China Petroleum Processing and Petrochemical Technology

Catalyst Research Volume 23, Issue 2, October 2023 Pp. 2100-2117

ANALYTICAL REPORT ON METAHEURISTIC AND NON-METAHEURISTIC ALGORITHMS FOR CLUSTERING IN WIRELESS NETWORKS

Nishant Tripathi[1], Charanjeet Singh[2], and Kamal Kumar Sharma[3]

Abstract – This analytical report delves into the comprehensive evaluation of both metaheuristic and non-metaheuristic algorithms utilized for clustering in wireless networks. Clustering techniques play a pivotal role in enhancing the efficiency and performance of wireless networks by organizing nodes into meaningful groups. Metaheuristic algorithms, inspired by natural processes, offer innovative solutions to complex optimization problems, while non-metaheuristic algorithms rely on traditional mathematical principles. This report systematically compares and contrasts the efficacy of various algorithms, considering key metrics such as convergence speed, scalability, robustness, and adaptability to dynamic network conditions. By scrutinizing both categories of algorithms, this report aims to provide a holistic understanding of their respective advantages, limitations, and applicability in wireless network clustering scenarios. The insights derived from this analysis can guide network engineers, researchers, and practitioners in selecting the most suitable algorithms based on specific network requirements, ultimately contributing to the advancement of wireless network clustering techniques.

Keywords **–** Wireless Networks, Clustering, Metaheuristic Algorithm, Low Energy Adaptive Clustering Hierarchy

1. INTRODUCTION

Clustering plays a vital role in Wireless Networks as it enables efficient data organization and communication among nodes. This analytical report provides an in-depth exploration of two main categories of clustering algorithms: metaheuristic and non-metaheuristic, and their applications in Wireless Networks. [1] The report examines the implementation methods, clustering approaches, efficiency, mathematical formulations (if any), and complexities of various algorithms, including K-Means, DBSCAN, Hierarchical Clustering, Particle Swarm Optimization (PSO), Genetic Algorithm (GA), and more. Non-metaheuristic algorithms, such as K-Means and DBSCAN, are simple, efficient, and commonly used in Wireless Networks, while metaheuristic algorithms, like PSO and GA, offer global optimization capabilities but come with higher computational costs. Moreover, soft clustering algorithms, like Fuzzy C-Means and Expectation-Maximization, allow

[1]Dept. of Electronics and Communication Engineering, School of Electronics and Electrical Engineering, Lovely Professional University, Phagwara, Punjab (India).

*Dept. of Electronics and Communication Engineering, Pranveer Singh Institute of Technology Kanpur.

[2]Dept. of Electronics and Communication Engineering, School of Electronics and Electrical Engineering, Lovely Professional University, Phagwara, Punjab (India)

[3]Ambala College Of Engineering and Applied Research Devsthali, Ambala Cantt – Jagadhari Road, P.O. Sambhalkha Ambala Haryana, India

[1] Corresponding author;
E-mail: rcharanjeet@gmail.com

for flexible data membership assignments, enhancing uncertainty modelling. The critical evaluation highlights trade-offs between efficiency, accuracy, and complexity, emphasizing the importance of selecting appropriate algorithms based on network size, data characteristics, and energy constraints. [1], [2] Clustering is a fundamental task in data mining and machine learning, aimed at grouping similar data points together in a dataset. Metaheuristic and non-metaheuristic algorithms are two broad categories of approaches used for solving clustering

problems. Let's explore each category and some representative algorithms from each:

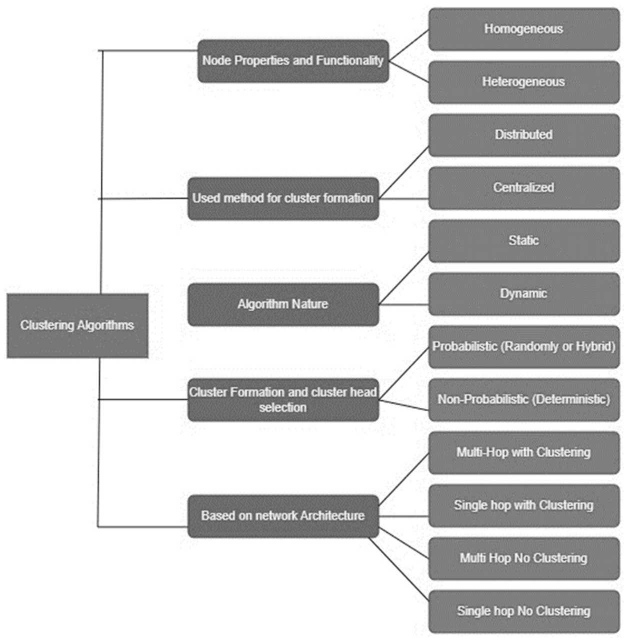

Fig.1. Types of Clustering in Wireless Sensor Networks

1.1 Non-Metaheuristic Algorithms for Clustering:

1.11 K-Means: K-Means is a popular and widely used clustering algorithm. It aims to partition data points into K clusters, where each data point belongs to the cluster with the nearest mean. It iteratively updates the cluster centroids until convergence.

