International Journal of Computer Sciences and Engineering Open Access
Review Paper Vol.-8, Issue-1, Jan 2020 E-ISSN: 2347-2693

Approaches to Block Rumors in Social Networks: A Review

P. K. Tiwari[1,*], M. K. Singh[2], A.K. Bharti[3]

[1,2]Department of CSE, IEC College of Engineering and Technology, Gr. Noida, India
[3]School of Computer Science, Maharishi University of Information Technology, Lucknow, India

Corresponding Author: prabhattiwari.cse@ieccollege.com, Tel.: +91-80767-24970

DOI: https://doi.org/10.26438/ijcse/v8i1.132136 | Available online at: www.ijcseonline.org

Accepted: 18/Jan/2020, Published: 31/Jan/2020

Abstract— As Online Social Networks have become the integral part of our lives, the pros and cons of using them have been reflected in the society. On one hand, these online networks are the easiest ways to connect with your peers, communities and good for social and professional collaborations; on the other hand, they are the most vulnerable means of spreading rumors, threats and gossips within no time. There are several information diffusion algorithms using which the rumors can be shared on these mediums. The ways companies target and increase their customer base and sales using these diffusion algorithms, similarly rumors and gossips among the communities can also be shared. Some algorithms are deterministic and some are stochastic in nature. In this paper, we have reviewed the methods for spreading and blocking the rumors and compared them in the context of dynamic social networks. We have categorized the approaches on the basis of various measures and analysed their behavioural differences. The impact of several social parameters have also been studied to find the factors which are preferable to block the rumors.

Keywords— Social Networks, Information Diffusion, Rumor Blocking, Dynamic Graphs, Anti-rumors

I. INTRODUCTION

In online social networks, people create various groups on the basis of friendship and other relations to connect with other users. These groups are virtual communities, having specific common features among the members [1][2]. When we visualize these networks as graphs as shown in fig. 1, they exhibit sparseness to reflect less inter-community interaction and have dense intra- community interactions.

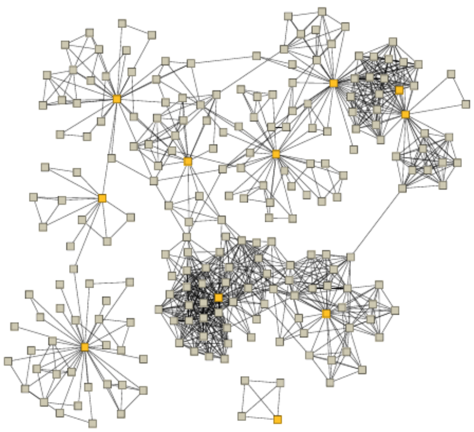

Figure 1: A social Network Graph

Social networks are **open** in nature; i.e., users can post their individual opinions. This leads to the misuse of freedom. To prohibit this, various networking sites have applied different types of checks to insure the validity of posts. But there are still various loopholes in the existing mechanisms. Another aspect of social networking sites is, their **reliability**. If malicious rumors are flooded over it, social network will not be considered as a reliable source of information. The primary strength of Social networks is rapid spreading of information throughout the network. But, considering rumors and valuable information same would be the problem. So we need to analyse the structure of network and the central nodes of that network which define how quick the information can be diffused throughout the network. Centrality can be a measure to find out the influential capabilities of nodes or edges in a network [3][4]. We can find influential nodes and restrict them using various properties like degree, closeness and between-ness centrality, diffusion mechanisms, anti-rumor techniques etc.

In the social network environment, as the network grows the distances between the nodes keep increasing. To address this situation we can perform clustering and apply various algorithms to find various influential nodes. If any message is passed through the central node, the chances of quick diffusion of the message would always be higher [5]. For example, in Facebook, a person having more number of friends has a larger network to influence. On Twitter or Instagram, if a celebrity having more followers, tweets or posts a photo, it immediately becomes a national news. This news has the power to start a world level campaign like #metoo, which ultimately empowered the women giving them a platform to raise their voices against the oppression. Any unauthentic news, like rumor about Ayodhya, is capable of setting up the religious riots. So we also need to find all those influential nodes which may or may not have the tendency of centrality, but are capable enough to affect the network flow.

In [6], authors have emphasised on various reliability issues and challenges of social networks. Their main focus is to utilize this information diffusion, online communication datasets for future research.

