International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 08, August 2026)

A Survey for Analysis of Optimization – Based Task Scheduling Algorithms in Cloud Computing

Dr. R. Pragaladan[1], R. Jamunarani [2]

[1]Head & Associate Professor, [2]Ph.D Research Scholar, Department of Computer Science, Sri Vasavi College, Erode, India

***Abstract*— Cloud computing provides scalable and on-demand computing resources through the internet. Efficient task scheduling is essential for improving resource utilization and reducing execution time in cloud environments. Numerous optimization-based algorithms have been proposed to solve the complex task allocation problem. This survey reviews existing research on task scheduling algorithms in cloud computing and classifies them into traditional, heuristic, and metaheuristic approaches. Widely used optimization techniques such as Particle Swarm Optimization, Ant Colony Optimization, and Genetic Algorithm are analyzed in terms of performance metrics including makespan, resource utilization, and execution cost. A comparative analysis of existing approaches is presented to highlight their advantages and limitations. Furthermore, the survey identifies research gaps in cost-aware scheduling, energy-efficient resource allocation, and scalable cloud scheduling techniques. Finally, future research directions are discussed to guide the development of efficient and lightweight scheduling algorithms for next-generation cloud computing environments.**

***Keywords* — Cloud computing, efficient task scheduling, make span, resource utilization, execution cost.**

I. INTRODUCTION

Cloud computing has emerged as a transformative computing paradigm that provides flexible and scalable computing resources over the internet. It enables organizations and individuals to access computational power, storage, and applications without maintaining physical infrastructure. The rapid growth of cloud services has resulted in large-scale data centers that host thousands.

Virtual machines and handle millions of user requests. One of the major challenges in cloud computing environments is efficient resource management complex optimization problems in distributed systems. This survey provides a comprehensive review of optimization-based task scheduling algorithms used in cloud computing systems. The study analyzes various approaches, compares their performance, and identifies existing research gaps.

II. LITERATURE REVIEW

Cloud computing has evolved into a powerful computing model that offers flexible, on-demand access to a shared pool of configurable resources over the internet.

One of the central challenges in this environment is task scheduling, which focuses on assigning tasks to virtual machines (VMs) in a way that optimizes system performance. Due to its complexity and combinatorial nature, task scheduling is classified as an NP-hard problem, making it difficult to solve efficiently using conventional methods, especially in large-scale and dynamic cloud environments.

This study provides a detailed review of 22 research articles that investigate various optimization-driven scheduling techniques. The discussion is organized into key thematic areas, including traditional scheduling methods, swarm intelligence approaches, evolutionary and genetic algorithms, hybrid optimization strategies, and advanced techniques such as meta heuristics and reinforcement learning. Each category is evaluated based on its methodology, objectives, and effectiveness.

The major performance goals identified across these studies include reducing makespan, minimizing operational costs, improving resource utilization, and ensuring compliance with Quality of Service (QoS) and Service Level Agreements (SLA). Simulation tools like CloudSim are widely used to assess algorithm performance. Recent trends highlight a growing interest in hybrid models and quantum-inspired algorithms, which demonstrate better scalability and faster convergence compared to traditional techniques. Emerging research directions include energy-efficient scheduling, integration with fog and edge computing, and the use of machine learning for predictive decision-making.

Task Scheduling and Optimization in Cloud Computing

Cloud computing delivers computing resources as scalable and metered services via the internet. Within this framework, task scheduling plays a crucial role in distributing workloads (cloudlets) across available resources efficiently and within acceptable time limits.

Because task scheduling problems are NP-hard, classical deterministic approaches often fail to provide optimal or near-optimal solutions in complex and heterogeneous systems. Poor scheduling decisions can lead to either underutilization or overloading of resources, both of which degrade system performance and QoS.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 08, August 2026)

To overcome these limitations, optimization-based approaches aim to enhance throughput, reduce execution delays, and improve overall system efficiency. Modern research increasingly relies on metaheuristic techniques, artificial intelligence, and hybrid models to address these challenges effectively.

An adaptive workflow scheduling model based on convex optimization has been proposed to improve task execution efficiency. This approach enhances resource allocation and load balancing while reducing execution time. Although it performs better than traditional methods, it may struggle in highly dynamic or real-time environments

1. Traditional Scheduling Techniques

Basic scheduling methods form the foundation for more advanced optimization strategies and are often used for comparison.

