Optimization
Iranian Journal of Optimization
Volume 16, Issue 4, 2024, 201-227
Research Paper

Online version is available on: www.ijo.rasht.iau.ir

Islamic Azad University
Rasht Branch
ISSN: 2588-5723
E-ISSN:2008-5427

A Review on the Combined Use of Partitional Clustering and Metaheuristic Algorithms

Saeb Morady Tootkaleh [1*] and Hossein Azgomi[1]

[1]Department of Computer Engineering, Ra.C., Islamic Azad University, Rasht, Iran

Revise Date: 04 February 2026
Accept Date: 11 February 2026

Abstract

Nowadays, a vast amount of data is being generated, and its volume continues to increase daily. Various methods exist for storing data, and alongside storage, it is essential to manage data efficiently. One of the common approaches for data management and control is clustering. Data clustering is performed in two main forms: hierarchical and partitional. Cluster analysis is an "unsupervised" learning approach used to identify data structures. The term "unsupervised" means that the data is not labeled beforehand. Clustering methods can be utilized for data aggregation in various fields.

In this paper, we explore the combined use of metaheuristic algorithms and partitional clustering. The most well-known and widely used partitional clustering algorithm is the k-means algorithm. In addition to this well-known algorithm, we also discuss other approaches such as the k-medoids and k-median algorithms. In many cases, clustering algorithms tend to get trapped in local search optima. To overcome this issue, many researchers have incorporated metaheuristic algorithms alongside clustering methods, which have proven to be effective in resolving this problem. One of the most well-known and oldest metaheuristic algorithms is the genetic algorithm, which has gained significant attention from researchers and technologists over the years.

Keywords:
Data Clustering
Partitional Clustering
Unsupervised Clustering
Hierarchical Clustering
Metaheuristic Algorithms

***Correspondence E-mail**: saeb.moradytootkaleh@iau.ir

INTRODUCTION

One of the major challenges in the world of technology is the vast volume of data, making data management a complex and demanding task. Clustering is one of the common techniques in data mining that plays a significant role in data management. Nowadays, data is collected more easily and is abundantly available everywhere. However, what is scarce is knowledge. Extracting valuable knowledge from massive amounts of data is crucial (Kenidra et al., 2016). In the field of software, data analysis is the most essential tool for processing large-scale data (Velmurugan & Santhanam, 2011). Therefore, to analyze such large datasets, various methods must be employed. Clustering is one such technique that helps us partition data efficiently. By clustering data, we can make informed decisions, which is highly beneficial for many large-scale businesses.

Clustering is a term used to describe numerical methods and multivariate data analysis techniques aimed at discovering data groups. This definition can include a collection of objects that are grouped based on similarity (De Castro & Daniel, 2016). Grouping refers to the partitioning of a dataset into subsets where the objects within each group share common characteristics, usually based on a measure of similarity or distance. A cluster can be defined based on internal cohesion or external separation of its objects (De Castro & Daniel, 2016). Clustering is a data analysis method that, when applied to a heterogeneous dataset, identifies homogeneous subgroups as defined by a given data model or a similarity measure (Downs et al., 2002). Data mining is the process of discovering knowledge from vast amounts of data stored in databases, data warehouses, or other information repositories (Agrawal et al., 1998). It involves using algorithms to identify predictive patterns within datasets. Automated data analysis is used for behavior prediction, risk assessment, relationship identification, and other analytical tasks, applying models to the data (DeRosa, 2004). In fact, when data mining techniques are used to solve real-world problems, they fall into two broad categories: predictive methods and descriptive methods. Clustering is a type of descriptive method (Abonyi & Feil, 2007). Metaheuristic algorithms are a subset of optimization methods that aim to find the best (near-optimal) solution. These algorithms are derivative-free techniques characterized by their simplicity, flexibility, and ability to avoid local optima (Mirjalili et al., 2014). Metaheuristic algorithms operate in a stochastic manner, initiating the optimization process by generating random solutions. Due to their intuitive concepts and ease of implementation, these algorithms can be deployed in the simplest way possible. Moreover, they can be easily modified to suit specific problems. A key feature of metaheuristic algorithms is their remarkable ability to prevent premature convergence in optimization processes (Agrawal et al., 2021). These algorithms are widely applied across various fields, including engineering and applied sciences. For example: Electrical engineering (finding optimal solutions for power generation), Industrial engineering (job scheduling, transportation, vehicle routing problems, location problems), Civil engineering (bridge and building design), Communications (radar design, networking), Data mining (classification, prediction, clustering, system modeling) and more (Agrawal et al., 2021).

In the following sections, we will discuss various aspects of our study:

- Section 2 covers partitional clustering and different types of center-based clustering methods.
- Section 3 explains metaheuristic algorithms that have been combined with different clustering techniques.
- Section 4, which is a key part of this paper, focuses on the combined use of partitional clustering and metaheuristic algorithms.

- Finally, Section 5 presents the conclusion and a comparison of the discussed approaches.

CLUSTRING

Over the past two decades, we have witnessed an exponential growth in data volume. This trend can be observed in various domains such as social media, online transactions, wireless sensor networks, satellite data, and astronomical information. The overwhelming increase in data volume poses significant challenges in decision-making processes.

Data analysis enables us to identify specific patterns within datasets and make more informed decisions. Therefore, developing new approaches and computational tools for data analysis is essential. Data mining is a subfield of computer science and a primarily computational process aimed at discovering and extracting knowledge from data.

Data representation involves determining how data should be modeled and structured. This includes aspects such as data size, the number of available features, the nature and scale of the data, and the number of clusters or classes. For example, in text data, each document can be represented as a " container of words" In this approach, words are considered as features, but their order within the document is disregarded.

Clustering is one of the most useful approaches for pattern recognition, image processing, decision-making, and data classification. This technique is widely utilized in various fields, including medical sciences, cybersecurity, signal analysis, life sciences classification, remote sensing, demography, social sciences, geology, anthropology, economics, finance, and planning (Nanda, S. J., & Panda, G., 2014).

Generally, clustering techniques can be categorized into two main types: partitional clustering and hierarchical clustering. In this paper, we focus exclusively on partitional clustering. If the opportunity arises in the future, we will also explore hierarchical clustering, which has numerous applications.

Partitional clustering

A partitional clustering algorithm generates a single partition of the data without any hierarchical structure. This means that the algorithm requires a predefined number of clusters as a fundamental rule. Typically, a partitional algorithm optimizes an objective function defined by the dataset. Most of these algorithms need to be executed multiple times with different initializations, and the best configuration obtained is selected as the (optimal) clustering result (Nanda & Panda, 2014).

In machine learning, achieving perfect optimization is not always possible, and we do not always have predefined labels for classifying input data points. Problems where no target labels are available for classification are referred to as unsupervised learning problems in the field of artificial intelligence. In an unsupervised learning problem, the goal is to model the hidden structured information within the data.