In a network if a link is deleted, it can disconnect the communication between various nodes and can make the network disjoint too. In [7], authors used a probabilistic way to delete links for identifying the critical nodes behind the information spread. They provide a weight called contribution value to every node in proportion to their bridge detection capability. They claim their method to be faster than the existing algorithms by showing the experimental results found by Blog, Enron and Road networks, respectively. In [8], a marginal decrement perspective is considered to minimize the rumor propagation. The approach aims at edges but blocks nodes to reduce the flow of rumor. In [9], an algorithm using community structure is proposed to find the potential nodes in social networks. The algorithm is parameter free and based on label propagation. The results on synthetic and real-world networks under common diffusion models demonstrate its efficiency. The influence of a node does not depends only upon its neighbors but also the distance between two nodes. Shortest path distance is a good measure of weighting the connectivity. In [10], a gravity centrality index is used to measure the node's influence on network. The experiments prove it to be better than various centrality measures.

Rest of the paper is organized as follows: In Section II, we define rumors and categorize them according to their feature like life span, content etc., Section III contains a taxonomy of the reviewed literature. In Section IV details of related works has been given and various categories of blocking algorithms are explained and their impact is described. Finally Section V concludes the paper.

II. RUMORS AND THEIR CHARACTERISTICS

Rumors can be defined as unverified information while posting. Rumors may or may not contain truth but cannot be trusted as they were not verified while posting. We can categorize rumors on various categories according to their veracity, durability, degree of credibility, textual content, spreading and mitigation behavior etc.

In [11], a temporal classification is given. Rumors may be (i) Long-standing rumors: which persist with significance without being established as truth, (ii) Emerging rumors: which are unseen. So this is a necessity for every rumor detection and blocking model to be able to automatically identify unseen rumors for which there is no training data available.

In [12], a consequence based classification is discussed, i.e. (i) Wish rumors: the consequence of their spread are hopes, (ii) Dread rumors: the consequences are disappointment and fear, and (iii) Wedge-driving rumors: consequences are hatred and polarization of society.
On the basis of veracity rumors can be classified as true, false and unverified [13].

III. TAXONOMY OF ALGORITHMS FOR RUMOR BLOCKING

As our behavioural psychology drives us to trust our friends' opinion on Facebook and other social networking sites, various malicious news are bombarded there, with the intention to spread negative publicity or influence harmony of the communities. In this paper we try to review all the mechanisms used for stopping such rumors and classify them differently using their impact on online social networks.

We can categorize the algorithms on the basis of nodes, links, community structure, spreading counter messages etc. We subcategorize the algorithms based on nodes and links on the basis of considering different measures for centrality and non-centrality. A single element or multiple elements can be identified on the basis of weights assigned to them. For example, in community structure, various centrality measure approaches are used to find the influential node but rarely have any approach considered the community as dynamic. Since in social networks, every individual participant is free to join or leave the community or media, the vagueness of network keeps affecting the centrality measure. This vagueness can be quantified and used as a measure. We can define the taxonomy of algorithms using categorization shown in fig. 2.

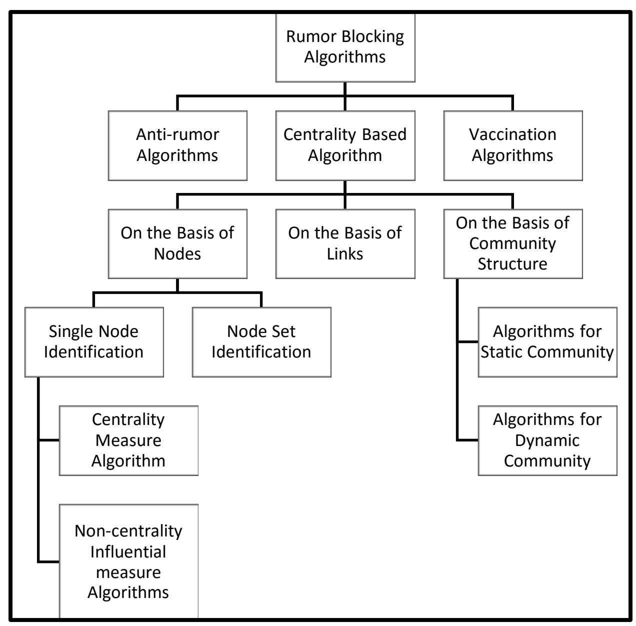

Figure 2: Taxonomy of Rumor Blocking Algorithms

IV. LITERATURE REVIEW

Apart from communication, social networking sites are used for marketing purposes by various business organizations and for promoting the content available on internet as well. In [14], the first study on influencing the customers using social network was described. Influence maximization approach given in [15] was based on greedy algorithms. Though the maximization of information flood was the first research area on which studies were focusing, the counter part of such approaches are also useful for minimizing the malignant and useless information. The following subsections brief about various approaches to reduce the rumors in social networks.