1.12 Hierarchical Clustering: Hierarchical clustering builds a tree-like structure of clusters by merging or splitting them based on their similarity. The two main types of hierarchical clustering are agglomerative (bottom-up) and divisive (top-down).

1.13 DBSCAN (Density-Based Spatial Clustering of Applications with Noise): DBSCAN is a density-based clustering algorithm that groups data points based on their density. It identifies core points, which have a minimum number of points within a specified radius, and forms clusters by connecting core points that are within each other's radius.

1.14 OPTICS (Ordering Points to Identify the Clustering Structure): OPTICS is an extension of DBSCAN that overcomes some of its limitations. It produces an ordering of the data points

Catalyst Research Volume 23, Issue 2, October 2023 Pp. 2100-2117

that represents the density-based clustering structure, allowing for flexible clustering extraction.

1.15 Gaussian Mixture Model (GMM): GMM assumes that the data points are generated from a mixture of several Gaussian distributions. The algorithm estimates the parameters of these distributions to identify the underlying clusters.

1.2 Metaheuristic Algorithms for Clustering:[10]-[12]

1.21 Particle Swarm Optimization (PSO): PSO is inspired by the social behaviour of bird flocking or fish schooling. In the context of clustering, data points are treated as particles that move through the search space to find optimal cluster centres.

1.22 Genetic Algorithm (GA): GA is inspired by the process of natural selection. It uses operators such as selection, crossover, and mutation to evolve a population of potential solutions to the clustering problem.

1.23 Ant Colony Optimization (ACO): ACO mimics the foraging behaviour of ants. In clustering, it can be used to find centroids or to optimize clustering parameters.

1.24 Simulated Annealing (SA): Simulated annealing is a probabilistic metaheuristic algorithm that explores the search space by accepting worse solutions initially and gradually decreasing the acceptance probability.

1.25 Harmony Search (HS): HS is inspired by musicians' improvisation process. It explores the search space by creating new solutions based on the memory of previously found good solutions.

Both metaheuristic and non-metaheuristic algorithms have their strengths and weaknesses. Non-metaheuristic algorithms like K-Means are simple, efficient, and easy to implement but may get stuck in local optima. Metaheuristic algorithms, on the other hand, offer the potential to find better solutions and explore the search space more extensively, but they are often computationally more expensive. The choice of algorithm depends on the specific clustering problem, the size of the dataset, the desired quality of clustering, and available computational resources. Researchers and practitioners often experiment with different algorithms to find the best fit for their data and requirements. Clustering in sensor networks is a critical task as it helps in organizing and managing the data collected from numerous sensors efficiently.

Metaheuristic and non-metaheuristic algorithms can both be applied to clustering problems in sensor networks. Let's examine these algorithms in the context of sensor networks based on their method of implementation, clustering approach, handling size, advantages, disadvantages, applications, and efficiency.

2. Non-Metaheuristic Algorithms for Clustering in Sensor Networks: [10], [12], [16]

2.1 K-Means Clustering:

- Method of Implementation: Iterative optimization approach that assigns each sensor to the nearest centroid and recalculates centroids until convergence.
- Clustering Approach: Hard clustering, where each sensor belongs to exactly one cluster.
- Handling Size: Suitable for medium to large-sized sensor networks but sensitive to the initial choice of centroids.
- Advantages: Simple, easy to implement, computationally efficient, and widely used.
- Disadvantages: Prone to getting trapped in local optima, may not work well with irregularly shaped clusters or varying cluster sizes.
- Applications: Sensor data compression, data aggregation, and anomaly detection in sensor networks.
- Efficiency: Efficient for small to medium-sized datasets but may struggle with large and high-dimensional data.

2.2 Hierarchical Clustering:

- Method of Implementation: Builds a tree-like structure of clusters through agglomerative (bottom-up) or divisive (top-down) merging.
- Clustering Approach: Can be used for both hard and soft clustering, providing a hierarchy of clusters.
- Handling Size: Can handle large sensor networks but computational complexity increases with the number of sensors.
- Advantages: Produces a hierarchy of clusters, no need to specify the number of clusters beforehand, flexible.
- Disadvantages: Computationally expensive for large datasets, difficult to interpret the results.
- Applications: Identifying nested patterns in sensor data, visualizing cluster hierarchies.
- Efficiency: Can be time-consuming for large sensor networks due to its hierarchical nature.

2.3 DBSCAN (Density-Based Spatial Clustering of Applications with Noise):

- Method of Implementation: Density-based algorithm that groups sensors based on their density and connectivity.
- Clustering Approach: Hard clustering with core points, border points, and noise points.
- Handling Size: Efficient for large sensor networks and can handle irregularly shaped clusters.
- Advantages: Robust to noise and can identify clusters of varying shapes and sizes.
- Disadvantages: Sensitive to the choice of distance and density parameters, may struggle with clusters of different densities.
- Applications: Outlier detection, spatial pattern recognition in sensor networks.

