Journal of Information and Organizational Sciences
Volume 48, Number 2 (2024)
Journal homepage: jios.foi.hr

JIOS

DOI: 10.31341/jios.48.2.5
Open Access

UDC 004.94:316.776.32:004.77
Original Scientific Paper

Modeling Information Diffusion in Online Social Networks Using a Modified Firefly Algorithm

Amjad Alloush[1*], Ghaida Rebdawi[2] and Mohammad Saeed Abou Trab[1]

[1] Faculty of Informatics and Engineering, University of Damascus, Damascus, Syria
[2] Faculty of Organization and Informatics, Higher Institute for Applied Sciences and Technology, Damascus, Syria
[*]Correspondence: amjad.alloush@damascusuniversity.edu.sy

PAPER INFO

Paper history:
Received 15 January 2024
Accepted 22 April 2024
Published 18 December 2024

Citation:
Alloush, A., Rebdawi, G. & Trab, M. S. A. (2024). Modeling Information Diffusion in Online Social Networks Using a Modified Firefly Algorithm. In Journal of Information and Organizational Sciences, vol. 48, no. 2, pp. 311-322

ABSTRACT

The dynamics and patterns of information propagation on online social networks are complex and challenging to model and to predict. This study proposes a novel algorithm for simulating the spread of information on online social networks using a swarm-based approach. The algorithm is based on the firefly algorithm, which incorporates a new term called Vantablack to represent the non-spreader nodes in the network. The proposed algorithm is validated on three real-world datasets extracted from Kaggle.com, covering different topics and domains. The proposed algorithm outperforms other baseline methods in terms of accuracy and efficiency in predicting information diffusion.

Keywords: social media, information dissemination, swarm intelligence, firefly model, Vantablack

1. Introduction

Over the past decade, there has been a tremendous rise in the dominance of social networking platforms such as Facebook, LinkedIn, and Twitter [1]. Social network models are represented as graphs [2], where nodes represent users and edges represent connections between these nodes [3]. Additionally, considering the widespread adoption of portable internet devices such as tablets and smartphones, it is evident that the popularity of social networking sites will continue to expand in the future [4].

The wonders of nature have served as a profound inspiration for the development of numerous metaheuristic algorithms. Nature has a remarkable ability to discover solutions to problems through experience without explicit instruction. The concept of natural selection and the survival of the fittest have been pivotal driving forces behind initial metaheuristic algorithms [5]. Various animal species employ diverse modes of communication to interact and exchange information among themselves [5].

The Firefly algorithm is a swarm intelligence-based optimization algorithm inspired by the flashing behavior of fireflies. It has been applied to various optimization problems, including feature selection, image segmentation, and clustering [6].

This study assumes that the dissemination of information in an online social network can be similar to the synchronized flashing patterns observed in firefly swarms. To investigate this analogy, we employed the firefly algorithm to model the propagation of information in real-life Twitter datasets, focusing on user-generated tweets associated with a specific hashtag. This study presents the following main contributions:

We introduced a new aspect of the traditional firefly model by incorporating a Vantablack node status. This classification allowed us to differentiate between two types of individuals within the infected population: spreaders, who actively disseminate information, and non-spreaders, who choose to retain information without further propagation.

The remainder of this paper is structured as follows:

Section 2 describes the exploration of traditional models in the information diffusion field. Section 3 focuses on the methodology employed and introduces the proposed algorithm. Section 4 provides a detailed description of the datasets used to validate the model. In Section 5, we present visualizations of the proposed algorithm applied to select portions of a Twitter network. Section 6 discusses the results obtained from diverse real-world datasets. Finally, Section 7 serves as the conclusion, highlighting the paper's contribution to the field of information diffusion.

2. Background and Diffusion Models

Information diffusion models are computational models designed to understand and predict the spread of information through social networks. These models have garnered significant attention across various disciplines, including social science, computer science, and marketing. By simulating information propagation, researchers can gain insights into the mechanisms driving viral phenomena, influence dynamics, and information cascades. In their survey, Guille et al. [7] categorize information diffusion models into two types: prediction models and explanatory models. Figure 1 illustrates this classification.

Prediction Models: Prediction models for information diffusion focus on forecasting the future spread of information based on historical data and network characteristics [7]. These models are valuable for making informed decisions, optimizing resource allocation, and designing targeted interventions [8][7][9][10]. Yang et al. [11] propose a prediction model that integrates both explicit and implicit networks to forecast information diffusion, demonstrating improved predictive accuracy compared to traditional diffusion models (IC – SIR) [11]. Kuhlman et al. [12] suggest a model that combines epidemic dynamics with network structure to predict the spread of information in social networks, providing insights into optimal control strategies for influencing information diffusion.