- *Enhanced Shortest Job First (SJF):*

This improved version incorporates dynamic priority adjustment, workload prediction, and task pre-emption to avoid starvation and improve efficiency.

- *Deadline-Oriented Scheduling:*

Tasks are assigned based on their deadlines, leading to improved response time and reduced makespan.

- *Conventional Algorithms (FCFS and RR):*

These simple techniques process tasks sequentially or using time slices but are less efficient in complex cloud environments.

- *Push–Pull Scheduling Mechanism:*

Tasks are dynamically migrated between VMs to balance load and meet QoS requirements.

- *Linked List-Based Push–Pull (LPP):*

This method enhances task migration using structured data handling, resulting in better SLA compliance and reduced response time.

2. Swarm Intelligence-Based Methods

Swarm intelligence techniques are inspired by the collective behavior of natural systems and are widely used for optimization problems.

- *Particle Swarm Optimization (PSO):* Uses particle movement strategies to find optimal task-resource mappings.
- *Ant Colony Optimization (ACO):* Mimics ant foraging behavior to determine efficient execution paths.
- *Bee Colony Optimization (BCO):* Applies cooperative search strategies to optimize resource allocation.
- *Grey Wolf Optimizer (GWO):* Provides effective global search capabilities for complex optimization problems.
- Additional approaches include the Slime Mold Algorithm (SMA), Whale Optimization Algorithm (WOA), and Equilibrium Optimizer (EO), which help improve convergence and avoid local optima.

3. Evolutionary and Genetic Approaches

These algorithms are based on biological evolution principles and are effective for solving complex optimization problems.

- *Genetic Algorithm (GA):* Uses selection, crossover, and mutation operations to optimize scheduling objectives such as time and cost.
- *Differential Evolution (DE):* Efficiently generates near-optimal solutions for complex scheduling problems.
- *Hybrid Differential Evolution (HDE):* Enhances standard DE by improving convergence speed and solution quality.
- *Quantum-Inspired Evolutionary Algorithms (FOG-QIEA):* Utilize quantum computing concepts to optimize multiple objectives, including energy consumption, execution time, and cost, with faster convergence rates.

4. Hybrid Optimization Strategies

Hybrid approaches combine multiple algorithms to leverage their strengths and overcome individual limitations

- *PSO-GWO Hybrid Model:* Integrates exploration and exploitation capabilities to improve performance metrics such as makespan and throughput.
- *PSOMC:* Combines PSO with membrane computing to increase population diversity and avoid local optima.
- *Genetic Hill Climbing (GHCA):* Improves resource utilization and reduces cost through iterative refinement.
- *Quantum-Based Dynamic Scheduler (QBDS):* Uses quantum-inspired techniques to optimize multiple objectives simultaneously.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 08, August 2026)

5. Advanced and Intelligent Scheduling Techniques

- *Reinforcement Learning-Based Scheduling (RL-MOTS):* Applies Deep Q-Networks to dynamically allocate resources based on real-time conditions.
- *Security-Aware Scheduling (SeSiC):* Considers both security constraints and deadlines to improve user satisfaction.
- *IRMTS:* Enhances load balancing and service quality in software-defined cloud environments.
- *Big Data Scheduling:* Uses mathematical optimization models to handle large-scale data processing efficiently.
- *Predictive Scheduling Approaches:* Utilize historical data to forecast resource requirements and ensure task completion reliability.

III. Cloud Computing Overview

Cloud computing is a modern computing paradigm in which computing resources and services are delivered over the internet in a distributed manner. These services typically include processing capabilities, data storage, networking infrastructure, and software applications, all of which can be accessed remotely without direct physical management.

Key Characteristics of Cloud Computing

1. On-Demand Self service

Users can independently provision computing resources such as servers, storage, and applications whenever required, without needing assistance from the service provider.

2.Broad Network Accessibility

Cloud services are accessible via the internet and can be used across multiple devices, including laptops, mobile phones, and tablets, through standard communication protocols.

3.Resource Pooling

Service providers maintain a shared pool of computing resources that serve multiple users simultaneously. These resources are dynamically allocated and reassigned based on user demand, following a multi-tenant model.