- Efficiency: Efficient for large and dense sensor networks, but may suffer from higher-dimensional data.

2.4 Fuzzy C-Means (FCM) Clustering:

- Method of Implementation: FCM extends K-Means to allow soft clustering, where each sensor is assigned a membership degree to multiple clusters.
- Clustering Approach: Soft clustering, where each sensor belongs to multiple clusters with varying degrees of membership.
- Handling Size: Suitable for small to medium-sized sensor networks, but may face scalability issues with large networks.
- Advantages: Robust to noise, provides flexibility in cluster membership assignment, and can handle overlapping clusters.
- Disadvantages: Sensitive to the choice of the fuzziness parameter, may converge to local optima.
- Applications: Fuzzy sensor data clustering, pattern recognition in sensor networks.
- Efficiency: May have higher computational overhead compared to K-Means, especially for large networks.[17]

2.5 Expectation-Maximization (EM) Clustering:

- Method of Implementation: EM is a statistical algorithm used to estimate parameters of probabilistic models, such as Gaussian Mixture Models (GMMs).
- Clustering Approach: EM aims to find the maximum likelihood estimates of the parameters in the GMM to cluster data.
- Handling Size: Suitable for small to medium-sized sensor networks but may become computationally expensive for large networks.
- Advantages: Probabilistic clustering, can handle data points with uncertain memberships.
- Disadvantages: Sensitive to the choice of the initial model parameters, may converge to local optima.
- Applications: Clustering sensor data with underlying probabilistic distributions.
- Efficiency: EM can be computationally expensive, especially when the number of clusters or model complexity increases.[16-17]

2.6 Affinity Propagation (AP):

- Method of Implementation: AP is a message-passing algorithm that determines the number of clusters and cluster centroids simultaneously.
- Clustering Approach: Hard clustering with exemplars, where sensors serve as exemplars for their respective clusters.

- Handling Size: Can handle small to medium-sized sensor networks but may face scalability issues with large networks.
- Advantages: Does not require the number of clusters to be specified beforehand, can handle uneven cluster sizes.
- Disadvantages: Sensitive to the choice of similarity measures and damping factor, may converge to suboptimal solutions.
- Applications: Clustering sensors with diverse data characteristics and varying connectivity.
- Efficiency: AP may have higher computational complexity for large sensor networks, especially during the message-passing phase.

2.7 LEACH (Low-Energy Adaptive Clustering Hierarchy): LEACH is one of the pioneering protocols for energy-efficient clustering in wireless sensor networks. It aims to extend the network lifetime by rotating cluster heads among the sensors.

- Method of Implementation: LEACH uses a randomized approach to elect cluster heads for each round. The sensors with low-energy nodes have higher chances of being selected as cluster heads.
- Clustering Approach: LEACH employs a decentralized, self-organizing, and adaptive clustering approach where sensors form clusters with a rotating cluster head.
- Handling Size: LEACH is suitable for medium to large-sized sensor networks.
- Advantages: Energy-efficient, easy to implement, adaptive to changes in network conditions.
- Disadvantages: Cluster head selection randomness can lead to uneven energy consumption, may suffer from early node failures.
- Applications: LEACH is commonly used in environmental monitoring, agriculture, and surveillance applications.[1],[2],[9],[10],[16],[21]

2.8 SEP (Stable Election Protocol): SEP is an improved version of LEACH that addresses the issue of early node death in the original LEACH protocol. It aims to provide better stability and prolong the network lifetime.

- Method of Implementation: SEP uses a combination of deterministic and probabilistic approaches for cluster head selection. Nodes with higher energy levels and better connectivity have higher probabilities of becoming cluster heads.
- Clustering Approach: SEP follows a hierarchical and distributed clustering approach, allowing for better energy balance.
- Handling Size: SEP is suitable for medium to large-sized sensor networks.
- Advantages: Enhanced stability, more balanced energy consumption, increased network lifetime.
- Disadvantages: Increased complexity compared to LEACH, but still relatively easy to implement.

Catalyst Research Volume 23, Issue 2, October 2023 Pp. 2100-2117

- Applications: SEP is commonly used in applications that require better stability and resilience to node failures.[1-2]

2.9 DEEC (Distributed Energy-Efficient Clustering): DEEC is another energy-efficient clustering protocol that aims to prolong network lifetime by dynamically adapting the cluster head selection process based on remaining energy levels.

- Method of Implementation: DEEC employs a stochastic process to select cluster heads, considering the remaining energy of the nodes.
- Clustering Approach: DEEC uses a distributed approach for clustering, allowing sensors to self-organize into clusters.
- Handling Size: DEEC is suitable for medium to large-sized sensor networks.
- Advantages: Energy-efficient, adaptive to node energy levels, better energy balance.
- Disadvantages: Complexity increases with larger networks, may suffer from cluster head failures.
- Applications: DEEC is commonly used in applications that require balanced energy consumption and network longevity. [22]

2.10 TEEN (Threshold-sensitive Energy Efficient Sensor Network Protocol): TEEN is a data-centric clustering protocol that aims to reduce communication overhead by transmitting data only when certain thresholds are crossed.