Partitional clustering is a type of unsupervised learning problem where we aim to group similar data points into clusters based on their underlying structure. One of the most well-known and widely used partitional clustering algorithms is the K-means algorithm. In K-means, K represents the number of clusters in which we aim to classify our data points.

K-Means Clustering

K-Means is one of the most popular clustering algorithms. In this section, we describe the K-Means algorithm, which is undoubtedly the most well-known and widely used partitional clustering algorithm. Specifically, we focus on the global K-Means algorithm, which computes clusters incrementally. Incremental-based algorithms start with a single cluster center and gradually add a new cluster center in each iteration. This approach helps to minimize the global clustering problem. The K-

Means algorithm follows an iterative procedure, beginning with a randomly selected initial cluster configuration and then iteratively updating cluster memberships to achieve a better clustering outcome. Initially, K-Means randomly selects K cluster centers based on a predefined K value. Then, the algorithm alternates between two main steps until a stopping criterion is met. These steps are:

1. Assigning data points to clusters based on the minimum squared Euclidean distance.
2. Recomputing the cluster centers by averaging the assigned data points.

In other words, K-Means iteratively updates data point assignments based on their similarity to cluster centers until there is minimal or no significant change in the clustering function. The popularity of the K-Means algorithm comes from its simplicity and ease of implementation.

```
## K-Means Clustering

1. Choose the number of clusters(K) and
obtain the data points
2. Place the centroids c_1, c_2, .....
c_k randomly
3. Repeat steps 4 and 5 until convergence
or until the end of a fixed number of
iterations
4. for each data point x_i:
       - find the nearest centroid(c_1,
c_2 .. c_k)
       - assign the point to that cluster
5. for each cluster j = 1..k
       - new centroid = mean of all
points assigned to that cluster
6. End
```

K-Medoids clustering

The K-Medoids algorithm was introduced by (Kaufman and Rousseeuw, 1987). A medoid can be defined as a point within a cluster that has the smallest dissimilarity with other points in the same cluster. K-Medoids is similar to K-Means, as both are partitional clustering algorithms that divide a dataset into separate groups. Their objective is to minimize the distance between labeled points within a cluster and a designated central point (the cluster's medoid).

Key Differences Between K-Means and K-Medoids

1. Choice of Cluster Centers:
   - Unlike K-Means, where the cluster center is the mean of the points, K-Medoids selects actual data points as cluster centers (medoids).
   - This makes K-Medoids more interpretable since the cluster centers are real points from the dataset, rather than artificial centroids.
2. Distance Measurement:
   - K-Means typically relies on Euclidean distance for efficient solutions.
   - K-Medoids can use arbitrary dissimilarity measures, making it more robust to non-Euclidean data distributions.
3. Robustness to Outliers:
   - Since K-Medoids minimizes pairwise dissimilarity instead of squared Euclidean distances (as in K-Means), it is more resistant to noise and outliers compared to K-Means.

Working of K-Medoids

K-Medoids is a classical partitioning method that divides n objects into k clusters, where k is predefined before running the algorithm. The selection of k can be refined using techniques like Silhouette analysis to determine the optimal number of clusters.

K-Median clustering

The K-Median algorithm is similar to K-Means and is used for data analysis and clustering. However, instead of computing the mean of each cluster to determine the center, it calculates the median, which helps in reducing errors (Nesamalar et al., 2012).

Key Features of K-Median Clustering

1. Robustness to Outliers:
   - Unlike K-Means, which uses the mean, K-Median selects the median, making it less sensitive to outliers and noisy data (Whelan et al., 2015).
2. Optimization Objective:
   - The goal of K-Median clustering is to minimize the distance between each point and its closest cluster center.
   - The cluster center is set to the median of all points in that cluster, rather than the mean.
3. Handling Large Datasets:
   - The mean is highly susceptible to extreme values, which can skew the cluster centers when working with large datasets.
   - The median is a more robust statistic that provides better clustering stability.

Use Cases

- K-Median is often used in robust clustering scenarios, such as finance, supply chain logistics, and anomaly detection, where the presence of outliers is common.

K-Modes clustering

Clustering algorithms employ various strategies to determine cluster centers based on the type of data. For numerical data, Euclidean distance is the most widely used metric. In contrast, for categorical data, Hamming distance is considered the simplest and most effective method for clustering. The k-modes clustering algorithm leverages this distance metric to improve clustering performance by selecting an appropriate similarity measure. Hamming distance is defined as the number of positions at which the corresponding symbols in two binary strings of equal length differ. The k-modes algorithm, a commonly used method for clustering categorical data, relies on Hamming distance to compute the dissimilarity between objects for clustering purposes (Kuo et al., 2021).

Fuzzy c-means clustering

Given a set of objects, the primary objective of clustering is to partition the dataset into distinct groups based on object similarity, while minimizing intra-cluster dissimilarity. The Fuzzy C-Means (FCM) algorithm is a clustering method that allows a single data point to belong to multiple clusters simultaneously. This approach is particularly useful in pattern recognition tasks, where the boundaries between clusters may not be clearly defined (Aydilek & Arslan, 2013).

METAHEURISTIC ALGORITHMS

In recent years, there has been a significant surge of interest in metaheuristic algorithms, with researchers extensively leveraging them across a wide range of applications. Metaheuristics are generally categorized into two major classes: single-solution-based (single-objective) and population-based (multi-objective) algorithms. This distinction is fundamental in the literature. Broadly speaking, single-solution-based metaheuristics tend to be more exploitative, whereas population-based metaheuristics emphasize exploration (De León-Aldaco et al., 2015).

Metaheuristic algorithms are primarily used for optimization, and numerous approaches have been proposed over the past years—many of which are inspired by natural phenomena. The fundamental concepts of metaheuristics can be described at an abstract level, without being tied to any specific problem. These algorithms range from simple local search techniques to complex learning-based procedures (De León-Aldaco et al., 2015).

The general mechanism of a metaheuristic algorithm involves initiating a guided search process that efficiently explores the entire search space. These algorithms typically involve multiple parameters that must be tuned according to the specific problem and may incorporate strategies to avoid premature convergence to local optima.

Over the past few decades, many metaheuristic algorithms have been developed, among which Genetic Algorithms (GA) and Particle Swarm Optimization (PSO) are among the most well-known. These two approaches offer significant advantages, including robustness in handling complex problem scenarios and inherent support for parallelism (De León-Aldaco et al., 2015). Population-based methods, in particular, are more effective in global optimization tasks and are capable of dealing with objective functions that are static or dynamic, linear or nonlinear, and continuous or discontinuous.