4.Rapid Elasticity

Cloud environments allow resources to be scaled up or down quickly in response to changing workload demands, ensuring flexibility and efficient handling of varying usage levels.

5.Measured Service

Resource consumption is continuously monitored, enabling a pay-per-use model where users are billed based on their actual usage.

Cloud Service Models

- *Infrastructure as a Service (IaaS)*

This model provides fundamental computing resources such as virtual machines, storage systems, and networking components. Users can manage operating systems and applications while the cloud provider handles the physical infrastructure.

- *Platform as a Service (PaaS)*

PaaS offers a complete development environment, including tools and frameworks for building, testing, and deploying applications. It eliminates the need to manage underlying hardware and system software.

- *Software as a Service (SaaS)*

In this model, software applications are delivered over the internet and accessed via web browsers. Users do not need to install or maintain the software locally, as all management is handled by the provider.

IV. Task Scheduling Cloud Computing

Task scheduling in cloud computing refers to the systematic allocation of user-submitted tasks to available computing resources, such as virtual machines, in an efficient manner. It plays a crucial role in ensuring optimal system performance and effective resource management.

1) Objectives of Task Scheduling
2) Reduce overall completion time (makespan)
3) Minimize task execution time
4) Maximize utilization of available resources
5) Maintain balanced workload distribution across resources
6) Lower operational and computational costs
7) Challenges in Task Scheduling
8) Dynamic Workloads
9) Task arrival rates and resource demands can change unpredictably, making scheduling decisions more complex.
10) Heterogeneous Resources
11) Cloud environments consist of diverse resources with varying capabilities, which complicates efficient allocation.

12) Scalability Issues
13) Managing scheduling in large-scale distributed systems requires algorithms that can handle high volumes of tasks and resources efficiently.
14) Energy Efficiency Constraints

Reducing energy consumption while maintaining performance is a major challenge in modern cloud data centres.

V. Classification Of Scheduling Algorithms

In cloud computing, scheduling algorithms are essential for efficiently allocating tasks to available resources. These algorithms can be broadly classified into different categories based on their design approach and complexity.

1. Traditional Algorithms

Traditional scheduling methods are basic techniques that follow simple rules for task allocation. They are easy to implement but often fail to perform efficiently in complex and dynamic cloud environments.

Examples include:

- *First Come First Served (FCFS):* Tasks are executed in the order they arrive.
- *Round Robin (RR):* Tasks are assigned time slices in a cyclic manner.
- *Shortest Job First (SJF):* Tasks with the shortest execution time are prioritized.

Although these methods are straightforward, they are not well-suited for large-scale systems with diverse workloads.

2. Heuristic Algorithms

Heuristic-based approaches are designed to produce reasonably good solutions within a short time. They do not guarantee optimal results but are effective for practical scheduling problems.

Common examples:

- *Min-Min Algorithm:* Selects tasks with the minimum completion time and assigns them first.
- *Max-Min Algorithm:* Prioritizes tasks with the maximum execution time to balance workload distribution.

These techniques improve performance compared to traditional methods but may still struggle with highly complex scenarios.

3. Metaheuristic Algorithms

Metaheuristic algorithms are advanced optimization techniques that are capable of exploring large solution spaces to find near-optimal results. They are particularly useful for solving complex scheduling problems in cloud environments.

Examples include:

- *Particle Swarm Optimization (PSO):* Uses a population of particles to search for optimal solutions.
- *Ant Colony Optimization (ACO):* Inspired by ant foraging behavior to determine efficient task paths.
- *Bee Colony Optimization (BCO):* Mimics the food-searching behavior of bees for resource allocation.

These methods are highly effective in improving scheduling efficiency and avoiding local optima.