A. Centrality based approaches

In a graph, if a node or an edge has tendency of being central to the graph, its influence on the graph would be maximal. Among various centrality measures like Betweenness, degree and closeness centrality, gravity centrality, Jordan centrality and eigenvector centrality are also used to find the central entities. In [10] gravity centrality index of a node *x* is calculated using equation 1:

$$G(x) = \sum_{y\in \eta_x} \frac{ks(x)ks(y)}{d_{xy}^2} \quad \ldots (1)$$

where, *y* node is maximum *r* (radius) hops distant from *x*, d_{xy} is shortest distance between nodes *x* and *y*, and *ks* is k-shell index of a node.

In [16], Jordan centrality of a node is defined as

$$J(x) = \frac{1}{max_{y\neq x}\{d(x,y)\}} \quad \ldots (2)$$

A node which has the smallest maximum distance to other contaminated and recovered nodes is Jordan central node. There may be multiple Jordan central nodes in a graph.

Eigenvector centrality of a node can be defined as a sum of degree centrality of all its neighbors [17]. A relative centrality index of node *x* can be defined as

$$\begin{aligned} \mathcal{E}_x &= \frac{1}{c}\sum_{y\in \eta_x} \mathcal{E}_y \\ &= \frac{1}{c}\sum_{y\in G} a_{xy}\mathcal{E}_y \\ \Rightarrow Ax &= cx \quad \ldots (3) \end{aligned}$$

where, *c* is a constant and a_{xy} is the adjacency value between *x* and *y*.

1) Nodes

In [18], a diffusion model is proposed to identify both positive and negative information influencing nodes. A greedy algorithm is used to find the set of positive influences so that the impact of negative influences can be reduced. In [19], a least cost algorithm for blocking the rumors is investigated. A minimal number of 'protector' nodes are used to stop the bad influence. A protection cascade is triggered to counter the rumor cascade. A set of protector nodes is identified to minimize the rumors in the neighborhood. Both one-activate-one and one-activate-many models have been theoretically analysed in the paper. In [20], a greedy algorithm is proposed which is based on maximum marginal gain concept. The algorithm is comparable with various centrality measure approaches. It is assumed that only few nodes get affected by rumors. K-uninfected nodes are used to minimize this influence using a greedy algorithm without giving a theoretical basis. Like degree, betweenness and closeness centrality, gravity centrality is used to find the central indexing in a network [10]. This measures performed better than other centrality measures. In [21], a randomized algorithm is proposed to block the rumors using the concept of reverse tuple for sampling.

2) Edges

In [22], an edge betweenness measure is compared with the node centrality. The edge betweenness approach is proved to be more cost effective than other centrality measures. It is also stated that removal of a node makes another node important, hence discarding the node centrality measures. In [23], a contamination minimization approach is used to delete a limited number of edges in social network to stop rumors. The containment degree is calculated using influence degree of nodes. In [24], flu control problem is abstracted into edge deletion problem. The problem is shown to be super modular under the linear threshold model. A scalable algorithm is designed to delete the edge with approximation guarantees. In [25], another scalable algorithms is proposed to optimize a key graph parameter like eigenvalue to solve dissemination problems and find edges to be deleted for rumor blocking. In [8], a rumor spread minimization approach is used to reduce the rumor spread value, which is probability of a node being activated by the influential seed node. A heuristic algorithm is used to calculate the functions value. In [26], a contagion blocking approach in social networks is applied by identifying edges to be removed from a network. Various problems for contagion spread minimization are formulated and found to be efficiently solvable. In [27], limited number of links are blocked to reduce the rumor spreading. If a rumor diffuses through the network under the independent cascade model, a combinatorial optimization approach is used to identify k links to minimize this spreading. In [28], a mixed generalization model is used for infection control which uses nodes and edges into account.