Explanatory Models: Explanatory models aim to understand the underlying mechanisms and factors that drive information diffusion. These models provide valuable insights into the dynamics of information propagation, the identification of influential nodes and the emergence of information cascades [8][7][9][10]. Aral et al. [13] investigated the factors that make individuals influential or susceptible to information in social networks. The authors developed a model that quantifies the influence of individuals based on their network position and susceptibility to influence. Weng et al. [14] examined the relationship between network structure and information diffusion, proposing a model that predicts the virality of information by considering both network topology and content features.

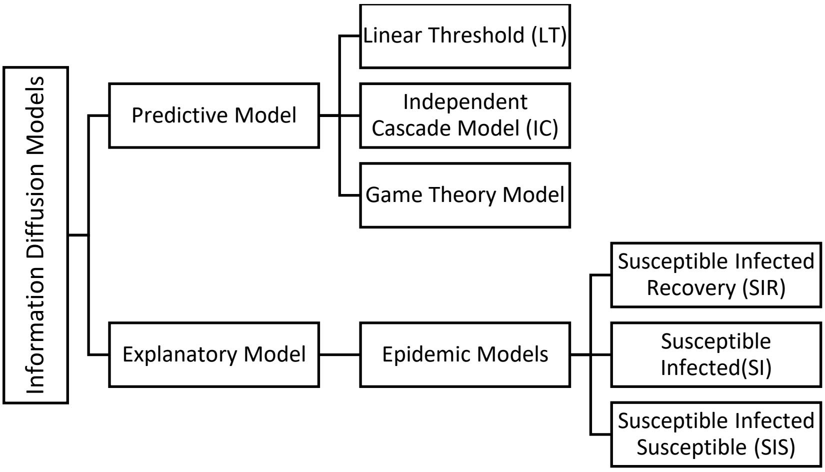

Figure 1. Models Classification

3. Related Works

Recently, other researchers have attempted to go beyond traditional epidemic models, such as SIR, SIS, and SI models, and have used nature-inspired algorithms to represent information dissemination.

Supriadi and colleagues [15] developed a retweet prediction system to analyze tweet propagation patterns. Their approach involved employing an Artificial Neural Network classification optimized with the Firefly Algorithm, utilizing both user-based and content-based features. In another study, Ghasemi et al. [16] introduced modifications to the Firefly Algorithm, known as Firefly Algorithm 1 to 3, aimed at enhancing global exploration and convergence characteristics through diverse firefly movement strategies. Li [17] enhanced a financial investment risk prediction model by integrating the Graph Convolutional Network (GCN) and Firefly Algorithm (FA), resulting in improved prediction accuracy. Chen [18] proposed a predictive firefly algorithm wherein fireflies exhibiting insufficient progress were eliminated, allowing new fireflies to replace them and explore for superior optimal values within the search function. Bei [19] presented an advanced hybrid firefly algorithm, the Improved Hybrid Firefly Algorithm with Probability Attraction Model (IHFAPA), addressing challenges related to computational efficiency and accuracy in complex optimization problem-solving.

Rui et al. [20] used the forest fire model to generate graph. Indu et al. [21] represented the interactions in social networks using the forest fire model. Kumar et al. [8] modified this model by adding the burnt state to represent the non-spreader. Wolfram [22] proposed biological-inspired Cellular Automata Models. Cellular automata models are inspired by biological systems and simulate local interactions to model global behavior. These models have been adapted to study the diffusion of information across spatial domains, resembling how signals propagate through biological tissues. Patterns emerging from cellular automata capture the intricate dynamics of information spreading. Xiao et al. [23] presented a swarm intelligence model for information propagation in mobile ad hoc networks. Their study combined the principles of ant colony optimization and particle swarm optimization to optimize the delivery of information packets in a dynamic network. Wang et al. [24] focused on dynamic information diffusion in social networks using swarm intelligence. The authors propose a dynamic swarm-based model that adapts to changes in network topology, optimizing the propagation of information through decentralized interactions. Moreover, an optimization algorithm inspired by the behavior of bird flocks or fish schools called Particle Swarm Optimization was introduced in [25]. It has been applied to information diffusion to model the collective behavior of individuals in a network. Each particle in the swarm represents an individual, and they explore the network space to find optimal information dissemination strategies. Also, Ant Colony Optimization [26] is inspired by the behavior of ant colonies when searching for food. It has been adapted to model information diffusion by considering the pheromone trail laid by ants as information spreads through a network. The behavior of ants and reinforcement of positive trails are used to optimize the dissemination process.

4. Methodology

In this section, we describe the firefly algorithm. Then, we propose a modification to this algorithm to model the spread of information, and the Vantablack node is added to distinguish the individual population into spreaders and non-spreaders. This helps identify the individual who plays a significant role in spreading information. Additionally, we establish and define the key parameters that influence the extent and dynamics of information spread.