- Method of Implementation: TEEN utilizes data-centric communication and event-driven data transmission.
- Clustering Approach: TEEN employs a hierarchical approach with data fusion capabilities.
- Handling Size: TEEN is suitable for medium to large-sized sensor networks.
- Advantages: Reduced communication overhead, energy-efficient data transmission.
- Disadvantages: May suffer from delays in transmitting critical data due to threshold-based approach.
- Applications: TEEN is commonly used in applications that require event-driven data reporting and efficient data fusion.

2.11 APTEEN (Adaptive Periodic Threshold-sensitive Energy Efficient Sensor Network Protocol): APTEEN is an extension of TEEN that addresses the limitation of fixed thresholds in TEEN by adapting them dynamically.

- Method of Implementation: APTEEN uses adaptive thresholds that change based on network conditions.
- Clustering Approach: Similar to TEEN, APTEEN follows a hierarchical clustering approach with data-centric communication.
- Handling Size: APTEEN is suitable for medium to large-sized sensor networks.
- Advantages: Improved adaptability to dynamic conditions, reduced data transmission delays.

- Disadvantages: Complexity increases compared to TEEN.
- Applications: APTEEN is commonly used in applications where adaptive data reporting is critical.

2.12 ZSEP (Zone-based Stable Election Protocol): ZSEP is a clustering protocol that divides the sensor network into zones and employs both energy level and distance-based metrics for cluster head selection.

- Method of Implementation: ZSEP uses a combination of energy and distance as selection metrics for cluster heads.
- Clustering Approach: ZSEP follows a zone-based clustering approach to organize the network.
- Handling Size: ZSEP is suitable for medium to large-sized sensor networks.
- Advantages: Improved energy balance, better network stability.
- Disadvantages: Requires additional localization information for distance-based calculations.
- Applications: ZSEP is commonly used in applications that require balanced energy consumption and network stability.

2.13 PEGASIS (Power-Efficient Gathering in Sensor Information Systems): PEGASIS is a chain-based clustering protocol that aims to reduce energy consumption during data aggregation.

- Method of Implementation: PEGASIS forms a chain of sensors to transmit data to a base station in a daisy-chained manner.
- Clustering Approach: PEGASIS follows a chain-based approach for data aggregation and transmission.
- Handling Size: PEGASIS is suitable for medium to large-sized sensor networks.
- Advantages: Improved energy efficiency during data aggregation, reduced communication overhead.
- Disadvantages: May suffer from single points of failure along the chain.
- Applications: PEGASIS is commonly used in applications that require efficient data aggregation and communication to a base station.

2.14 MODLEACH (Modified Low-Energy Adaptive Clustering Hierarchy): MODLEACH is an enhanced version of LEACH that addresses some of its limitations by introducing a more stable and energy-efficient clustering protocol.

- Method of Implementation: MODLEACH improves on LEACH by using fixed and dynamic probabilities for cluster head selection.
- Clustering Approach: MODLEACH follows a hierarchical and distributed clustering approach with improved energy balancing.
- Handling Size: MODLEACH is suitable for medium to large-sized sensor networks.

- Advantages: Better energy efficiency, improved network stability.
- Disadvantages: Increased complexity compared to the original LEACH.
- Applications: MODLEACH is commonly used in applications that require improved energy efficiency and network stability.

3. Metaheuristic Algorithms for Clustering in Sensor Networks:

3.1 Particle Swarm Optimization (PSO):

- Method of Implementation: Swarm-based optimization inspired by the social behavior of birds.
- Clustering Approach: PSO aims to find optimal cluster centroids by iteratively updating particle positions.
- Handling Size: Can handle large sensor networks, but the convergence speed depends on the swarm size.
- Advantages: Global optimization capability, suitable for large-scale problems, can handle irregularly shaped clusters.
- Disadvantages: Sensitivity to parameter settings, may get stuck in local optima.
- Applications: Sensor placement optimization, energy-efficient clustering in wireless sensor networks.
- Efficiency: Can be efficient for large sensor networks, especially when parallelized.

3.2 Genetic Algorithm (GA):

- Method of Implementation: Evolutionary approach based on natural selection and genetic operators.
- Clustering Approach: GA evolves a population of potential cluster configurations and selects the best ones.
- Handling Size: Suitable for medium to large-sized sensor networks but may face scalability issues.
- Advantages: Global optimization capability, flexible and adaptable to different clustering objectives.
- Disadvantages: Computationally expensive, requires tuning of genetic operators and parameters.
- Applications: Sensor network coverage optimization, fault-tolerant clustering.
- Efficiency: May suffer from efficiency issues in large sensor networks due to its population-based nature.