Genetic algorithm (ga)

The Genetic Algorithm (GA), inspired by the principles of natural selection and genetic evolution, has been successfully applied to the optimization of complex processes. Starting from an initial population, three primary genetic operators—selection, crossover, and mutation—are iteratively applied to evolve solutions across successive generations (Sastry et al., 2005). Through the selection operator, individuals with higher fitness (better objective function values) are probabilistically chosen for reproduction. The crossover operator combines segments of two parent solutions to generate offspring, while the mutation operator randomly alters the offspring to introduce diversity. The newly generated offspring population, created through these genetic operations, replaces the parent population in the next generation. Various replacement strategies exist, such as elitism, which ensures that the best individuals are carried forward (Sastry et al., 2005).

One of the most effective ways to enhance the performance of genetic algorithms is hybridization. The most common form of hybridization involves integrating local search techniques within the genetic algorithm as a decisive component of the evolutionary process, often incorporating domain-specific knowledge into the search procedure (Assunção et al., 2011). These hybrid approaches are commonly referred to as Memetic Algorithms (MA) (Sastry et al., 2005).

The GA optimization process typically involves the following key parameters:

1. Initialization: The GA encodes the problem parameters as genes, assembling them into chromosomes to form the initial population. Each chromosome represents a potential solution to the optimization problem. Larger populations tend to promote greater genetic diversity, while smaller populations execute faster, particularly when dealing with computationally expensive fitness functions (De León-Aldaco et al., 2015).

2. Fitness Evaluation: Fitness values are assigned to individuals in the population based on their quality with respect to the objective function.

3. Selection: Pairs of fit individuals are selected from the current population based on a specific selection strategy for reproduction.

4. Crossover (Recombination): Selected parent pairs are recombined to generate offspring. The crossover operator ensures that each child inherits specific genes from both parents. This is the main mechanism

through which GA explores new and potentially better solutions.

5. Mutation: Mutation randomly alters a gene in each child, enabling the discovery of solutions not present in the current population. Poorly performing offspring may be discarded during selection in future generations.
6. Replacement and Iteration: A portion or the entirety of the current generation is replaced with the new offspring. This process is repeated starting from the fitness evaluation step until a predefined stopping criterion is met (Chong & Zak, 2001).

Particle swarm optimization (pso)

The Particle Swarm Optimization (PSO) algorithm is a metaheuristic optimization method introduced by Dr. Eberhart and Dr. Kennedy (Kennedy & Eberhart, 2001). PSO has been widely applied to various problems, including classification, where it has demonstrated impressive performance. For instance, PSO-based classification has been evaluated on cancer datasets, showing high accuracy and being recognized as one of the most efficient and robust classification methods (Mahapatra et al., 2011).

PSO is a population-based algorithm that utilizes a swarm of individuals to explore the optimal positions in the search space. In this approach, each individual, referred to as a particle, navigates the search space with an adaptable velocity. The movement of each particle is influenced by two key factors: its own historical best position and the best-known position found by the entire swarm (Linyi & Deren, 2010). The PSO methodology has also been employed to enhance classification performance through fusion strategies, yielding superior results compared to standard population averaging methods (Veeramachaneni et al., 2007).

PSO is grounded in the concept of Swarm Intelligence (SI), a subfield of artificial intelligence that models the collective behavior of decentralized, self-organizing systems. SI consists of a population of agents that interact locally and globally within a randomized environment. These agents follow simple behavioral rules, and their interactions lead to the emergence of complex global behavior without centralized control.

Firefly algorithm (FA)

The Firefly Algorithm (FA) is a swarm-based metaheuristic optimization technique proposed by (Xin-She Yang, 2008). This algorithm is inspired by the natural flashing behavior of fireflies, where the brightness of a firefly represents the quality of a solution in the search space. According to this metaphor, brighter fireflies attract others, enabling an efficient exploration of the solution space. FA has proven effective in solving a wide range of optimization problems (Yang, 2009; Jaradat & Banihamad, 2018).

Yang introduced the following assumptions as the foundational rules of FA (Yang, 2008; Yang, 2009):

- Attractiveness: All fireflies are unisex, meaning that one firefly is attracted to another regardless of its sex.
- Brightness and Distance: The attractiveness of a firefly is proportional to its brightness, and the perceived brightness decreases as the distance between fireflies increases.
- Random Movement: If a firefly does not find any brighter companion, it moves randomly in the search space.

Yang applied the FA to nonlinear design problems (Yang, 2010, p. 133-144) and multimodal optimization tasks (Yang, 2009, p. 169-178), demonstrating its capability in locating global optima in two-dimensional environments.

The Firefly Algorithm operates based on three idealized rules:

1. Unisex attraction: All fireflies can attract each other, independent of sex.

2. Relative brightness: The movement of a less bright firefly is directed toward a brighter one. If there is no brighter firefly in the neighborhood, the firefly moves randomly.

Light intensity: The brightness of a firefly is determined by the objective function value in the given optimization problem (Hassanzadeh & Meybodi, 2012).

Cuckoo search algorithm (CS)

The Cuckoo Search (CS) algorithm was developed based on the brood parasitism behavior of certain cuckoo species and their unique reproduction strategies. It has gained significant popularity and has been successfully applied to various real-world optimization problems. Given its effectiveness, several binary versions of the CS algorithm have also been developed to tackle binary optimization tasks.

In 2012, Tiwari (Tiwari, V., 2012) applied the Cuckoo Search algorithm to the problem of face recognition. In this approach, Discrete Cosine Transform (DCT) was used to extract features, which served as the host nests in the CS framework. The algorithm demonstrated its effectiveness in finding the most matching face image. To obtain an optimal feature subset, Rodriguez et al. (Rodrigues et al., 2013) proposed a Binary Cuckoo Search (BCS) algorithm by introducing a transformation function that maps continuous variables to binary representations.

As a nature-inspired swarm intelligence algorithm, Cuckoo Search has been applied to a broad range of optimization problems. Initially, it was tested on several benchmark test functions, and later, many researchers extended its application to complex tasks such as the Traveling Salesman Problem (TSP), clustering problems, multi-objective optimization, and structural optimization. Modified versions of the CS algorithm have also been proposed by various researchers to enhance performance in specific tasks, including data clustering (Girsang et al., 2017).

Bat algorithm (BA)

The Bat Algorithm (BA) is a nature-inspired metaheuristic optimization algorithm, which belongs to the broader family of Swarm Intelligence (SI) techniques. It draws inspiration from the biological echolocation behavior exhibited by bats (Yang & Gandomi, 2012). Specifically, the algorithm mimics the behavior of microbats, which use echolocation to navigate and detect prey by adjusting both the pulse emission rate and the loudness of sound vibrations (Kaufman & Rousseeuw, 1990).

