To identify fake news online, the proposed approach employs a Genetic Algorithm-based detector optimization scheme for the Negative Selection Algorithm, a vital component of Artificial Immune Systems. The experimental results demonstrate the effectiveness of the Genetic Algorithm-based optimization scheme for generating the Negative Selection Algorithm detector in identifying fake news articles online.**

***Index Terms*—artificial immune systems, fake news, negative selection algorithm, self/nonself recognition, pattern detector**

I. Introduction

The advancement and easy access to the internet and technology have revolutionized news platforms. Social media and micro-blogging platforms have become the primary news source for many people, replacing traditional media channels like newspapers and television. According to a study, approximately two-thirds of Americans now access news through social media [1]. While social networks are helpful during crisis situations, there needs to be more control and fact-checking over posts. With millions of articles published every minute, manually checking them is not feasible. This absence of control or fact-checking makes social media a breeding ground for spreading unverified news, rumours or fake news at a fast pace and with little to no cost [2]. This can seriously affect society, commerce, national security, and democracy worldwide. For instance, when a fake news report about former US President Barack Obama being injured in an explosion was circulated, it triggered a US $130 billion loss in the stock market [3]. Consequently, there has been a growing interest in the automatic detection of fake news in social media and online platforms in recent years.

The majority of works in the literature have formulated fake news detection as either a classification or anomaly detection problem. A comprehensive survey is available in [4]–[7]. The existing approaches can be broadly categorized into four groups: natural language processing, data mining, social network analysis, and machine/deep learning. Among these approaches, machine/deep learning techniques have shown promising results [5]. Traditional machine/deep learning algorithms utilized for fake news detection include Support Vector Machines [8], Decision Trees [9], Random Forests [4], k-Nearest Neighbors [10], Naïve Bayes Classifiers [11], Convolutional Neural Networks [3], Recurrent Neural Networks [3], and Long Short-Term Memory [11]. Over the past few years, bio-inspired algorithms modeled after the principles of biological evolution have led to new and effective techniques for similar classification and detection tasks in various domains [12]–[14]. Inspired by the success, this work explores the potential of utilizing a hybrid approach combining two bio-inspired algorithms, Artificial Immune Systems (AISs) and Genetic Algorithms (GAs), to detect fake news articles online.

In recent years, AIS has emerged as a newer computational paradigm based on principles derived from the Biological Immune System (BIS) [15]. The BIS is a robust, highly parallel, distributed, self-organize, and adaptive system that defends our body from foreign pathogens by distinguishing between *self* (body cells) and *nonself* (foreign materials/pathogen). It uses learning, memory, and associative retrieval to solve recognition and classification tasks. By employing similar sophisticated pattern recognition and response mechanisms against foreign invaders as the BIS, AIS has been successfully applied to practical pattern recognition and anomaly detection tasks [16]–[19]. Accordingly, AIS provides an attractive option for improving the classification performance of fake news articles by utilizing the *self/nonself* distinction. GAs, on the other hand, are search and optimization heuristics inspired by natural selection, where the strongest individuals are selected for reproduction to produce the next generation of offspring [20].

One of the BIS's most intriguing and adaptive immunological mechanisms is *self/nonself recognition*. It allows the BIS to identify cells that belong to its own body (self/safe) from the foreign cells (nonself). The nonself cells that BIS recognizes could be harmful cells originating from the body (such as cancer and tumour cells) or foreign agents causing diseases (such as viruses and bacteria). This recognition enables BIS to develop a defence mechanism against these attackers without harming its own cells. Based on the principle of self/nonself recognition, researchers advanced the AIS design by introducing the Negative Selection Algorithm (NSA) [21], [22]. This algorithm has already demonstrated its effectiveness in detecting anomalies [23], [24].