3.3 Ant Colony Optimization (ACO):

- Method of Implementation: Inspired by the foraging behavior of ants to find optimal paths.
- Clustering Approach: ACO can be used to optimize cluster centroids or clustering parameters.

Catalyst Research Volume 23, Issue 2, October 2023 Pp. 2100-2117

- Handling Size: Suitable for medium-sized sensor networks, but large-scale applications may face scalability challenges.
- Advantages: Robust to local optima, adaptive, and decentralized nature.
- Disadvantages: May require fine-tuning of parameters, efficiency decreases with larger networks.
- Applications: Sensor placement optimization, data routing in wireless sensor networks.
- Efficiency: Generally suitable for moderate-sized sensor networks but may be less efficient for large networks.

3.4 Simulated Annealing (SA):

- Method of Implementation: Probabilistic optimization inspired by the annealing process in metallurgy.
- Clustering Approach: SA explores the search space by accepting worse solutions initially and gradually reducing the acceptance probability.
- Handling Size: Can handle moderate-sized sensor networks but may struggle with very large networks.
- Advantages: Escape local optima, flexible in accepting worse solutions initially, can handle various clustering objectives.
- Disadvantages: Slower convergence rate, requires careful tuning of cooling schedule and parameters.
- Applications: Data clustering in dynamic sensor networks, multi-objective clustering.
- Efficiency: May be less efficient for large-scale sensor networks due to its iterative nature.

3.5 Harmony Search (HS):

- Method of Implementation: Music-inspired algorithm to find optimal solutions through improvisation.
- Clustering Approach: HS explores the search space to find optimal cluster configurations.
- Handling Size: Generally suitable for moderate-sized sensor networks.
- Advantages: Escape local optima, simplicity, and ease of implementation.
- Disadvantages: May converge slowly, sensitivity to parameter settings.
- Applications: Sensor placement optimization, load balancing in sensor networks.
- Efficiency: Can be efficient for moderate-sized sensor networks, but larger networks may pose challenges.

3.6 Firefly Algorithm (FA):

- Method of Implementation: FA is inspired by the flashing patterns of fireflies and uses their attraction to find optimal solutions.

- Clustering Approach: FA aims to optimize cluster centroids by simulating the flashing behavior of fireflies.
- Handling Size: Can handle small to medium-sized sensor networks but may face efficiency issues with large networks.
- Advantages: Global optimization capability, adaptive and robust, and can handle complex objective functions.
- Disadvantages: Convergence speed may vary with different problems, requires fine-tuning of parameters.
- Applications: Clustering sensors in dynamic and changing environments, coverage optimization.
- Efficiency: Efficiency depends on problem complexity and parameter settings but can be competitive for moderate-sized sensor networks.

3.7 Artificial Bee Colony (ABC) Algorithm:

- Method of Implementation: ABC is inspired by the foraging behaviour of honeybee colonies, where bees communicate to find food sources.
- Clustering Approach: ABC uses the search behaviour of bees to optimize cluster centroids.
- Handling Size: Suitable for small to medium-sized sensor networks, but may face scalability challenges for larger networks.
- Advantages: Robust and adaptive, can handle multi-modal and non-linear optimization problems.
- Disadvantages: Convergence speed may vary with different problems, parameter tuning required.
- Applications: Clustering sensors in dynamic environments, optimizing sensor coverage and connectivity.
- Efficiency: ABC can be computationally efficient for moderate-sized sensor networks but may become less efficient for larger networks.

3.8 Bat Algorithm (BA):

- Method of Implementation: BA is inspired by the echolocation behavior of bats, where they emit sounds and use feedback to find prey.
- Clustering Approach: BA uses bats' echolocation and movement strategies to optimize cluster centroids.
- Handling Size: Suitable for small to medium-sized sensor networks but may become computationally expensive for larger networks.
- Advantages: Adaptive, flexible, and can handle continuous and discrete optimization problems.
- Disadvantages: Convergence speed may vary with different problems, requires parameter tuning.

- Applications: Clustering sensors in dynamic environments, optimizing sensor placement for coverage.
- Efficiency: BA can be efficient for small to medium-sized sensor networks, but its efficiency may decrease for larger networks.

Table 1 Summary chart of the key points for the mentioned clustering methods in wireless sensor networks

Clustering Method	Implementation Method	Clustering Approach	Efficiency	Complexity
LEACH	Randomized cluster head selection	Decentralized and adaptive	Energy-efficient, easy to implement	Suitable for medium to large-sized networks
SEP	Combination of deterministic and probabilistic approaches	Hierarchical and distributed	Enhanced stability, increased network lifetime	Suitable for medium to large-sized networks
DEEC	Stochastic cluster head selection	Distributed	Energy-efficient, better energy balance	Suitable for medium to large-sized networks
TEEN	Data-centric communication, event-driven data transmission	Hierarchical and data fusion	Reduced communication overhead	Suitable for medium to large-sized networks
APTEEN	Adaptive thresholds	Hierarchical and data-centric	Improved adaptability, reduced data transmission delays	Suitable for medium to large-sized networks
ZSEP	Combination of energy and distance-based metrics	Zone-based clustering	Improved energy balance, better network stability	Suitable for medium to large-sized networks