Although most bat species employ echolocation at a basic level, microbats are particularly notable due to their highly sophisticated echolocation capabilities (Richardson & Elphick, 2011). The Bat Algorithm simulates this adaptive behavior through frequency-tuning and dynamic adjustment of pulse emission rates, enabling effective exploration and exploitation of the search space.

BA has gained popularity for its simplicity, ease of implementation, and its superior convergence performance compared to other metaheuristic algorithms (Gan & Lai, 2019). It has been successfully applied to a wide range of real-world optimization problems.

Shuffled frog leaping algorithm (SFLA)

The Shuffled Frog Leaping Algorithm (SFLA) is a population-based metaheuristic approach for solving complex optimization problems. SFLA is inspired by the natural behavior of frogs searching for food and is considered a hybrid metaheuristic algorithm. The algorithm simulates a population of virtual frogs that are partitioned into several groups known as "memeplexes." Each memeplex represents a subpopulation with distinct search cultures that independently explore the solution space. Within each memeplex, frogs influence one another and undergo a series of local searches for a predefined number of iterations.

After local exploration within memeplexes, the algorithm performs a global mixing (shuffling) process, where ideas and solutions are shared among different memeplexes. This process combines local exploration, similar to the Particle Swarm Optimization (PSO) approach, with global information exchange, enhancing the overall search capability towards the global optimum.

The SFLA is designed to perform guided exploratory searches using heuristic functions (mathematical functions) to address combinatorial optimization problems. The algorithm can be visualized as a group of frogs moving within a swamp that contains several stones at various locations. Frogs leap from one stone to another, attempting to land on the stone with the highest amount of available food, which represents the optimal solution in the search space (Eusuff et al., 2006).

Artificial bee colony algorithm (ABC)

The Artificial Bee Colony (ABC) algorithm is inspired by the foraging behavior of honey bees. In this algorithm, each food source represents a potential solution to an optimization problem, and the nectar amount of a food source corresponds to its fitness or quality with respect to the given problem. Similar to most swarm intelligence-based metaheuristic algorithms, ABC is population-based and imitates the collective behavior of decentralized systems.

The ABC algorithm consists of three main components:

1. **Food Sources:** A bee evaluates various aspects of a food source, such as its distance from the hive, ease of extraction, and amount of nectar. These factors collectively influence the selection and exploitation of the food source.
2. **Employed Bees:** Each employed bee is associated with a specific food source currently being exploited. It carries information about this source—such as its profitability, direction, and distance—and shares it with other bees waiting in the hive.
3. **Unemployed Bees:** These bees are searching for new food sources. They can either be *scouts*, exploring the search space randomly, or onlookers, selecting a food source based on the shared information provided by employed bees. On average, the exploitation rate of unemployed bees is relatively low (Karaboga & Akay, 2009).

The ABC algorithm has demonstrated notable performance in solving a variety of optimization problems, particularly due to its simple implementation, flexibility, and balance between exploration and exploitation.

Ant colony optimization algorithm (ACO)

The Ant Colony Optimization (ACO) algorithm is a metaheuristic inspired by the collective behavior of social insects (Dorigo, 1992; Dorigo, 1994). In ACO, an artificial ant mimics the exact behavior of a real ant to find the shortest path between food sources and the nest. Each ant collects necessary information about the problem, makes decisions stochastically, and generates solutions in a step-by-step manner. The emergent behavior results in a group of relatively "unintelligent" ants interacting through simple rules, dynamically adjusting themselves, and maintaining their positions along the shortest paths.

Ants leave their nest and initially move randomly without any knowledge about the location of food sources. While doing so, they deposit a substance called pheromone on the ground. The pheromone trail represents a potential solution to the problem and is positively reinforced in subsequent iterations to become more attractive. Therefore, the pheromone concentration indicates how effective a solution has been based on the previous movement history of the ants. In addition to pheromone concentration, ants may also use heuristic performance values, which generally represent the influence of more informative local knowledge (Assunção et al., 2011).

APPLICATION OF METAHEURISTIC ALGORITHMS IN PARTITIONAL CLUSTERING

Metaheuristic algorithms have gained significant attention for solving optimization problems by finding optimal methods for a specific issue. These methods can be identified through multiple factors, essentially forming a procedure of evolved solutions based on a set of rules or mathematical equations over numerous iterations.

Heuristic algorithms provide a reasoning approach for problem-solving that allows tackling a problem through trial and error. However, metaheuristic algorithms represent a higher-level innovation capable of offering more powerful capabilities to address complex problems. Computations in metaheuristic algorithms are of a consensual nature, incorporating general heuristic rules to solve a class of computational problems.

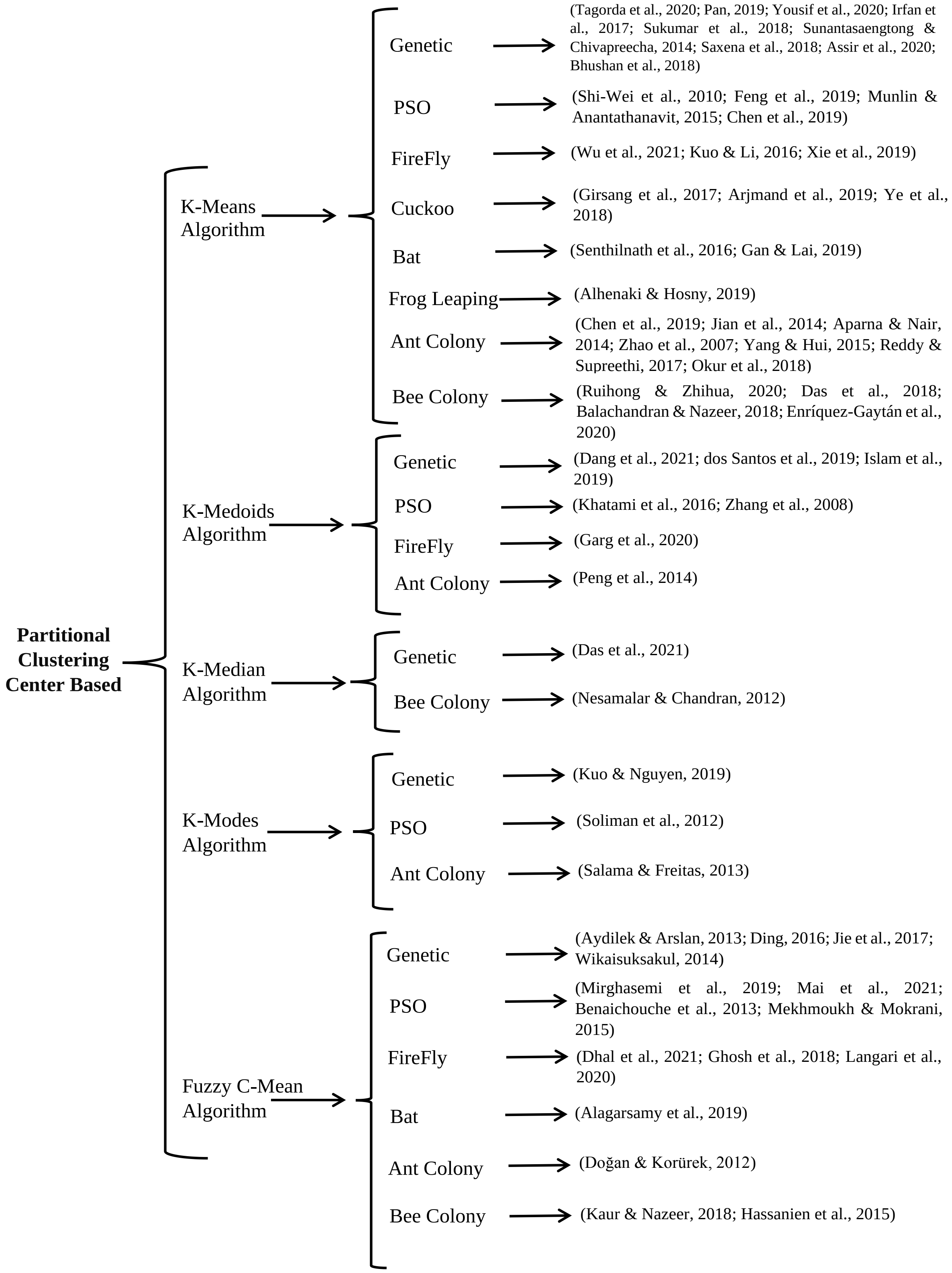
In partitional clustering, due to the nature of this method, an initial partitioning is first created. In the next stage, an iterative relocation approach is used, aiming to improve the clustering by moving objects (data) from one group to another. In other words, the algorithm keeps reassigning data to various clusters until a desirable (optimal) grouping is achieved.