In brief, NSA involves creating a set of memory detectors through negative selection. Initially, immature detectors that match a self pattern are removed, and the ones that survive the negative selection process become mature detectors for nonself pattern. Then, a positive selection process is used for nonself-matching. Any sample pattern that matches a mature detector is identified as nonself and saved as a memory detector for future use. However, the original NSA randomly generates detectors in the first phase, which may not cover the nonself space and have overlaps [23]. To address this issue, various detector types and generation schemes have been proposed, but they tend to be computationally expensive. Inspired by the concept presented in [25], this work proposes a simple GA-based optimization scheme for NSA detectors to identify fake news articles online. Experimental results demonstrate the superiority of the GA-based approach over conventional detector generation methods and highlight its efficiency in detecting fake news.

The structure of the paper is as follows: Section II provides information regarding NSA as a background, followed by Section III, which discusses the proposed GA-based NSA model. Section IV presents the experimental outcomes, and Section V concludes the paper with points to future works.

II. Negative Selection Algorithm

The initial proposal by Forrest et al. [21] was for the Negative Selection Algorithm (NSA) to imitate the BIS's mechanism of T cell maturation, specifically through the negative selection of T cells in the thymus. From the information processing viewpoint, NSA offers a different approach for pattern recognition where information about the nonself set (in our case, fake news) is stored to recognize self set (in our case, authentic news). This approach can be described conceptually in two phases, as shown in Fig. 1.

Assuming an adequate problem representation is available:

1) The initial step involves identifying the set of patterns to be protected (*self*) and gathering a dataset comprising all the representative self samples. Afterwards, candidate detectors that adhere to the same representation are generated randomly and compared with the self set to determine if any matches occur. If a match is found, signifying that a randomly generated detector has been recognized by a self sample, it is discarded. Comparable to the negative selection of T cells, only those randomly generated detectors that are not matched with any element of the self set are saved as *mature detectors* in the mature detector set.
2) The algorithm in the second phase observes the system in use for the existence of *nonself* patterns. It evaluates the matching (*affinity*) level between the sample pattern and all components of the mature detector set comprising nonself patterns. The sample pattern is recognized as a nonself pattern whenever the matching level exceeds a specified threshold value. Eventually, appropriate action is initiated.

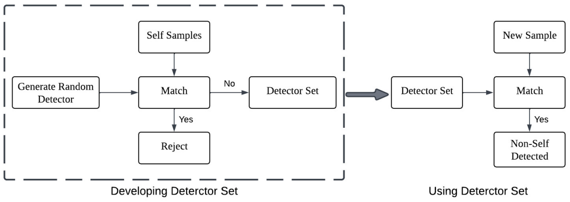


Fig. 1: The Negative Selection Algorithm (NSA) process.

The effectiveness of nonself pattern recognition by the NSA detectors relies on two critical factors [26]: the extent of nonself space coverage and the level of overlap among existing detectors. As a result, the ideal detectors should cover the maximum nonself space and have minimal overlapping among each other. Typically, conventional NSA generates detectors through a random search, which results in an exhaustive and time-consuming trial-and-error process [27]. Furthermore, detectors generated through random search cannot be guaranteed to be optimal [25]. This work utilizes GA to optimize NSA detectors to achieve the best possible anomaly detection performance in the case of fake news identification.

III. The Proposed Model

The main concept involves creating a set of mature detectors that encompass the entire negative space. The initial step consists in gathering samples labeled as "True" in the dataset (explained in Section IV) and using them as self patterns to define the positive space. Fig. 2 illustrates this, which depicts a two-dimensional example from [25]. In the figure, the black circles correspond to the self patterns and are centered around specific sample data points with a specific radius. Dotted circles represent the nonself patterns (detectors), each with a different radius. It is worth noting that the dimensions of the search space (for example, two in Fig. 2) will vary depending on the number of features in the problem domain. The primary aim is to generate the maximum number of mature detectors, each centered around a detector with a varying radius. The goal is to optimize two objectives: maximizing the coverage by these detectors (negative/nonself space) and minimizing the overlaps among themselves. This ensures that any incoming nonself pattern, such as fake news in this work, is detected by at least one mature detector (i.e., lies within a dotted circle). However, similar to [25], this work employs a single objective GA, which mainly minimizes the overlapping among the detectors to cover the maximum negative space.