PEGASIS	Chain-based data aggregation	Chain-based	Improved energy efficiency during data aggregation	Suitable for medium to large-sized networks

The table 1 presents a comprehensive overview of clustering methods tailored for wireless sensor networks, evaluating their implementation strategies, clustering paradigms, operational efficiency, and associated complexities. Each method demonstrates distinct characteristics that render them suitable for specific network configurations. LEACH employs randomized cluster head selection in a decentralized and adaptive framework, proving to be energy-efficient and easily implementable, making it apt for medium to large-sized networks. SEP, combining deterministic and probabilistic approaches within a hierarchical and distributed context, exhibits enhanced stability and prolonged network lifespan, particularly advantageous for medium to large-scale deployments. DEEC, employing stochastic cluster head selection in a distributed manner, showcases energy efficiency and superior energy balance, making it a fitting choice for medium to large-sized networks. TEEN adopts data-centric communication with event-driven data transmission in a hierarchical and data fusion approach, resulting in reduced communication overhead, well-suited for medium to large-sized networks. APTEEN, integrating adaptive thresholds in a hierarchical and data-centric mode, demonstrates heightened adaptability and decreased data transmission delays, making it a favorable option for similar network scales. ZSEP, through a combination of energy and distance-based metrics in zone-based clustering, achieves improved energy balance and network stability, catering to medium to large-sized networks. Finally, PEGASIS, leveraging chain-based data aggregation, enhances energy efficiency during data processing, making it a pragmatic choice for medium to large-sized networks.

The comparison table 2 provides valuable insights into the implementation methods and efficiency levels of various clustering approaches. It is evident that different methods exhibit distinct characteristics that can be crucial in selecting the appropriate approach for specific applications. Fuzzy C-Means (FCM) Clustering and Expectation-Maximization (EM) both utilize soft clustering techniques. FCM employs iterative optimization and demonstrates a moderate to high level of efficiency, potentially requiring fine-tuning. Its efficiency is derived from an objective function that considers distances between data points and centroids. On the other hand, EM relies on probabilistic estimation and exhibits moderate to high efficiency with iterative convergence. It maximizes log-likelihood or the expectation of data under the model. Affinity Propagation (AP) adopts a message-passing implementation method for hard clustering. It demonstrates moderate to high efficiency with iterative convergence, achieved through similarity propagation and responsibility-update message passing.

Swarm intelligence-based approaches, including Firefly Algorithm (FA), Artificial Bee Colony (ABC), and Bat Algorithm (BA), all implement soft clustering. They exhibit a low to moderate level of efficiency, providing global optimization. FA's efficiency is governed by the intensity of firefly flashing and attractiveness between fireflies. ABC's efficiency is determined by a fitness

Catalyst Research Volume 23, Issue 2, October 2023 Pp. 2100-2117

function based on the quality of food sources. BA achieves efficiency through bat position updates using echolocation and movement strategies. Simulated Annealing (SA) employs probabilistic optimization for both soft and hard clustering. It achieves a moderate level of efficiency through an energy function and acceptance probability during temperature reduction.

Table 2 Comparison stats of implementation method and efficiency of various clustering approach

Clustering Method	Implementation Method	Clustering Approach	Efficiency	Mathematical Formula
Fuzzy C-Means (FCM) Clustering	Iterative optimization	Soft clustering	Moderate to high, may require fine-tuning	Objective function involving distances between data points and centroids.
Expectation-Maximization (EM)	Probabilistic estimation	Soft clustering	Moderate to high, iterative convergence	Maximizing log-likelihood or expectation of data under the model
Affinity Propagation (AP)	Message-passing	Hard clustering	Moderate to high, iterative convergence	Similarity propagation and responsibility-update message passing
Firefly Algorithm (FA)	Swarm intelligence	Soft clustering	Low to moderate, global optimization	Intensity of firefly flashing and attractiveness between fireflies
Artificial Bee Colony (ABC)	Swarm intelligence	Hard clustering	Low to moderate, global optimization	Fitness function based on the quality of food sources
Bat Algorithm (BA)	Swarm intelligence	Soft clustering	Low to moderate, global optimization	Bat position update using echolocation and movement strategies
Simulated Annealing (SA)	Probabilistic optimization	Soft or hard clustering	Moderate, global optimization	Energy function and acceptance probability during temperature reduction

Catalyst Research Volume 23, Issue 2, October 2023 Pp. 2100-2117

In conclusion, this comparative analysis underscores the diversity of clustering methods, each with its unique implementation and efficiency profile. This information is invaluable for making informed decisions in selecting the most suitable clustering approach for specific applications and datasets.