Metaheuristic algorithms have been increasingly employed to enhance the performance of partitional clustering methods by improving initial cluster centers and avoiding local optima. Algorithms such as Genetic Algorithm, Particle Swarm Optimization, and Ant Colony Optimization can refine the clustering structure by iteratively updating cluster assignments based on global search strategies. These approaches have proven effective in high-dimensional data and complex clustering landscapes where traditional methods like k-means may fail to find optimal partitions (Maulik & Bandyopadhyay, 2000).

In partitional clustering, the effectiveness of the final solution highly depends on the initial choice of cluster centers. Metaheuristic algorithms help overcome this limitation by performing a global search to identify near-optimal initial points, thereby reducing sensitivity to random initialization. Moreover, these algorithms are capable of escaping local optima by balancing exploration and exploitation, which is particularly beneficial when dealing with noisy or non-linearly separable data. As a result, integrating metaheuristics with partitional clustering can significantly improve clustering accuracy and robustness in real-world applications such as bioinformatics, image segmentation, and customer segmentation (Xiao et al., 2008; Omran et al., 2007).

In this section, we focus on the metaheuristic algorithms that are applicable in solving partitional clustering problems and contribute to obtaining optimal (or near-optimal) solutions. Accordingly, a categorization is illustrated in the figure below, representing the overall process. This categorization consists of three levels, each of which will be examined individually.

In the first level, which comprises five branches, we include partitional clustering algorithms, such as k-means, k-medoids, k-median, k-modes, and fuzzy c-means. The second level consists of metaheuristic algorithms. This group includes eight metaheuristic methods: genetic algorithm, particle swarm optimization, firefly algorithm, cuckoo search algorithm, bat algorithm, shuffled frog leaping algorithm, artificial bee colony, and ant colony optimization. The third category encompasses previous studies that have been conducted in this area, which will be described in the following sections.



In this section, we will describe the selected studies that were numbered in the previous figure. We begin with the k-means algorithm, which has been utilized in conjunction with metaheuristic algorithms to solve various problems.

In (Tagorda et al., 2020), the authors proposed an optimized method for faster and more efficient delivery of goods from one location to another, designing an automatic vehicle routing system. According to the results obtained, the proposed method demonstrates superior performance in providing customer service within their respective regions. In this study, the k-means algorithm was used to assign couriers to a specific region or area, while the genetic algorithm was employed to select the optimal route.

In (Pan, 2019), the author Sinuo Pan investigates the optimal scheduling of an unmanned drone fleet in a delivery scenario. The objective is to construct a scheme that includes mission assignment for each drone, routing plans, and determining the number of drones required. Ultimately, an efficient method is proposed to solve the mission allocation and routing problem for multi-drone and multi-task delivery cases. In this study, k-means++ clustering is employed to determine the mission area for each drone, while the genetic algorithm is used to generate the flight paths.

In (Yousif et al., 2020), the authors proposed a novel method to improve the accuracy of license plate recognition for Arabic-Egyptian and English plates. The paper applies several image processing techniques such as edge detection and morphological operations for localization. By implementing a new method based on the genetic algorithm for optimizing these operations, the most prominent features were extracted. Additionally, the k-means clustering algorithm was used to segment the characters on the vehicle license plates.

In (Irfan et al., 2017), Shadab Irfan and colleagues employed various k-means clustering methods to categorize similar objects. To reduce the number of iterations required in the k-means clustering algorithm, a solution based on the genetic algorithm was proposed. By integrating k-means with evolutionary computation, the overall workflow and execution time are optimized. This approach contributes significantly to improving computational efficiency and reducing complexity.

In (Sukumar et al., 2018), the authors addressed the topic of network intrusion. Consequently, they introduced an intrusion detection system that utilizes an improved k-means algorithm for detection. In this system, the number of clusters is not predefined; instead, the optimal number is determined using a fitness function. This optimized value enhances the ability to identify different types of attacks.

In (Sunantasaengtong & Chivapreecha, 2014), the authors proposed the use of Wireless Sensor Networks (WSN) for indoor localization using the IEEE 802.15.4 standard. The proposed algorithm employs k-means clustering and the Genetic Algorithm (GA) as the engine for generating offline data, which leads to improved accuracy and reduced computational cost in the fingerprinting-based localization approach. The k-means clustering algorithm is used to estimate the Received Signal Strength Indicator (RSSI) across multiple classes for anomaly position estimation. Consequently, the Genetic Algorithm is applied to search for the optimal weights for each reference sensor to achieve higher accuracy in position estimation.

In (Saxena et al., 2018), an improved multi-objective Genetic Algorithm was proposed for the optimal separation (grouping) of heterogeneous numerical data. This method is designed for clearly defined clusters. The proposed approach integrates the Genetic Algorithm into the k-means algorithm with the aim of improving the cost function to better handle numerical data.