Our model aims to identify the nodes that have the potential to propagate information further, as well as those that are less likely to contribute to the spread.

4.1.1. The Firefly Algorithm

Fireflies generate brief flashes characterized by distinct patterns through bioluminescence [27]. These flashes serve the purpose of attracting potential mates and signaling potential threats. By observing the light patterns, compatible firefly mates respond by either imitating the same pattern or producing their own unique patterns in return. Additionally, it is important to recognize that the intensity of light diminishes over distance. Consequently, the flashing light emitted by a firefly elicits responses from nearby fireflies within the visual range of the flash.

The Firefly Algorithm is inspired by the flashing behavior of fireflies, which they use for signaling and mating purposes. The algorithm assumes that the intensity of the flash is proportional to the quality of the solution and that fireflies are attracted to brighter flashes. The algorithm creates a population of fireflies, each representing a possible solution to the optimization problem. The fireflies move towards the brighter ones in each iteration, updating their positions and intensities. The algorithm also introduces some randomness to avoid getting stuck in the local optima.

The algorithm begins by initializing a set of random solutions (fireflies) and assigning them random brightness levels (objective function values). The fireflies then move towards each other in search of brighter ones, and their positions are updated based on a combination of their current position, the distance to other fireflies, and their brightness. The movement is determined by a set of parameters, such as the attraction coefficient, absorption coefficient, and step size, which control the rate of convergence and exploration.

The algorithm continues iterating until a stopping criterion is met, such as reaching the maximum number of iterations or achieving a desired level of convergence. At the end of the optimization process, the best solution is returned as the global optimum.

The Firefly Algorithm has been shown to be effective in solving a wide range of optimization problems, including engineering design, image processing, and data mining. Its strengths lie in its simplicity, scalability, and ability to avoid getting stuck in local optima.

We used this algorithm with modifications to be used with the information spread over social networks.

4.1.2. The Modified Firefly Algorithm

The basic idea of modeling information diffusion using the firefly algorithm is to simulate the spread of information from one node to another in a social network. firefly is represented in the network as a node, and the intensity of its light represents its information or news. The nodes exchange information based on a set of rules, which are updated according to the intensity of their light.

In the context of our network model, individual fireflies serve as symbolic representations of network nodes, and their emitted light symbolizes the transfer of information. These nodes actively engage with incoming information and subsequently relay it to their connected peers. However, within this network, we have identified a distinct state, which we have called ‘Vantablack.' Nodes in the Vantablack state receive incoming information but abstain from disseminating it to other connected nodes. Consequently, when a node transitions into the Vantablack state, the propagation of information comes to a dead-end at that specific node, while continuing its course throughout other non-Vantablack nodes within the network. This unique state plays a critical role in understanding and analyzing information diffusion dynamics in our network.

The algorithm also considers absorption parameters, which encompass factors such as the degradation of information's relevance over time or due to external events like earthquakes or crises. These parameters reflect the transient nature of information and acknowledge that its significance may diminish as time elapses or as specific events lose their relevance over time. By incorporating these absorption parameters, the algorithm effectively adapts to the dynamic nature of information and its contextual importance, ensuring a more accurate and responsive handling of data in a variety of scenarios.

One of the significant advantages of using the firefly algorithm for information diffusion modeling is its ability to handle large-scale datasets [28], as highlighted by Kumar as a potential area of algorithm application. Additionally, the Firefly algorithm can generate insights into the dynamics of social networks, such as the spread of rumors, opinions, and trends.

The modified firefly Algorithm is as follow:

1. *Initialize the population of firefly's network (G).*
2. *Define the fitness function that measures the spread of information in the diffusion model (f).*
3. *The initial light intensity for each firefly was set based on its fitness value (provide the information for the initial set of the influencers).*
4. *Set the attractiveness parameter (β) and absorption coefficient (γ).*
5. *Repeat for all fireflies:*
    a. *Evaluate the fitness of neighbors (the connected nodes).*
    b. *For each firefly that neighboring the brighter fireflies in the population:*
        i. *Calculate the distance between the fireflies (link weight).*
        ii. *Calculate the attractiveness between the fireflies using the distance and the light intensities.*
        iii. *Update the firefly position (information) of the firefly based on the attractiveness and randomness.*
        iv. *Apply the probability function (r) on the firefly to decide if it's going to spread or to become a Vantablack.*
        v. *Absorption applied to reduce the light intensity of each firefly.*

Where G represent the network graph, f is the fitness function, β is the attractiveness parameter, γ is the absorption coefficient, r is the probability function.