Table 3 Complexity Chart of different clustering schemes

Fuzzy C-Means (FCM) Clustering	Medium to high, depending on data size
Expectation-Maximization (EM)	Medium to high, depending on data size
Affinity Propagation (AP)	Medium to high, depending on data size
Firefly Algorithm (FA)	Low to moderate, depending on swarm size
Artificial Bee Colony (ABC)	Low to moderate, depending on swarm size
Bat Algorithm (BA)	Low to moderate, depending on swarm size
Simulated Annealing (SA)	Moderate, depending on temperature schedule
Fuzzy C-Means (FCM) Clustering	Medium to high, depending on data size

4 CONCLUSION

In conclusion, the discussed clustering algorithms can be categorized into two main groups: non-metaheuristic and metaheuristic algorithms. Each group offers unique approaches to clustering in wireless sensor networks, and the choice of the algorithm depends on the specific requirements of the application and characteristics of the dataset. Non-metaheuristic algorithms like K-Means, Hierarchical Clustering, and DBSCAN are widely used and relatively easy to implement. They provide efficient and deterministic clustering solutions, making them suitable for scenarios where simplicity and speed are important. However, they may struggle with complex data structures, varying cluster sizes, and noisy data. Additionally, non-metaheuristic algorithms may be sensitive to the initial parameters, leading to suboptimal solutions or local optima. On the other hand, metaheuristic algorithms such as PSO, GA, ACO, SA, FA, ABC, and BA offer the advantage of global optimization capabilities. They can handle complex and non-linear objective functions, making them suitable for applications with irregularly shaped clusters and varying cluster sizes. Metaheuristic algorithms explore the search space more extensively, allowing them to potentially find better solutions than non-metaheuristic approaches. However, they generally come at a higher computational cost, especially for large-scale sensor networks. Fuzzy C-Means (FCM) and Expectation-Maximization (EM) are soft clustering algorithms that assign data points to multiple clusters with varying degrees of membership. They are suitable for situations where data points may belong to more than one cluster, providing more flexibility in modelling uncertainty in the data. However, they require careful parameter tuning and may suffer from higher computational complexity. Affinity Propagation (AP) is a message-passing algorithm that can automatically determine the number of clusters and cluster exemplars. It offers the advantage of adaptability and stability but may have higher computational demands.

In critical evaluation, it is essential to consider the trade-offs between efficiency, accuracy, complexity, and scalability when choosing a clustering algorithm for sensor networks. Non-metaheuristic algorithms like K-Means and DBSCAN are efficient and straightforward but may lack adaptability. Metaheuristic algorithms provide better optimization capabilities but may require more computational resources. Deciding on the most appropriate algorithm depends on factors like the network size, data characteristics, energy constraints, and the specific clustering objective. Furthermore, researchers and practitioners should explore hybrid approaches or algorithm modifications tailored to their specific needs. Additionally, the clustering algorithms discussed here represent only a fraction of the vast range of clustering techniques available, and ongoing research and advancements in the field may lead to new and improved algorithms in the future. As the field of clustering in sensor networks continues to evolve, it is essential to stay updated with the latest developments and adapt the chosen algorithm to the specific requirements of each application.

ACKNOWLEDGEMENT

In our research article, " Analytical Report on Metaheuristic and Non-Metaheuristic Algorithms for Clustering in Wireless Networks" we extend our appreciation to the key contributors and our research facility Centre at School of electronics and Electrical Engineering, Lovely Professional University, which facilitated its completion. There are no conflicts of interest among anyone. Our gratitude extends to the broader research community for their contributions, enriching our literature review specifically in next generation capability models for upcoming telecom sector. Colleagues and friends provided insightful discussions, enhancing the quality of this work. Lastly, our families and friends deserve recognition for their unwavering support and encouragement throughout this research journey.

REFERENCES

[1] Nishant Tripathi, Charanjeet Singh, Charanjeet Singh, “Hybrid Cloud Network Security Issues and Challenges for Sensor as a Service in Virtual Sensor Networks” in “European Chemical Bulletin” ,Eur. Chem. Bull. 2023,12 (Special Issue 6), 7576 – 7590.

[2] N.Tripathi, C.Singh, and K. K.Sharma, “Enhancement of Performance Metrics of Heterogeneous Wireless Camera Sensor Network with Functionalization of Extensive Zonal Stable Election Protocol using Threshold Amplification and Residual Energy”, IJRITCC, vol. 11, no. 8s, pp. 657–669, Aug. 2023.

[3] Sheng W., Hui T., Chen Q., Jiang M., Wang G. (2016) A Survey of Load Balancing in Wireless Sensor Networks: Taxonomy, Challenges, and Solutions. Sensors, 16(3), 430.

[4] Khalid R., Javaid N., Akhtar N., Niaz I.A., Khan Z.A., Guizani M. (2018) A Comprehensive Survey of Load Balancing Techniques in Wireless Sensor Networks. IEEE Communications Surveys & Tutorials, 20(1), 65-115.

[5] Chen J., Li Q., Zhao W., Wu D. (2017) Load Balancing Techniques in Wireless Sensor Networks: A Survey. Mobile Information Systems, 2017, 7940745.