4. Evolutionary Algorithms

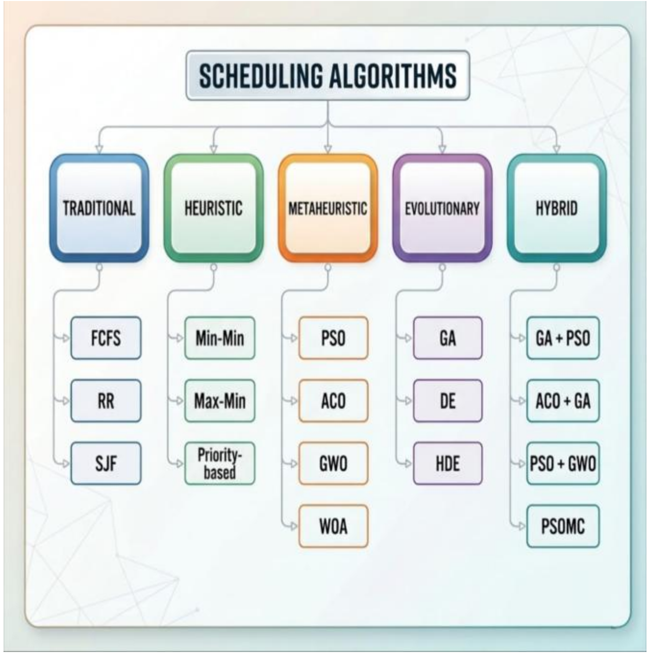
Evolutionary algorithms are inspired by natural selection and genetic processes. They iteratively improve solutions through operations such as selection, mutation, and crossover.

Examples include:

- *Genetic Algorithms (GA):* Use population-based search and genetic operations to optimize scheduling objectives.
- *Differential Evolution (DE):* Focuses on improving candidate solutions through vector-based mutation and recombination.

These algorithms are well-suited for handling complex, multi-objective scheduling problems.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 08, August 2026)

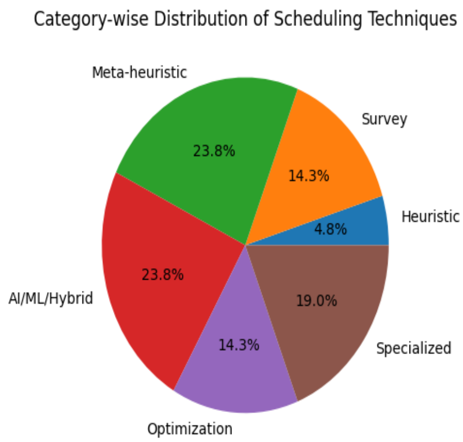


VI. Comparative Analysis

Ref	Technique / Algorithm	Category	Objective	Key Advantages	Limitations
[1]	Enhanced Shortest Job First (SJF)	Heuristic	Minimize waiting time & response time	Simple, improves execution time	Starvation of long tasks
[2]	SDN-based Scheduling	Network-aware	Load balancing & traffic optimization	Better network utilization	Complex architecture
[3]	General Survey (2018)	Survey	Overview of scheduling methods	Broad coverage	No implementation
[4]	Scheduling Review (2021)	Survey	Classification of algorithms	Updated taxonomy	Lacks experimental validation
[6]	Optimization Algorithms Study	Meta-heuristic	Improve makespan & cost	Covers multiple techniques	No unified framework
[7]	Hybrid AI-based Scheduling	AI / Hybrid	Intelligent decision making	Better adaptability	High computational cost
[8]	Differential Evolution (DE)	Evolutionary	Optimize task allocation	Good global search ability	Slow convergence
[9]	Meta-heuristic Survey	Meta-heuristic	Classification & challenges	Identifies open problems	No practical validation
[10]	PSO + Grey Wolf Optimization	Hybrid Meta-heuristic	Minimize makespan & cost	Improved convergence	Parameter tuning complexity