The modified Firefly Algorithm (MFA) with Information Diffusion can be described as follows:

1. **Initialize the population of fireflies' network (G):** At the beginning of the simulation, a network of fireflies is created, where fireflies represent the social network users. This network serves as the basis for modeling the diffusion of information.
2. **Define the fitness function that measures the spread of information in the diffusion model (f):** A fitness function is established as below to quantify the effectiveness of information diffusion within the network.

$$f(x) = \frac{N_{inf}(x)}{N} + \sum_{k=0}^{y} \beta_{x.k} - \gamma . N_{inf}(x) - \sum_{k=0}^{y} \frac{1}{W_{con,k}(x)}$$

Where x represents the node, $N_{inf}(x)$ is the number of infected nodes by the information propagated by x, N is the total number of nodes in the network, $\beta_{x.k}$ is the attractiveness parameter, γ is the absorption coefficient, $W_{con,k}(x)$ is the weight link between connected nodes with the mean one.

1. **Set the initial light intensity for each firefly based on its fitness value:** Each firefly in the population is assigned an initial light intensity value based on its fitness score. Symbolizing their capacity to influence information diffusion.
2. **Set the attractiveness parameter (β) and absorption coefficient (**γ): Parameters β and γ are introduced to control the behavior of the fireflies in the network. β determines the attractiveness between fireflies, while γ influences the rate at which fireflies' light intensity diminishes over time.
3. **Repeat for all fireflies:** The following steps are executed for each firefly in the network:
    a. **Evaluate the fitness of neighbors (the connected nodes):** The algorithm assesses the fitness of neighboring fireflies (the connected users), which are essentially the nodes connected to the current firefly in the network.
    b. For each firefly neighboring the brighter fireflies in the population:
        i. Calculate the distance between the fireflies (link weight): The distance or link weight between the current firefly and its neighboring fireflies is computed. This distance signifies the link weight of the connection between fireflies.
        ii. Calculate the attractiveness between the fireflies using the distance and the light intensities: An attractiveness measure is calculated between the current firefly and its neighbors. This measure takes into account both the distance between fireflies and their respective light intensities. The attractiveness measurement can be denoted as the follow:

Let β_{ij} be the attractiveness measure between the current firefly i and its neighbor j. Let d_{ij} represent the distance between fireflies i and j, and let I_i and I_j denote their respective light intensities.

$$\beta \text{ij} = \frac{Ij}{\left(1 + \delta . d_{ij}^2\right)}$$

In this formula, δ is a parameter that determines the impact of the distance on the attractiveness measure. Adjusting δ allows us to control the influence of distance relative to light intensity in the attractiveness measure calculation.

iii. **Update the firefly position (information) of the firefly based on the attractiveness and randomness:** The algorithm updates the information (light intensity) of the current firefly based on the attractiveness calculated in the previous step, as well as a degree of randomness. This update represents how information is shared or modified between connected nodes.

iv. **Apply the probability function r on the firefly to decide if it's going to spread or become a Vantablack:** A probability function, denoted as 'r,' is applied to each firefly. This function determines whether the firefly will continue to spread information or enter a state referred to as 'Vantablack,' where it no longer contributes to information diffusion.

$$P(Spread) = \frac{\mathbf{1}}{\mathbf{1} + \boldsymbol{e}^{-(\textbf{intensity\_difference})}}$$

This random function that represents the node status will fall in the range

$$\begin{Bmatrix} 0 < P(Spread) \leq 1 \\ if\ the \text{ node probability } > 0.5 \end{Bmatrix}$$

then it will spread the information or it will be converted to Vantablack.

v. **Absorption was applied to reduce the light intensity of each firefly:** As time progresses, the light intensity of each firefly (representing information) decreases due to absorption. This step mimics the natural fading of information's relevance over time or in response to external factors.

5. Experiment and Dataset Description

We have evaluated our proposed model using the extracted datasets from Twitter network. This dataset contains the users and relationships between them. Our evaluation process involved subjecting our model to a comprehensive examination across three distinct real-time Twitter datasets, each focusing on specific topics of interest. This different dataset topics to validate our model across different user domains and interests. These datasets served as the foundation for our extensive validation process, allowing us to assess the model's performance and effectiveness in different contexts and subject matters. These dataset topics are: NFT Tweets [29], Bitcoin [30], and Pfizer Vaccine [31].

"Bitcoin is a decentralized digital currency, often referred to as a cryptocurrency, that was created in 2009 by an anonymous person or group of people using the pseudonym Satoshi Nakamoto."

"Pfizer Vaccine Is one of the early vaccines developed to combat the COVID-19 pandemic"

"A Non-Fungible Token (NFT) is a digital asset that represents ownership or proof of authenticity of a unique item or piece of content using blockchain technology."