3) Community Structure

In [29], two competing campaigns are studied simultaneously which spread rumor and truth in a social network. A subset of individuals is identified which spread truth such that majority of individual adopt the true news rather than being misled by rumors. This problem is shown to be NP-hard and a greedy algorithm is used to solve approximate results. In [21], a randomized approximation algorithm is used to identify *k* seed users to trigger the spread of a positive cascaded information to maximize the number of users not influenced by rumors. In [30], an algorithm to reduce rumors in dynamic graphs is proposed. In this algorithm, seed users are selected using an adaptive strategy using dynamic independent cascade model.

B. Anti-Rumor based approaches

Rumor can be combatted with rumor-like anti-rumor messages. Various rumor spreading models like Susceptible Infected Recovered model, Susceptible Infected Susceptible model and Independent Cascade (IC) model are used for spreading anti-rumors. In [31][32], three i*ndependent cascade model* based anti-rumor approaches were proposed to control the rumors from spreading throughout the network. The first model called *Delayed Start Model* assumes that there is a delay between a node being infected and its detection. From the pool of all infected nodes at a moment only one node is selected to spread the anti-rumor message. The message propagates through the neighboring nodes. The second model named *Beacon Model* selects multiple agent nodes called beacons on the basis of connectivity, trust, and specific topics and makes them in charge of spreading anti-rumor message, thus reduces the delay and speeds up the anti-rumor spreading. Finally the *Neighborhood Model* allows any node take decisions regarding combating rumors. It was analysed that the rumor spreading gets reduced by 60% using the Beacon model.

In [33], an IBM problem of rumor minimization is solved by finding the positive influencers and diffusing positive information to block negative ones. A cascaded simultaneous diffusion called *Multi-Campaign Independent Cascade Model* is used to achieve this. In [29], a counter-campaign is used to limit the rumors using predictive hill climbing approach. Given the states of few nodes, states of all the nodes are predicted first and then hill climbing approach is used to choose the influences. The algorithm has the ability to learn parameters with missing data as well. This problem is proved to be NP-Hard and a feasible exact solution is not possible. In [34], linear threshold and independent cascade were combined to create a model. It was concluded that in large communities few highly influential nodes should be used to spread true information. In case of large number of influential nodes selection, the method is capable of minimizing rumor spread in small communities.

C. Vaccination approaches

In [35] a social vaccination mechanism called Pulse Vaccination for Rumor Control is discussed, which is borrowed from epidemic model. They identify the vaccination population at every time interval for repeated vaccination and thus control the rumor. Their assumption is that a vaccinated user can not catch or spread a particular rumor again. They have considered the dynamic nature of social networks and identify different groups of influencing individuals after every time interval. The influencing capability of a node has been found using equation 2

$$IC_x(t) = \alpha(t) + |N_x| + \sum_{y \in N_x} \frac{\rho_{xy}}{<\rho>} \qquad \text{... (2)}$$

where, $IC_x(t)$ is the influential capability of individual node x at time t, $\alpha(t)$ is rumor attraction factor, ρ_{xy} is the probability of acceptance of node y on x, N_x is a set of nodes directly connected to x.

V. CONCLUSION

In this paper, we briefly studied various rumor diffusion and blocking mechanisms. Though, blocking rumors or maximizing truth is a NP-Hard problem, most of the algorithms use greedy approach to solve it. The centrality measure is used to identify seed nodes and influential links in almost all the approaches. Counter-campaigning is an effective way to stop rumors. This approach advertises the positive news to reduce the flow of negative information. There is still a need to find effective rumor blocking algorithms for dynamic social networks.

REFERENCES

[1] Harris, Lisa, and Alan Rae. *"Social networks: the future of marketing for small business."* Journal of business strategy 2009.

[2] Watts, Duncan J. *"The "new" science of networks."* Annu. Rev. Sociol. 30 (**2004**): **243-270.**

[3] Verbeke, Wouter, David Martens, and Bart Baesens. *"Social network analysis for customer churn prediction."* Applied Soft Computing 14 (**2014): 431-446.**

[4] Centola, Damon. *"The spread of behavior in an online social network experiment."* science 329, no. 5996 (**2010): 1194-1197.**

[5] Ghosh, Rumi, and Kristina Lerman. *"Predicting influential users in online social networks."* arXiv preprint arXiv:1005.4882 (2010).