[11]	AI + Meta-heuristic	AI-based	Resource scheduling	Intelligent optimization	High overhead
[12]	Security-aware Scheduling	Security-based	Improve user satisfaction & safety	Enhances trust	Increases execution time
[13]	Adaptiveand Convex Otimization-based Workflow Scheduling	Optimization-based (Convex + Adaptive)	Improve workflow scheduling efficiency and resource allocation	Enhances resource utilization, improves load balancing	less focus on real-time adaptation
[14]	PSOMC Algorithm	Swarm Intelligence	Task scheduling optimization	Efficient resource usage	Local optima problem
[15]	Cost-aware Scheduling	Economic model	Minimize cost in heterogeneous cloud	Cost-efficient	Ignores QoS factors
[16]	Multi-cloud Scheduling	Distributed	Optimize multi-cloud performance	Scalability	Coordination overhead
[17]	IoT Fog-Cloud Scheduling	Edge/Fog	Real-time task handling	Low latency	Resource constraints
[18]	Quantum-inspired Scheduler	Advanced optimization	Multi-objective scheduling	High optimization capability	Theoretical complexity
[19]	Reinforcement Learning	AI/ML	Energy & cost optimization	Adaptive learning	Training time high
[20]	Improved PSO	Swarm Intelligence	Efficient task scheduling	Faster convergence	Risk of premature convergence
21]	Predictive Scheduling	ML-based	Resource availability prediction	Improves reliability	Requires historical data
5],[22]	Push-Pull Optimization (LPP)	Hybrid heuristic	Deadline-aware scheduling	Reduces SLA violations	Limited scalability analysis

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 08, August 2026)



VII. CONCLUSION

Efficient task scheduling plays a crucial role in enhancing both resource utilization and overall system performance within cloud computing environments. This survey examined a range of optimization-based scheduling techniques and evaluated their respective advantages and limitations. The analysis shows that advanced approaches, particularly metaheuristic and hybrid algorithms, outperform conventional scheduling methods in terms of efficiency and adaptability. Despite these improvements, challenges such as scalability, cost management, and handling dynamic workloads still persist. Therefore, future research should prioritize the development of lightweight and adaptive scheduling strategies that can efficiently operate in large-scale and rapidly changing cloud environments.

REFERENCES

[1] Y. Pachipala, K. S. Sureddy, A. B. S. S. Kaitepalli, N. Pagadala, S. S. Nalabothu, and M. Iniganti, "Optimizing task scheduling in cloud computing: An enhanced shortest job first algorithm," Procedia Computer Science, vol. 233, pp. 604–613, 2024.

[2] M. Mahdizadeh, A. Montazerolghaem, and K. Jamshidi, "Task scheduling and load balancing in SDN-based cloud computing: A review of relevant research," Journal of Engineering Research, 2024, doi: 10.1016/j.jer.2024.11.002

[3] Aarti and S. Kaur, "A survey for task scheduling in cloud computing," International Journal of Current Research, vol. 10, no. 9, pp. 73660–73664, 2018.

[4] I. M. Ibrahim, A. H. Radie, K. Jacksi, S. R. M. Zeebaree, H. M. Shukur, Z. N. Rashid, M. A. M. Sadeeq, and H. M. Yasin, "Task scheduling algorithms in cloud computing: A review," Turkish Journal of Computer and Mathematics Education, vol. 12, no. 4, pp. 1041–1053, 2021

[5] R. B. S., M. Namratha, N. Srikanth, N. Halatti, and P. Srinivasan, "A novel push–pull optimization algorithmfostered deadline-aware task scheduling in virtualized cloud computing," Franklin Open, vol. 14, Art. no. 100499, 2026, doi: 10.1016/j.fraope.2026.100499.

[6] G. S. L. Namitha and B. M. Siraparapu, "A study on various optimization algorithms for task scheduling in cloud computing," International Journal of Advance Research, Ideas and Innovations in Technology, vol. 7, no. 4, 2021.

[7] V. R. S. P. Alla, N. R. Medikondu, L. S. Parige, K. Satyanarayana, V. S.

[8] Kankhva, N. Dhaliwal, and A. K. Saxena, "Optimizing task scheduling in cloud computing: A hybrid artificial intelligence approach," Cogent Engineering, vol. 11, no. 1, Art. no. 2328355, 2024, doi: 10.1080/23311916.2024.2328355.7.

[9] M. Abdel-Basset, R. Mohamed, W. A. Elkhalik, M. Sharawi, and K. M. Sallam, "Task scheduling approach in cloud computing environment using hybrid differential evolution," Mathematics, vol. 10, no. 21, Art. no. 4049, 2022, doi: 10.3390/math10214049.

[10] E. H. Houssein, A. G. Gad, Y. M. Wazery, and P. N. Suganthan, "Task scheduling in cloud computing based on meta-heuristics: Review, taxonomy, open challenges, and future trends," Swarm and Evolutionary Computation, vol. 62, Art. no. 100841, 2021, doi: 10.1016/j.swevo.2021.100841.