Table 1 presents the hashtags used to collect tweets of a topic and the duration and the total number of tweets, retweets for which the tweets were collected for each dataset.

Topic	Hashtag used	Duration	Number of tweets	Number of retweets
Bitcoin	Cryptocurrency – BITCOIN - BTC	6/2/2021 till 31/5/2022	4569721	5252
Pfizer Vaccine	PfizerBioNTech	12/12/2020 till 24/11/2021	11013	20
NFT	NFT	01/01/2021 till 14/11/2021	12,396	104

Table 1. represents the hashtags and topics during the time

Our experiment on these datasets started from representing the extracted dataset then Define parameters such as attractiveness parameter (β), absorption coefficient (γ), and any other relevant parameters. Tune these parameters through preliminary experiments if necessary. We define the baseline comparison models

Independent Cascade (IC) and the Susceptible-Infected-Recovered (SIR) model. Then we run the explements on each dataset separately then we visualize and interpret the results as described in the next section.

6. Result and Discussion

In our study, we conducted a comprehensive analysis of information dissemination using three distinct datasets sourced from the Twitter platform. These datasets encompassed a diverse range of topics and trends. Our research helps to better understand the dynamics of information propagation within these datasets. Moreover, we leveraged the rich social network data available on Twitter by extracting user-follower relationships dataset from kaggle.com. By collecting and processing these valuable data, we were able to construct a graph that represents the intricate web of connections and interactions among Twitter users within our datasets. This graph served as a foundational element in our research, allowing us to gain deeper insights into how information flows through the Twitter network and how various user communities engage with each other.

We conducted a comparative analysis of the predicted spreader counts generated by our model in contrast to well-established models. Specifically, we benchmarked our model against the Independent Cascade (IC) Model and the Susceptible-Infected-Recovered (SIR) Model, both widely recognized and accepted frameworks for examining information diffusion phenomena. These classical models have been extensively employed in prior research and have demonstrated their effectiveness in modeling information propagation dynamics.

The figures presented below provide a visual representation of the users counts predicted by three distinct models: the Modified Firefly Algorithm (MFA) model, the Independent Cascades (IC) model, and the Susceptible-Infected-Recovered (SIR) model. These predictions are Adjacent with the actual spreader counts derived from the comprehensive Twitter dataset we employed in our study.

Our actual dataset encompasses all tweets and retweets posted by users within the examined timeframe. For each specific topic or hashtag, the actual data is constructed as the cumulative sum of tweets and retweets over time. In our analysis, individuals who engaged in posting tweets or retweets related to a particular topic or hashtag are considered spreaders of information. Consequently, the actual spreader count comprises the sum of tweets and retweets by these individuals.

Figure 2. concerns the tweets that have hashtags about Bitcoin, figure 3. represents the tweets that contain Pfizer hashtags over the time, and figure 4. represents all the tweets about the non-fungible token hashtags.

The Independent Cascades (IC) and Susceptible-Infected-Recovered (SIR) models exhibit a degree of variability in their outcomes, with instances of both overestimating and underestimating the population of spreaders in certain scenarios. It can be seen that our MFA model predicts values close to the actual data followed by IC Model and SIR Model.

Time	Actual Data	IC	SIR	MFA
Jan-21	100	100	100	100
Mar-21	3526	90228	79567	242901
Apr-21	485814	897244	379306	879117
May-21	721051	915874	618520	897054
Jun-21	1229834	1077766	1509387	1231213
Jul-21	1260405	1245860	1568901	1470781
Aug-21	1470857	1661543	2374832	1728520
Sep-21	2194391	2512572	2440984	2012793
Oct-21	2200999	2671017	3320486	2040515
Nov-21	2373685	2806536	3388186	2683695
Dec-21	2614560	3841915	3724677	2883781
Jan-22	2883684	3850127	3735007	3088370

Feb-22	3674973	4048055	3957713	3493173
Mar-22	3808568	4303516	4413864	4056498
Apr-22	3980051	4519839	4545073	4453644
May-22	4,574,973	4,574,973	4,574,973	4,574,973

Table 2. represents the data from experiment for Bitcoin

	IC	SIR	MFA
Correlation	0.978099199	0.968467	0.99002
Euclidean distance	1974818.594	2466594	970382
Cosine Similarity	0.993270782	0.990309	0.996615