[6] Xiong N., Liu X., Ma J., Cheng J., Chen S. (2015) A Survey of Load Balancing in Wireless Sensor Networks: Relevance, Methods and Challenges. International Journal of Distributed Sensor Networks, 11(5), 516269.

[7] Guo M., Ding Y., Shen L., Li Y., Wang W. (2019) A Survey on Load Balancing in Wireless Sensor Networks. Wireless Communications and Mobile Computing, 2019, 9310572.

[8] Ding Y., Zhang S., Tian S., Shen L., Wang W. (2017) Load Balancing in Wireless Sensor Networks: State-of-the-Art and Future Perspectives. Sensors, 17(12), 2773.

[9] Silvani, Xavier et al. "IEEE 802.11 Wireless sensor network for hazard monitoring and mitigation." Natural Hazards 114 (2022): 3545-3574.

[10] Chaudhari, S.S. (2021). A survey on multipath routing techniques in wireless sensor networks. Int. J. Netw. Virtual Organisations, 24, 267-328.

[11] Venkateswararao, M., & Malladi, S. (2021). Improving Packet Delivery Ratio in Wireless Sensor Network with Multi Factor Strategies. International Journal of Advanced Computer Science and Applications, 12.

[12] Adeel A, Gogate M, Farooq S, Ieracitano C, Dashtipour K, Larijani H, Hussain A (2019) A survey on the role of wireless sensor networks and IoT in disaster management. In: Durrani TS, Wang W, Forbes SM (eds) Geological disaster monitoring based on sensor networks. Springer, Singapore, pp 57–66

[13] Ahmadi A, Bigdeli A, Baktashmotlagh M, Lovell B (2012) A wireless mesh sensor network for hazard and safety monitoring at the port of Brisbane, https://doi.org/10.1109/LCN.2012.6423601

[14] Alsamhi SH, Ma O, Ansari MS, Gupta SK (2019) Collaboration of drone and internet of public safety things in smart cities: an overview of QoS and network performance optimization. Drones 3(1):13

[15] Carlotti T, Silvani X, Innocenti E, Morandini F, Bulté N, Dang T (2019) An OCARI-based wireless sensor network for heat measurements during outdoor fire experiments. Sensors 19(1):158

[16] Erdelj M, Król M, Natalizio E (2017) Wireless sensor networks and Multi-UAV systems for natural disaster management. Comp Netw. https://doi.org/10.1016/j.comnet.2017.05.021

[17] Kandris D, Nakas C, Vomvas D, Koulouras G (2020) Applications of wireless sensor networks: an up-to-date survey. Appl Sys Innovat 3(1):14

[18] Kirichek, R., Paramonov, A., Koucheryavy, A.: Flying ubiquitous sensor networks as a quening system. In: Proceedings, International Conference on Advanced Communication Technology, ICACT 2015, Phoenix Park, Korea, July 01–03, 2015

[19] Attarzadeh, N., Mehrani, M.: A New Three Dimensinal Clustering Method for Wireless Sensor Networks. Global Journal of Computer Science and Technology 11(6), April 2011. version 1.0

[20] Hooggar, M., Mehrani, M., Attarzadeh, N., Azimifar, M.: An Energy Efficient Three Dimensional Coverage Method for Wireless Sensor Networks. Journal of Academic and Applied Studies 3(3), March 2013.

[21] Abakumov, P., Koucheryavy, A.: The cluster head selection algorithm in the 3D USN. In: Proceedings, International Conference on Advanced Communication Technology, ICACT 2014, Phoenix Park, Korea.

[22] Dang Thanh Hai, Le Hoang Son, Vinh Trong Le , « Novel Fuzzy Clustering Scheme for 3D Wireless Sensor Networks" , "Applied Soft Computing Journal http://dx.doi.org/10.1016/j.asoc.2017.01.021 .

[23] I.F. Akyildiz, W. Su*, Y. Sankarasubramaniam, E. Cayirci , "Wireless sensor networks: a survey" , "Computer Networks 38 (2002) 393–422".

[24] Antonio-Jesus Yuste-Delgado et. al. "A Distributed Clustering Algorithm Guided by the Base Station to Extend the Lifetime of Wireless Sensor Networks", www.mdpi.com/journal/sensors.

[25] Amrita Ghosal et. al. "Distributed on-demand clustering algorithm for lifetime optimization in wireless sensor networks", "wireless sensor networks, Journal of Parallel and Distributed Computing (2020), doi: https://doi.org/10.1016/j.jpdc.2020.03.014."

[26] Lijun Xie and Xi Zhang "3D Clustering-Based Camera Wireless Sensor Networks for Maximizing Lifespan With Minimum Coverage Rate Constraint" , "Globecom 2013 - Ad Hoc and Sensor Networking Symposium".

[27] Noman Shabbir and Syed Rizwan Hassan , "Routing Protocols for Wireless Sensor Networks", http://dx.doi.org/10.5772/intechopen.70208 ".