In (Assir et al., 2020), the authors addressed the issue of breast cancer in women. The article presents a fully automated Computer-Aided Diagnosis (CAD) system designed to assist radiologists in making accurate decisions by highlighting potentially suspicious regions. The use of CAD systems can enhance diagnostic accuracy by reducing the number of false positives and false negatives. CAD, as a computer-aided detection system, is an effective tool that can help radiologists identify tumor cells at early stages. In this study, the k-means algorithm was utilized to detect sensitive regions, while the Genetic Algorithm (GA) was employed to determine the appropriate center for each mammographic image.

In (Bhushan et al., 2018), Shashi Bhushan and colleagues proposed a hybrid method combining the Genetic Algorithm (GA) and the k-means algorithm to design an energy-efficient clustering protocol for heterogeneous wireless sensor networks. One of the main challenges in clustering lies in finding the optimal number of clusters within a large search space. In the proposed method, inter-cluster distance, intra-cluster distance, and the number of cluster heads are used to determine the optimal number of clusters. Energy conservation and network lifetime are considered as the two primary performance criteria.

In (Shi-Wei et al., 2010), the authors proposed a method based on k-means clustering and the Particle Swarm Optimization (PSO) algorithm. This paper presents a model that adopts Principal Component Analysis (PCA) to reduce the dataset to lower dimensions in order to avoid the effects of multicollinearity and the curse of dimensionality on clustering the same dataset. The article assumes that a higher amount of information is equivalent to greater variance.

In (Feng et al., 2019), the authors introduced a novel Class-based Evolutionary Extreme Learning Machine (CEELM) for extracting monthly operational rules of water reservoirs. In this method, the k-means clustering algorithm is used to partition the large-scale input space into several distinct subspaces, followed by the application of the Particle Swarm Optimization algorithm to identify the complex input-output relationships of samples within each cluster.

In (Munlin & Anantathanavit, 2015), the authors proposed a method for the Traveling Salesman Problem (TSP) using a divide-and-conquer strategy. In this method, the k-means algorithm is used to cluster the cities, and then the sequence of other cities is solved based on the order determined by the Particle Swarm Optimization (PSO) metaheuristic algorithm.

In (Chen et al., 2019), the authors presented a method to achieve accurate and efficient image segmentation. This article employs the k-means clustering algorithm alongside the PSO algorithm to enhance image segmentation. The goal of this method is to resolve the issue of initial cluster center selection in the k-means algorithm and to prevent getting trapped in local optima.

In (Wu et al., 2021), the authors addressed the issue of water resource management and proposed a novel learning machine model that incorporates the k-means clustering algorithm and the Firefly Algorithm. The study used subsets of data (5, 10, 15, 20, 25, 30, and 40) to estimate monthly averages via parallel computation for Poyang Lake in southern China using temperature data collected from 26 meteorological stations. Two inputs, namely average temperature and the maximum and minimum temperatures, were considered.

In (Kuo & Li, 2016), the authors proposed a three-stage predictive model to develop a forecasting system for export trade value. The model employs the k-means algorithm and the Firefly Algorithm in the first two stages. In the third stage, a Support Vector Machine (SVM) is used for final prediction.

In (Xie et al., 2019), the authors proposed two variants of the Firefly Algorithm, namely IIEFA and CIEFA, to address the issues of initial sensitivity and local optima traps commonly encountered in the standard k-means clustering algorithm. Two novel strategies were introduced to enhance diversity and improve the search efficiency.

In (Girsang et al., 2017), the issue of optimization in clustering is addressed. The authors integrated the k-means clustering algorithm with the Cuckoo Search (CS) metaheuristic algorithm. The CS algorithm was employed to overcome the local optimum problem—one of the known limitations of k-means. Cuckoo Search was used to generate robust initial cluster centers, while k-means was applied to accelerate convergence towards a solution.

In (Arjmand et al., 2019), the authors focused on breast tumor detection. They proposed a clustering-based algorithm for the automatic segmentation of tumors in MRI samples. The method utilized k-means clustering for segmentation and Cuckoo Search Optimization (CSO) for initializing the cluster centers in the k-means algorithm.

In (Ye et al., 2018), the authors proposed a k-means clustering algorithm based on Cuckoo Search. Experimental results demonstrated that the algorithm achieved better clustering performance, faster convergence rates, and higher accuracy compared to traditional methods.

In (Senthilnath et al., 2016), a new clustering method based on the Bat Algorithm (BA) was proposed to address product type classification using multispectral satellite imagery. The divisive clustering algorithm presented in this study extracts information by determining optimal cluster centers from training samples.

In (Gan & Lai, 2019), a method was introduced for the use of machine vision in the automated grading of Edible Bird's Nests (EBN), employing a novel set of features derived through image processing techniques. The classification of EBN was performed using a Bat Algorithm-based clustering method built upon the k-means algorithm.

In (Alhenaki & Hosny, 2019), the authors proposed an optimized model for text document clustering using a hybrid of the Genetic Algorithm (GA) and Shuffled Frog Leaping Algorithm (SFLA), referred to as GA-SFLA. This approach effectively clusters text documents based on selected features, where GA is responsible for feature selection and SFLA performs the clustering task.

In (Chen et al., 2019), a novel clustering algorithm was introduced based on a horizontal layout and a constrained version of k-means to enhance clustering accuracy. The Ant Colony Optimization (ACO) algorithm was employed to address the problem of reconstructing shredded textual documents.

In (Jian et al., 2014), the original k-means algorithm was analyzed, and due to its known limitation of getting trapped in local optima, an improved clustering algorithm based on Ant Colony Optimization was proposed. This enhanced algorithm was applied in the performance management system of student sports.

In (Aparna & Nair, 2014), a new method was introduced to improve cluster quality in high-dimensional data using Ant Colony Optimization

(ACO) alongside k-means. ACO, as one of the most widely used probabilistic metaheuristics, was applied to discover optimal clustering paths and enhance clustering efficiency.

In (Zhao et al., 2007), an image segmentation method was presented based on the integration of Ant Colony Optimization and k-means clustering. This approach accurately segmented objects, reduced segmentation time, and improved overall segmentation effectiveness.

In (Yang & Hui, 2015), the authors addressed information security in networks. A data mining approach was proposed to reduce the rate of false alarms generated by intrusion detection systems and improve detection accuracy. The k-means algorithm was used to accelerate the convergence rate within the Ant Colony Optimization process.

In (Reddy & Supreethi, 2017), enhancements to the k-means clustering algorithm were proposed through the application of the Ant Colony Optimization metaheuristic. Two strategies were explored: the first allowed ants to randomly select data points, while in the second, data selection was computationally handled instead of being collected by ants.