Table 3. represents the statistics data for Bitcoin

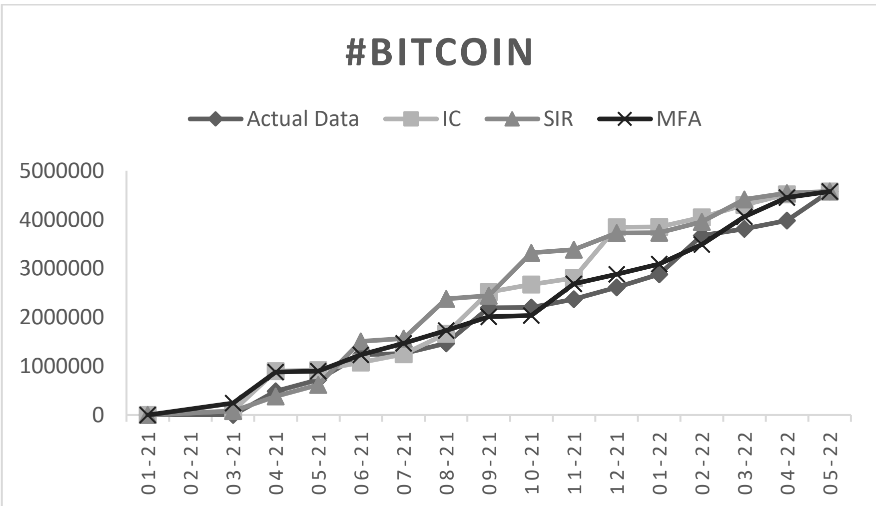

Figure 2. Bitcoin hashtag dataset comparison

Time	Actual Data	IC	SIR	MFA
Dec-20	100	100	100	100
Jan-21	142	1070	3684	1554
Feb-21	905	1402	4326	1879
Mar-21	2085	1585	6218	2682
Apr-21	2791	2628	7446	3051
May-21	6171	2838	7604	6769
Jun-21	6610	4610	7692	7298
Jul-21	7004	6836	7979	7595

Aug-21	8108	7304	8127	8684
Sep-21	8175	7847	8491	9263
Oct-21	10569	9413	10727	10063
Nov-21	11033	11433	11633	11033

Table 4. represents the data from experiment for Pfizer

	IC	SIR	MFA
Correlation	0.955127549	0.905538	0.991962
Euclidean distance	4332.690965	8225.554	2513.606
Cosine Similarity	0.984163102	0.955284	0.995356

Table 5. represents the statistics data for Pfizer

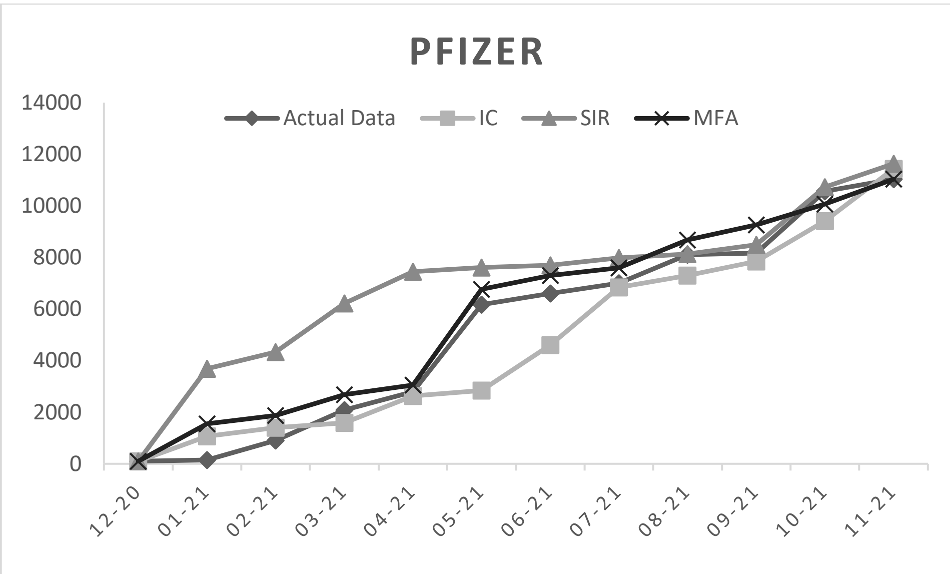

Figure 3. Pfizer hashtag dataset comparison

Time	Actual Data	IC	SIR	MFA
Jan-21	100	100	100	100
Mar-21	1028	513	302	1157
Apr-21	2191	843	4947	2533
May-21	2784	999	6625	3230
Jun-21	3507	1224	7563	3710
Jul-21	4197	4852	7813	5219

Aug-21	5873	6806	9928	7817
Sep-21	7857	9574	10438	8914
Oct-21	9450	11825	12477	10145
Nov-21	12,500	12,500	12,500	12,500

Table 6. represents the data from experiment for NFT

	IC	SIR	MFA
Correlation	0.966523604	0.909405	0.989333
Euclidean distance	4513.188562	9201.619	2607.225
Cosine Similarity	0.979869479	0.965062	0.995742