[11] R. Prasad, A. Roy, and S. Kumari, "Enhancing cloud task scheduling using a hybrid particle swarm and grey wolf optimization approach," arXiv preprint, arXiv:2505.15171v1 [cs.DC], May 2025.

[12] R. Arora and A. Abraham, "Resource scheduling methods for cloud computing environment: The role of meta-heuristics and artificial intelligence," Engineering Applications of Artificial Intelligence, vol. 114, Art. no. 105345, 2022, doi: 10.1016/j.engappai.2022.105345. [12] B. Wang, J. Cao, C. Wang, W. Huang, and Y. Song, "Security-aware task scheduling for improving the user satisfaction in hybrid clouds," Journal of Parallel and Distributed Computing, 2026, doi: 10.1016/j.jpdc.2026.105252.

[13] K. Lakhwani, G. Sharma, R. Sandhu, N. K. Nagwani, S. Bhargava, V. Arya, and A. Almomani, "Adaptive and convex optimization inspired workflow scheduling for cloud environment," International Journal of Cloud Applications and Computing, vol. 13, no. 1, 2023.

[14] K. Li, L. Jia, and X. Shi, "Research on cloud computing task scheduling based on PSOMC," Journal of Web Engineering, vol. 21, no. 6, pp. 1749–1766, 2022, doi: 10.13052/jwe1540-9589.2161

[15] A. Ayyadi and A. Jahani, "Addressing cost and resource variability for big data task scheduling in heterogeneous cloud environments," The Journal of Supercomputing, vol. 81, p. 762, 2025, doi: 10.1007/s11227-025-07248-0.

[16] Q. Zhang, S. Geng, and X. Cai, "Survey on task scheduling optimization strategy under multi-cloud environment," Computer Modeling in Engineering & Sciences, vol. 136, no. 3, pp. 2703–2728, 2023, doi: 10.32604/cmes.2023.022287.

[17] D. Bradeen and D. Tong, "Intelligent task scheduling framework for cloud platforms based on improved particle swarm optimization," Journal of Computer Technology and Software, vol. 5, no. 1, 2026.

[18] A. Mindil, A. Y. Hamed, M. R. Hassan, and M. K. Elnahary, "A novel approach for dynamic task scheduling for IoT in fog-cloud environment," Scientific Reports, vol. 16, Art. no. 5501, 2026, doi: 10.1038/s41598-026-35156-7

[19] N. G. Hammouda, M. Shalaby, R. H. C. Alfilh, and N. S. S. Singh, “A quantum-driven multi-objective scheduler for scalable task orchestration in fog-based cyber-physical-social systems,” Scientific Reports, vol. 16, Art. no. 6874, 2026, doi: 10.1038/s41598-025-33627-x.

[20] X. Yu, J. Mi, L. Tang, L. Long, and X. Qin, “Dynamic multi-objective task scheduling in cloud computing using reinforcement learning for energy and cost optimization,” Scientific Reports, vol. 15, Art. no. 45387, 2025, doi: 10.1038/s41598-025-29280-z.

[21] D. Bradeen and D. Tong, “Intelligent task scheduling framework for cloud platforms based on improved particle swarm optimization,” Journal of Computer Technology and Software, vol. 5, no. 1, 2026.

[22] U. Sen, M. Sarkar, and N. Mukherjee, “A predictive and availability-aware job scheduling algorithm for resource management in cloud,” International Journal of Computing, vol. 24, no. 4, pp. 717–726, 2025, doi: 10.47839/ijc.24.4.4337.”

[23] R.M.Aravind and R.Pragaladan, “ A task scheduling in the cloud computing environment based on bacterial colony optimisation”, International Journal of Communication Networks and Distributed Systems, pp-591-613, vol 31, No 6, 2025.

[24] Dr.Nirmaladevi Krishnan, “Big data and fuzzy logic for demand forecasting in supply chain management: A data-driven approach”, Journal of fuzzy extension and applications 6 (2), 260-283, 2025.

[25] Nirmaladevi, K., and K. Prabha. "A selfish node trust aware with Optimized Clustering for reliable routing protocol in Manet." Measurement: Sensors 26 (2023): 100680.