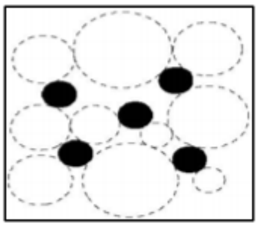


Fig. 2: Positive and negative detectors in NSA [25].

Assume that N self samples contain L features and are centered around $S_i = [x_1^i, x_2^i, ..., x_L^i]$ $(i = 1, 2, ..., N)$, with radii R_i $(i = 1, 2, ..., N)$ defining their self-regions. The aim is to generate M mature detectors with centers $md_i = [w_1^i, w_2^i, ..., w_L^i]$ $(i = 1, 2, ..., M)$ and radii r_i $(i = 1, 2, ..., M)$ so that the optimized detectors can cover the maximum nonself space without overlap. Specifically, for each detector j, we determine S_k as the center of its nearest self sample k:

$$|S_k - md_j| = \min_j^N |S_i - md_j| \tag{1}$$

Where $|S_k - md_j|$ and $|S_i - md_j|$ are Euclidean distances. If there is a self-sample i, such that $|S_i - md_j| \leqslant R_i$, the detector md_j becomes invalid since it overlaps with the self sample. To ensure that the detectors have the maximum radius r without any overlap with the N *self* samples, the radius of detector j, r_j, is chosen as follows:

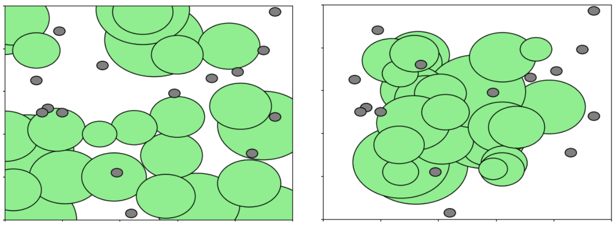
$$r_j = |S_k - md_j| - R_k \tag{2}$$

Thus, we aim to search for the optimal md_j that offers the largest r_j to cover the nonself space. Accordingly, radius of the detector is the fitness of the GA, and larger valid radii detectors are given higher fitness for evolution in the GA (as shown in (2)). In our implementation, each run of the GA optimization loop results in only one optimal detector. In each successive iteration of the GA, the most recent best detector is added to the set of previously discovered best detectors and verified for its validity (i.e., whether it lies within the radii of the other detectors). If valid, it is added to the mature detector set; otherwise, it is ignored. To ensure that the detectors do not overlap, the optimal detector obtained in one iteration is considered a “new” self sample and included in the “varying” self sample set. It is important to note that after adding the best detector, the radii of the other detectors may change, so they must be recalculated. As per (1), the final set of mature detectors generated in this approach does not overlap and provide the maximum coverage of the nonself space.

IV. Experiments and Results

To evaluate the proposed GA-based NSA model, a publicly available dataset was utilized, which was compiled by Ahmed, Traore, and Saad [28]. The dataset comprises 12600 authentic articles and 12600 fake news articles, all obtained from real-world sources. The authentic articles were gathered from Reuters.com, a news agency website. In contrast, the fake news articles were sourced from a dataset available on kaggle.com containing articles from unreliable websites marked as such by Politifact, a fact-checking organization based in the USA. Each article in the dataset is described by five features: article text, article type, article title, article date, and article label indicating whether the article is fake or authentic.

The dataset was divided into training and testing sets using 5-fold cross-validation. In each validation cycle, 75% of the data was utilized for training, while the remaining 25% was used for testing purposes. We employed 25 detectors initially with a radius size of 2 for GA-based NSA. The GA had a population size of 30 and ran for 20 generations, with crossover and mutation probabilities of 0.7 and 0.3, respectively. For the random generating scenario, we also utilized 25 detectors with a threshold value of 1.



(a) GA-baed NSA

(b) Random NSA

Fig. 3: Search space with self samples and generated detectors.