[6] Bargar, Alicia, Stephanie Pitts, Janis Butkevics, and Ian McCulloh. *"Challenges and Opportunities to Counter Information Operations Through Social Network Analysis and Theory."* In 2019 11th International Conference on Cyber Conflict (CyCon), vol**. 900**, pp. 1-18. IEEE, **2019.**

[7] Ohara, Kouzou, Kazumi Saito, Masahiro Kimura, and Hiroshi Motoda. *"Critical Node Identification based on Articulation Point Detection for Uncertain Network."* International Journal of Networking and Computing 9, no. 2 (**2019): 201-216.**

[8] Yan, Ruidong, Yi Li, Weili Wu, Deying Li, and Yongcai Wang. *"Rumor blocking through online link deletion on social networks."* ACM Transactions on Knowledge Discovery from Data (TKDD) 13, no. 2 (**2019): 1-26.**

[9] Zhao, Yuxin, Shenghong Li, and Feng Jin. *"Identification of influential nodes in social networks with community structure based on label propagation."* Neurocomputing 210 (2016): 34-44.

[10] Ma, Ling-ling, Chuang Ma, Hai-Feng Zhang, and Bing-Hong Wang. *"Identifying influential spreaders in complex networks based on gravity formula."* Physica A: Statistical Mechanics and its Applications 451 (**2016): 205-212.**

[11] Zubiaga, Arkaitz, Ahmet Aker, Kalina Bontcheva, Maria Liakata, and Rob Procter. *"Detection and resolution of rumours in social media: A survey."* ACM Computing Surveys (CSUR) 51, no. 2 (**2018): 1-36.**

[12] Robert H. Knapp. *"A psychology of rumor. "* Public Opin. Q. 8, 1 (**1944), 22–37.**

[13] Zubiaga, Arkaitz, Maria Liakata, Rob Procter, Geraldine Wong Sak Hoi, and Peter Tolmie. *"Analysing how people orient to and spread*

rumours in social media by looking at conversational threads." PloS one 11, no. 3 (**2016).**

[14] Domingos, Pedro, and Matt Richardson. *"Mining the network value of customers."* In Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, pp. **57-66. 2001.**

[15] Kempe, David, Jon Kleinberg, and Éva Tardos. *"Maximizing the spread of influence through a social network."* In Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, pp. **137-146. 2003.**

[16] Dekker, A.H., 2008. *"Centrality in social networks: Theoretical and simulation approaches."* Proceedings of SimTecT **(2008), pp.12-15.**

[17] Bonacich, Phillip. *"Power and centrality: A family of measures."* American journal of sociology 92, no. 5 (**1987): 1170-1182.**

[18] Kaur, Harneet, and Jing He. *"Blocking negative influential node set in social networks: from host perspective."* Transactions on Emerging Telecommunications Technologies 28, no. 4 (**2017**): e3007.

[19] Arazkhani, Niloofar, Mohammad Reza Meybodi, and Alireza Rezvanian. *"Influence Blocking Maximization in Social Network Using Centrality Measures."* In 2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI), pp. **492-497**. IEEE, **2019.**

[20] Budak, Ceren, Divyakant Agrawal, and Amr El Abbadi. *"Limiting the spread of misinformation in social networks."* In Proceedings of the 20th international conference on World wide web, pp. **665-674.** ACM, **2011.**

[21] Tong, Guangmo, Weili Wu, Ling Guo, Deying Li, Cong Liu, Bin Liu, and Ding-Zhu Du. *"An efficient randomized algorithm for rumor blocking in online social networks."* IEEE Transactions on Network Science and Engineering (**2017).**

[22] Dey, Paramita, and Sarbani Roy. *"Centrality based information blocking and influence minimization in online social network."* In 2017 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), pp. **1-6.** IEEE, **2017.**

[23] M. Kimura, K. Saito, and H. Motoda, *"Minimizing the spread of contamination by blocking links in a network."* in AAAI, vol. **8**, **2008**, pp. 1175–1180.