Table 7. represents the statistics data for NFT

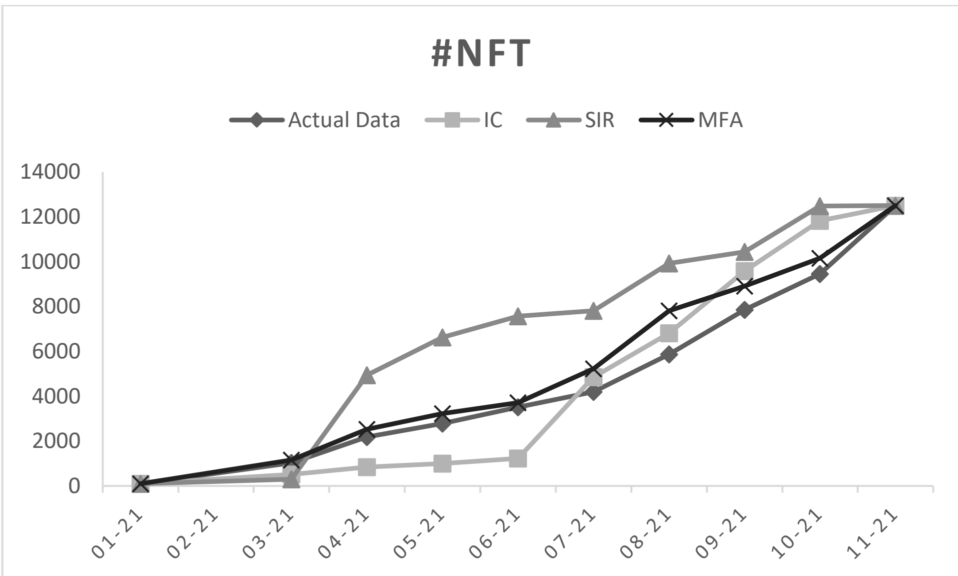

Figure 4. NFT hashtag dataset comparison

Thus, it can be concluded that our model predicts active spreaders efficiently and can be used for modeling for information diffusion in real-life scenarios after calculating the correlation, Euclidean distance and Cosine Similarity. Between the actual data and the baseline algorithms and our model.

7. Conclusion

In this study, we explored how information spreads on online social networks, such as Twitter, using a model inspired by swarming behavior, known as the modified firefly algorithm. We used fireflies to represent different types of users with light-symbolizing information. Brighter fireflies transmitted information, while others remained Vantablack and did not share any information. Our model yielded favorable results when compared with other models in simulating real-life information spreading, as demonstrated on three distinct datasets. While our findings contribute significantly to the understanding of online information diffusion, there remains vast space for further exploration and enhancement of our model. Future work could involve refining the firefly algorithm by incorporating additional parameters to better mimic the complexities of user behavior and interactions. Additionally, exploring the impact of external factors, such as user engagement, content relevance, and temporal variations, could provide a more comprehensive understanding of information spread dynamics.

Acknowledgements

We acknowledge the support of Damascus University for funding this research.

References

[1] E. Bakshy, I. Rosenn, C. Marlow, and L. Adamic, "The role of social networks in information diffusion," in Proc. 21st Annu Conf World Wide Web (WWW'12), 2012, pp. 519–528.

[2] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," Computer Networks and ISDN Systems, vol. 30, no. 1-7, pp. 107–117, 1998.

[3] A. L. Barab´asi, "Network science," Cambridge University Press, 2016.

[4] G. Cai, R. Wang, and B. Qiang, "Online social network evolving model based on damping factor," Procedia Comput Sci, vol. 9, no. 1, pp. 1338–1344, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.procs.2012.04.147

[5] A. Jain, S. Sharma, and S. Sharma, "Firefly Algorithm," in Nature-Inspired Algorithms Applications, 2021, pp. 157–180.

[6] S. L. Tilahun and H. C. Ong, "Modified firefly algorithm," J Appl Math, vol. 2012.

[7] A. Guille and H. Hacid, "Information Diffusion in Online Social Networks: A Survey," 2013.

[8] S. Kumar, M. Saini, M. Goel, and B. S. Panda, "Modeling information diffusion in online social networks using a modified forest-fire model," J Intell Inf Syst, vol. 56, no. 2, pp. 355–377, 2021.

[9] M. Li, X. Wang, K. Gao, and S. Zhang, "A survey on information diffusion in online social networks: Models and methods," Inf., vol. 8, no. 4, 2017.

[10] D. Li, Z. M. Xu, S. Li, T. Liu, and X. W. Wang, "A survey on information diffusion in online social networks," Chin. J. Comput., vol. 254, pp. 31–36, 2014.