Fig. 3 presents the detectors generated by GA-based NSA (Fig. 3a) and NSA with a random strategy (Fig. 3b). To enhance their visualization, the detectors are presented in two dimensions and represented by black circles for self patterns and coloured circles for detectors. The figures demonstrate that GA-based NSA generates mature detectors with significantly less overlap than those generated by the NSA with a random generating strategy. While the mature detector set obtained by GA-based NSA (Fig. 3a) does not cover the complete negative space, it covers more negative space compared to the mature set generated by the random strategy (Fig. 3b). It is important to note that the GA-based NSA in this work had a single objective: minimizing overlap among mature detectors. Having taken this into account, the results obtained are highly promising.

TABLE I: Performance Comparison In Detecting Fake News

Model	Precision	Recall	F1-Score	Accuracy
GA-based NSA	94.11%	100%	96.96%	98%
NSA (random)	45.65%	87.5%	60%	44%

Table I presents the classification accuracy achieved by the two approaches in detecting fake news articles, serving as evidence of the effectiveness of the proposed approach. The results clearly indicate that GA-based NSA outperforms NSA with a random strategy across all accuracy measures considered. The precision and recall values obtained by GA-based NSA are substantially higher at 94.11% and 100%, respectively. These findings demonstrate that GA-based NSA can almost perfectly distinguish between "True" and "Fake" news samples, with minimal risk of misprediction. The high level of accuracy (98%) achieved by the proposed approach supports its ability to predict and classify the majority of new samples accurately. The high F1 score (96.96%) obtained by the GA-based NSA further demonstrates its capacity to achieve high precision and recall for a significant proportion of unknown samples. Furthermore, this validates the effectiveness of the proposed approach when dealing with "real" fake news.

V. Conclusion

One of the primary challenges in identifying fake news is distinguishing it from trustworthy news. AIS-based techniques have demonstrated impressive performance in classification, pattern recognition, and anomaly detection applications. Building on this success, this work proposes a GA-based NSA, modeled after the BIS's self/nonself recognition, to identify online fake news articles. The experimental results demonstrate the effectiveness of this approach in accurately identifying and classifying news articles. The proposed approach generates mature detectors that exhibit minimal overlap among themselves. However, as a single-objective GA was employed to optimize the detectors, the mature detectors could not fully cover the negative space. Future research aims to use a multi-objective GA to maximize negative space coverage and minimize overlapping among detectors simultaneously. Ultimately, the goal is to develop an integrated online news identification system that automatically generates training data for identifying fake news on social media. Advancements in generative modeling, such as Generative Adversarial Networks, may help achieve this objective.

References

[1] X. Zhang and A. A. Ghorbani, "An overview of online fake news: Characterization, detection, and discussion," *Information Processing & Management*, vol. 57, no. 2, p. 102025, 2020.

[2] A. Zubiaga, A. Aker, K. Bontcheva, M. Liakata, and R. Procter, "Detection and resolution of rumours in social media: A survey," *ACM Computing Surveys (CSUR)*, vol. 51, no. 2, pp. 1–36, 2018.

[3] J. A. Nasir, O. S. Khan, and I. Varlamis, "Fake news detection: A hybrid cnn-rnn based deep learning approach," *International Journal of Information Management Data Insights*, vol. 1, no. 1, p. 100007, 2021.

[4] M. S. Raja and L. A. Raj, "Fake news detection on social networks using machine learning techniques," *Materials Today: Proceedings*, vol. 62, pp. 4821–4827, 2022.

[5] A. Bondielli and F. Marcelloni, "A survey on fake news and rumour detection techniques," *Information Sciences*, vol. 497, pp. 38–55, 2019.

[6] A. Nagaraja, S. KN, A. Sinha, J. V. RAJENDRA KUMAR, and P. Nayak, "Fake news detection using machine learning methods," in *International Conference on Data Science, E-learning and Information Systems 2021*, 2021, pp. 185–192.

[7] D. De Beer and M. Matthee, "Approaches to identify fake news: A systematic literature review," *Integrated Science in Digital Age 2020*, pp. 13–22, 2021.

[8] K. M. Yazdi, A. M. Yazdi, S. Khodayi, J. Hou, W. Zhou, and S. Saedy, "Improving fake news detection using k-means and support vector machine approaches," *International Journal of Electronics and Communication Engineering*, vol. 14, no. 2, pp. 38–42, 2020.