[24] E. B. Khalil, B. Dilkina, and L. Song, *"Scalable diffusion-aware optimization of network topology,"* in Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, **2014, pp. 1226–1235.**

[25] H. Tong, B. A. Prakash, T. Eliassi-Rad, M. Faloutsos, and C. Faloutsos, *"Gelling, and melting, large graphs by edge manipulation,"* in Proceedings of the 21st ACM international conference on Information and knowledge management. ACM, **2012, pp. 245–254.**

[26] Kuhlman, Chris J., Gaurav Tuli, Samarth Swarup, Madhav V. Marathe, and S. S. Ravi. *"Blocking simple and complex contagion by edge removal."* In 2013 IEEE 13th International Conference on Data Mining, pp**. 399-408**. IEEE, **2013.**

[27] Yao, Qipeng, Chuan Zhou, Linbo Xiang, Yanan Cao, and Li Guo. *"Minimizing the negative influence by blocking links in social networks."* In International conference on trustworthy computing and services, pp. 65-73. Springer, Berlin, Heidelberg, **2014.**

[28] He, Jing, Hongyu Liang, and Hao Yuan. *"Controlling infection by blocking nodes and links simultaneously."* In International workshop on internet and network economics, pp. **206-217**. Springer, Berlin, Heidelberg, **2011.**

[29] S. Wang, X. Zhao, Y. Chen, Z. Li, K. Zhang, and J. Xia, *"Negative influence minimizing by blocking nodes in social networks."* In AAAI (Late-Breaking Developments), **2013, pp. 134–136.**

[30] Tong, Guangmo, Weili Wu, Shaojie Tang, and Ding-Zhu Du. *"Adaptive influence maximization in dynamic social networks."* IEEE/ACM Transactions on Networking (TON) 25, no. 1 (**2017): 112-125.**

[31] Tripathy, R.M., Bagchi, A. and Mehta, S., 2010, October. *"A study of rumor control strategies on social networks."* In Proceedings of the 19th ACM international conference on Information and knowledge management (pp. **1817-1820).**

[32] Tripathy, R.M., Bagchi, A. and Mehta, S., 2013. *"Towards combating rumors in social networks: Models and metrics."* Intelligent Data Analysis, 17(1), pp.**149-175.**

[33] L. Fan, Z. Lu, W. Wu, B. Thuraisingham, H. Ma, and Y. Bi, *"Least cost rumor blocking in social networks,"* in Distributed Computing Systems (ICDCS), 2013 IEEE 33rd International Conference on. IEEE, **2013, pp. 540–549.**

[34] Nguyen, N.P., Yan, G., Thai, M.T. and Eidenbenz, S., 2012, June. *"Containment of misinformation spread in online social networks."* In Proceedings of the 4th Annual ACM Web Science Conference (pp. **213-222).**

[35] Santhoshkumar, S. and Babu, L.D., 2019. *"An Effective Rumor Control Approach for Online Social Networks."* In Information Systems Design and Intelligent Applications (pp. 63-73). Springer, Singapore.

AUTHORS PROFILE

P. K. Tiwari completed his Bachelor of Technology in Information Technology from Dr. A.P.J. Abdul Kalam Technical University, Lucknow, Uttar Pradesh in year 2012. He is currently pursuing Master of Technology (CSE) from Maharishi University, Lucknow under the supervision of Professor A. K. Bharti and working as an Assistant Professor in the Department of Computer Sciences, IEC Group of Institution, Greater Noida since Sep 2019. Previously he worked as a lecturer in Accurate Institute of Management and Technology, Greater Noida from 2012 to 2019 making it approximately 7 years of academic experience. His main research work focuses on social network mining, web mining and literature mining etc.

M. K. Singh pursed his Bachelor of Technology in Information Technology from Dr. A.P.J. Abdul Kalam Technical University, Lucknow, Uttar Pradesh in year 2005 and Master of Technology in Computer Science and Engineering from Jamia Hamdard University, New Delhi in year 2012. He is currently working as an Assistant Professor in the Department of Computer Science, IEC College of Engineering and Technology, Greater Noida since 2017. Previously he has worked as a lecturer and assistant professor in Institutes of state technical university since 2005; making it approximately 14 years of academic experience. He has published many research papers in reputed journals and conferences. His main research work focuses on algorithm, wireless networking and Machine Learning etc.

Dr A. K. Bharti is working as professor and Dean of Maharishi School of computer science. He has over 17 years of rich experience in Research, Education and Industry He has rich experience in IT, OS, DBMS, Computer Graphics and computer networks etc. He was doctorate from NAAC 'A' grade Central University Babasaheb Bhimrao Ambedkar University Lucknow. He completed his Masters in Computer Application from state engineering college KNIT, Sultanpur, Graduation in B.Sc. (hon.) in mathematics from the leading Banaras Hindu University (BHU), Varanasi.