[11] J. Yang, J. Leskovec, and J. M. Hofman, "Modeling information diffusion in implicit networks," in Proc. 10th ACM Int. Conf. Web Search Data Mining, 2016, pp. 233–242.

[12] C. J. Kuhlman and A. Sudbury, "Predicting and controlling the spread of information in social networks," Journal of Complex Networks, vol. 6, no. 1, pp. 80–98, 2018.

[13] S. Aral and D. Walker, "Identifying influential and susceptible members of social networks," Science, vol. 337, no. 6092, pp. 337–341, 2012.

[14] L. Weng, F. Menczer, and Y. Y. Ahn, "Virality prediction and community structure in social networks," Scientific Reports, vol. 3.

[15] M. R. Supriadi, J. Jondri, and I. Indwiarti, "Retweet Prediction Using Artificial Neural Network Method Optimized with Firefly Algorithm," JINAV: Journal of Information and Visualization, vol. 4, no. 2, pp. 179-188, 2023. [Online]. Available: https://doi.org/10.35877/454RI.jinav1903

[16] M. Ghasemi et al., "A new firefly algorithm with improved global exploration and convergence with application to engineering optimization," Decision Analytics Journal, vol. 5, p. 100125, 2022. [Online]. Available: https://doi.org/10.1016/j.dajour.2022.100125

[17] M. Li, "Financial investment risk prediction under the application of information interaction Firefly Algorithm combined with Graph Convolutional Network," PLoS ONE, vol. 18, no. 9, p. e0291510, 2023. [Online]. Available: https://doi.org/10.1371/journal.pone.0291510

[18] L. Chen and J. Li, "A Firefly Algorithm Based on Prediction and Hybrid Samples Learning," in Advanced Intelligent Computing Technology and Applications. ICIC 2023, D.-S. Huang, P. Premaratne, B. Jin, B. Qu, K.-H. Jo, and A. Hussain, Eds. Singapore: Springer, 2023, pp. 363-372. [Online]. Available: https://doi.org/10.1007/978-981-99-4755-3_23

[19] J.-L. Bei et al., "Improved Hybrid Firefly Algorithm with Probability Attraction Model," Mathematics, vol. 11, p. 389, 2023. [Online]. Available: https://doi.org/10.3390/math11020389

[20] X. Rui, S. Hui, X. Yu, G. Zhang, and B. Wu, "Forest fire spread simulation algorithm based on cellular automata," Natural Hazards, vol. 91, no. 1, pp. 309–319, 2018.

[21] V. Indu and S. M. Thampi, "A nature-inspired approach based on Forest Fire model for modeling rumor propagation in social networks," Journal of Network and Computer Applications, vol. 125, pp. 28–41, 2019.

[22] S. Wolfram, "Cellular automata as models of complexity," Nature, vol. 311, no. 5985, pp. 419–424, 1984. DOI: 10.1038/311419a0

[23] J. Xiao, Y. Zheng, and X. Wang, "A swarm intelligence model for information propagation in mobile ad hoc networks," International Journal of Computer Mathematics, vol. 84, no. 11, pp. 1597–1609, 2007. DOI: 10.1080/00207160701363959

[24] J. Wang and K. Xing, "Swarm intelligence for dynamic information diffusion in social networks," IEEE Transactions on Industrial Informatics, vol. 14, no. 1, pp. 33–42, 2018. DOI: 10.1109/TII.2017.2727114

[25] Z. Cui, J. Ren, X. S. Yang, et al., "Particle swarm optimization for optimizing information diffusion in complex networks," Physica A: Statistical Mechanics and its Applications, vol. 392, no. 4, pp. 983–993, 2013. DOI: 10.1016/j.physa.2012.10.032

[26] F. Xiong, X. Liu, X. Wang, et al., "An ant colony optimization algorithm for information diffusion in complex networks," Applied Soft Computing, vol. 60, pp. 273–283, 2017. DOI: 10.1016/j.asoc.2017.07.011

[27] X. S. Yang, "Firefly algorithms for multimodal optimization," in Proc. 5th Int. Conf. Stochastic Algorithms: Foundation Applications (SAGA ’09), 2009, vol. 5792 of Lecture notes in Computer Science, pp. 169–178.

[28] V. Kumar and D. Kumar, "A Systematic Review on Firefly Algorithm: Past, Present, and Future," Arch Computat Methods Eng, vol. 28, pp. 3269–3291, 2021. DOI: 10.1007/s11831-020-09498-y

[29] https://www.kaggle.com/datasets/mathurinache/nft-tweets

[30] https://www.kaggle.com/datasets/kaushiksuresh147/bitcoin-tweets

[31] https://www.kaggle.com/datasets/gpreda/pfizer-vaccine-tweets