In (Okur et al., 2018), the authors proposed a method for segmenting digital mammography images to assist specialists or radiologists in identifying cancerous regions. In this study, the Ant Colony Optimization (ACO) algorithm was integrated with k-means clustering to accurately detect malignant areas in mammographic images.

In (Ruihong & Zhihua, 2020), the authors addressed the clustering of user-required products. This new restriction-based approach is built upon the k-means algorithm, and the Bee Colony (BC) algorithm is proposed to overcome the local optimization problem associated with k-means.

In (Das et al., 2018), a data clustering method was proposed using the Bee Colony Optimization algorithm. To improve efficiency, the authors integrated this method with the k-means algorithm.

In (Balachandran & Nazeer, 2018), a method for outlier detection was introduced. The paper proposed a simple and robust algorithm for outlier detection based on mean and standard deviation. The k-means clustering algorithm and the Bee Colony algorithm were employed, resulting in improved cluster quality.

In (Enríquez-Gaytán et al., 2020), a new approach was presented to identify cluster centers in complex datasets. The Bee Colony metaheuristic was applied, with the objective function formulated using the k-means algorithm. The goal was to design a clustering method for gas detection.

In (Dang et al., 2021), the authors proposed two pilot allocation schemes to reduce interference and enhance the performance of massive MIMO systems free from inter-cell coordination. A Genetic Algorithm was used to determine the test sequence for all users, and the k-medoids algorithm was applied for faster convergence.

In (dos Santos et al., 2019), a hybrid network topology for IoT sensors was proposed using a combination of mesh and star topologies. The aim was to ensure better network coverage for sensor allocation. For network planning, an integration of the Genetic Algorithm and the k-medoids algorithm was suggested.

In (Islam et al., 2019), a novel clustering method was presented. A Hybrid Genetic Algorithm (HGA) was proposed for data clustering, in which a genetic encoding of the clustering problem was introduced. Both k-means and k-medoids algorithms were utilized in the process.

In (Khatami et al., 2016), the authors introduced a new fire detection method using image processing techniques. The Particle Swarm Optimization (PSO) algorithm, along with sampled image pixels, was used to derive the weight matrix of color difference transformation.

The k-medoids algorithm was employed to provide readiness criteria for the PSO method.

In the older study (Zhang et al., 2008), the authors explored spatial clustering with obstacle constraints. PSO was used to compute obstructed distances in a network, while k-medoids was applied for clustering spatial data.

The paper (Garg et al., 2020) focused on anomaly detection in the Internet of Things (IoT). The authors proposed a multi-stage model for anomaly detection aiming to address spatial clustering limitations of DBSCAN in noisy applications. The method integrates the Firefly Algorithm with the k-medoids algorithm for data partitioning.

In (Peng et al., 2014), the focus is on routing in wireless sensor networks. The authors proposed a novel clustering algorithm based on the Ant Colony Optimization (ACO) algorithm and the k-medoids clustering method.

In (Das et al., 2021), a clustering approach was presented using the k-medians-based clustering algorithm in combination with a Genetic Algorithm for clustering data points.

In (Nesamalar & Chandran, 2012), a flexible ligand-protein docking method was proposed, utilizing k-medians clustering integrated with Bee Colony Optimization. The molecular docking problem involves finding the appropriate position and orientation for binding a small ligand molecule to a large receptor molecule. The authors combined the Bee Colony Optimization algorithm with k-medians clustering to address this challenge.

In (Kuo & Nguyen, 2019), the authors tackled the clustering of data with categorical attributes. A novel clustering method was introduced based on the k-modes clustering algorithm in combination with a Genetic Algorithm.

In (Soliman et al., 2012), a clustering approach for categorical data was proposed using the Particle Swarm Optimization (PSO) algorithm along with the k-modes clustering method. The method utilized the FK-Modes algorithm to address uncertainty phenomena and applied PSO to obtain globally optimal solutions.

In (Salama & Freitas, 2013), the authors made two contributions in a proposed approach that categorizes datasets into separate subsets to facilitate asymmetric local Bayesian network classification. The k-modes algorithm was used to generate clusters prior to training the Bayesian network classifier, and the Ant Colony Optimization algorithm was employed for learning in multiple Bayesian networks.

In (Aydilek & Arslan, 2013), a hybrid clustering method based on Fuzzy C-Means (FCM) was proposed, combining Support Vector Regression (SVR) with a Genetic Algorithm (GA). In this approach, FCM clustering parameters, cluster size, and weighting factors were optimized to estimate missing values.

In (Ding, 2016), aiming to address the limitations of the standard FCM algorithm, the authors proposed a kernel-based fuzzy clustering algorithm (KFCM), optimized through a Genetic Algorithm (GA) to enhance fuzzy clustering performance.

In (Jie et al., 2017), to overcome the low efficiency and improve the performance of FCM, two novel fuzzy clustering algorithms were proposed, based on an improved version of a self-adaptive cellular genetic algorithm (IDCGA).

In (Wikaisuksakul, 2014), a multi-objective Genetic Algorithm was introduced to address the problem of data clustering. The given dataset is automatically partitioned into several fuzzy groups using the parameters of the Fuzzy C-Means clustering method.

In (Mirghasemi et al., 2019), a feature enhancement method based on Particle Swarm Optimization (PSO) in the wavelet domain was proposed for segmenting noisy images. This approach employs adaptive wavelet shrinkage, using the FCM clustering function both as an evaluation mechanism and as the segmentation algorithm.

In (Mai et al., 2021), the authors focused on satellite image analysis and proposed a hybrid method incorporating semi-supervised Fuzzy C-Means clustering and Particle Swarm Optimization (PSO).

In (Benaichouche et al., 2013), an improved image segmentation method was introduced using the Fuzzy C-Means clustering algorithm. To enhance the segmentation performance, the PSO algorithm was also applied.

In (Mekhmoukh & Mokrani, 2015), the authors addressed brain image segmentation and proposed a novel method based on Particle Swarm Optimization (PSO) and Fuzzy C-Means (FCM) clustering for enhanced segmentation performance.

In (Dhal et al., 2021), a histogram-based fuzzy C-means clustering method (HBFC) was introduced, incorporating an improved Firefly Algorithm (FA) to enhance segmentation accuracy.

In (Ghosh et al., 2018), the authors applied Fuzzy C-Means clustering along with the Firefly Algorithm to segment brain tissue images more effectively.

In (Langari et al., 2020), to preserve user privacy in social networks—such as protecting databases from identity disclosure and minimizing attack risks—the authors proposed a method combining Fuzzy C-Means clustering and the Firefly Algorithm.

In (Alagarsamy et al., 2019), a hybrid approach combining the Fuzzy C-Means algorithm with the Bat Algorithm was proposed for improved automatic identification of heterogeneous brain tumors.