[9] I. Ahmad, M. Yousaf, S. Yousaf, and M. O. Ahmad, "Fake news detection using machine learning ensemble methods," *Complexity*, vol. 2020, pp. 1–11, 2020.

[10] A. Kesarwani, S. S. Chauhan, and A. R. Nair, "Fake news detection on social media using k-nearest neighbor classifier," in *2020 international conference on advances in computing and communication engineering (ICACCE)*. IEEE, 2020, pp. 1–4.

[11] S. Senhadji and R. A. San Ahmed, "Fake news detection using naïve bayes and long short term memory algorithms," *IAES International Journal of Artificial Intelligence*, vol. 11, no. 2, p. 746, 2022.

[12] A. Yadav and D. K. Vishwakarma, "A comparative study on bio-inspired algorithms for sentiment analysis," *Cluster Computing*, vol. 23, pp. 2969–2989, 2020.

[13] R. Aswani, A. K. 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*, vol. 20, pp. 515–530, 2018.

[14] J. Batra, R. Jain, V. A. Tikkiwal, and A. Chakraborty, "A comprehensive study of spam detection in e-mails using bio-inspired optimization techniques," *International Journal of Information Management Data Insights*, vol. 1, no. 1, p. 100006, 2021.

[15] L. N. De Castro, L. N. Castro, and J. Timmis, *Artificial immune systems: a new computational intelligence approach*. Springer Science & Business Media, 2002.

[16] P. Saurabh and B. Verma, "An efficient proactive artificial immune system based anomaly detection and prevention system," *Expert Systems with Applications*, vol. 60, pp. 311–320, 2016.

[17] D. Dasgupta, *Artificial immune systems and their applications*. Springer Science & Business Media, 2012.

[18] J. H. Carter, "The immune system as a model for pattern recognition and classification," *Journal of the American Medical Informatics Association*, vol. 7, no. 1, pp. 28–41, 2000.

[19] N. Bayar, S. Darmoul, S. Hajri-Gabouj, and H. Pierreval, "Fault detection, diagnosis and recovery using artificial immune systems: A review," *Engineering Applications of Artificial Intelligence*, vol. 46, pp. 43–57, 2015.

[20] J. Holland, "Adaptation in natural and artificial systems mit press," *Cambridge, MA*, 1975.

[21] S. Forrest, A. S. Perelson, L. Allen, and R. Cherukuri, "Self-nonself discrimination in a computer," in *Proceedings of 1994 IEEE computer society symposium on research in security and privacy*. IEEE, 1994, pp. 202–212.

[22] S. A. Hofmeyr and S. Forrest, "Architecture for an artificial immune system," *Evolutionary computation*, vol. 8, no. 4, pp. 443–473, 2000.

[23] C. Laurentys, G. Ronacher, R. M. Palhares, and W. M. Caminhas, "Design of an artificial immune system for fault detection: a negative selection approach," *Expert Systems with Applications*, vol. 37, no. 7, pp. 5507–5513, 2010.

[24] M. Gong, J. Zhang, J. Ma, and L. Jiao, "An efficient negative selection algorithm with further training for anomaly detection," *Knowledge-Based Systems*, vol. 30, pp. 185–191, 2012.

[25] X.-Z. Gao, S. Ovaska, and X. Wang, "Genetic algorithms-based detector generation in negative selection algorithm," in *2006 IEEE mountain workshop on adaptive and learning systems*. IEEE, 2006, pp. 133–137.

[26] F. A. Gonzalez, *A study of artificial immune systems applied to anomaly detection*. the University of Memphis, 2003.

[27] D. Dasgupta, K. KrishnaKumar, D. Wong, and M. Berry, "Negative selection algorithm for aircraft fault detection," in *Artificial Immune Systems: Third International Conference, ICARIS 2004, Catania, Sicily, Italy, September 13-16, 2004. Proceedings 3*. Springer, 2004, pp. 1–13.

[28] H. Ahmed, I. Traore, and S. Saad, "Detecting opinion spams and fake news using text classification," *Security and Privacy*, vol. 1, no. 1, p. e9, 2018.