In (Doğan & Korürek, 2012), to address challenges in continuous domains such as sensitivity to initialization and local optima entrapment, the authors utilized a combination of Fuzzy C-Means clustering and the Ant Colony Optimization (ACO) algorithm.

In (Kaur & Nazeer, 2018), to overcome the limitation of FCM getting trapped in local optima, the authors proposed the use of the Artificial Bee Colony (ABC) optimization algorithm, which is capable of yielding globally optimal solutions.

In (Hassanien et al., 2015), a method was proposed for accurate segmentation of retinal blood vessels using a hybrid of Artificial Bee Colony optimization, Fuzzy C-Means clustering, and pattern matching techniques.

Here is a summary of the reviewed articles in the table below.

Num	**Article author**	**Year**	**Number of citations**	**Type of clustering algorithm**	**Type of metaheuristic algorithm**	**main topic the article**	Ref
1	Ian Paolo Tagorda	2015	5	K-Means	Genetic	Vehicle Routing System	Tagorda et al., 2020
2	Sinuo Pan	2019	8	K-Means	Genetic	UAV Delivery Planning	Pan, 2019
3	BEDIR BEDIR YOUSIF	2020	71	K-Means	Genetic	Vehicle License Plate Recognition	Yousif et al., 2020

4	Shadab Irfan	2017	21	K-Means	Genetic	Optimization of K-Means clustering	Irfan et al., 2017
5	Anand Sukumar J V	2018	83	K-Means	Genetic	Network Intrusion Detection	Sukumar et al., 2018
6	Panya Sunantasaengtong	2014	8	K-Means	Genetic	Indoor Localization System	Sunantasaengtong & Chivapreecha, 2014
7	Ankur Saxena	2018	-	K-Means	Genetic	Validate Cluster Generation	Saxena et al., 2018
8	Assir, A.	2020	3	K-Means	Genetic	Detection of Breast Cancer	Assir et al., 2020
9	Shashi Bhushan	2018	48	K-Means	Genetic	heterogeneous wireless sensor network	Bhushan et al., 2018
10	Li Shi-Wei	2010	4	K-Means	PSO	Date Clustering	Shi-Wei et al., 2010
11	Zhong-kai Feng	2019	165	K-Means	PSO	rule derivation of hydropower reservoir	Feng et al., 2019
12	Mud-Armeen Munlin	2015	6	K-Means	PSO	Traveling Salesman Problem	Munlin & Anantathanavit, 2015
13	Xuexin Chen, Pu Miao	2019	14	K-Means	PSO	Image Segmentation	Chen et al., 2019
14	Lifeng Wu	2021	78	K-Means	FireFly	parallel computation	Wu et al., 2021
15	R.J. Kuo	2016	59	K-Means	FireFly	export trade forecasting	Kuo & Li, 2016
16	Hailun Xie	2019	209	K-Means	FireFly	Improving K-means clustering	Xie et al., 2019
17	Abba Suganda Girsang	2017	11	K-Means	Cuckoo	Clustering Problem	Girsang et al., 2017
18	Amir Arjmand	2019	31	K-Means	Cuckoo	Breast Tumor Segmentation	Arjmand et al., 2019
19	Shuce Ye	2018	17	K-Means	Cuckoo	K-Means Clustering	Ye et al., 2018
20	J. Senthilnath	2016	134	K-Means	Bat	Multispectral Satellite Image Classification	Senthilnath et al., 2016
21	Jack En Gan	2019	10	K-Means	Bat	Automated Grading of Edible Birds Nest	Gan & Lai, 2019
22	Lubna Alhenaki	2019	1	K-Means	Frog Leaping	Large Dataset Document Clustering	Alhenaki & Hosny, 2019
23	Junhua Chen	2019	16	K-Means	Ant Colony	text documents	Chen et al., 2019

24	Wang Jian	2014	2	K-Means	Ant Colony	Sports Performance	Jian et al., 2014
25	Aparna K	2014	18	K-Means	Ant Colony	high dimensional data	Aparna & Nair, 2014
26	Bo Zhao	2007	27	K-Means	Ant Colony	Image Segmentation	Zhao et al., 2007
27	Xu Yang	2015	19	K-Means	Ant Colony	Intrusion Detection	Yang & Hui, 2015
28	T Namratha Reddy	2017	25	K-Means	Ant Colony	Optimization of K-Means Algorithm	Reddy & Supreethi, 2017
29	AGHİ Okur	2018	1	K-Means	Ant Colony	Segmentation On Digital Mammogram Images	Okur et al., 2018
30	Zhang Ruihong	2020	10	K-Means	Bee Colony	Collaborative Filtering Recommendation	Ruihong & Zhihua, 2020
31	Pranesh Das	2018	71	K-Means	Bee Colony	data clustering	Das et al., 2018
32	Anu Balachandran	2018	2	K-Means	Bee Colony	Datasets that Contain Outliers	Balachandran & Nazeer, 2018
33	J Enríquez-Gaytán	2020	6	K-Means	Bee Colony	Gas Sensing	Enríquez-Gaytán et al., 2020
34	Xuan-Toan Dang	2021	18	K-Medoids	Genetic	Pilot Assignment strategy	Dang et al., 2021
35	WGV dos Santos	2019	10	K-Medoids	Genetic	Star-Mesh IoT Network	dos Santos et al., 2019
36	Md. Touhidul Islam	2019	20	K-Medoids	Genetic	Data clustering	Islam et al., 2019
37	Amin Khatami	2016	9	K-Medoids	PSO	Fire Flame Detection	Khatami et al., 2016
38	Xueping Zhang	2008	8	K-Medoids	PSO	Spatial Clustering	Zhang et al., 2008
39	Sahil Garg	2020	125	K-Medoids	FireFly	Security in IoT	Garg et al., 2020
40	Li Peng	2014	16	K-Medoids	Ant Colony	New Clustering Algorithm	Peng et al., 2014
41	Sumita Das	2021	2	K-Median	Genetic	Data Pre-Processing	Das et al., 2021
42	E. Kiruba Nesamalar	2012	4	K-Median	Bee Colony	flexible protein-ligand docking	Nesamalar & Chandran, 2012
43	R. J. Kuo	2019	28	K-Modes	Genetic	Categorical Data	Kuo & Nguyen, 2019
44	Omar S. soliman	2012	5	K-Modes	PSO	hybrid fuzzy particle swarm and fuzzy k-modes clustering	Soliman et al., 2012
45	Khalid M. Salama	2013	17	K-Modes	Ant Colony	Bayesian Multi-net Classifier	Salama & Freitas, 2013

46	Ibrahim Berkan Aydilek	2013	398	Fuzzy C-Mean	Genetic	imputation of missing	Aydilek & Arslan, 2013
47	Yi Ding	2016	263	Fuzzy C-Mean	Genetic	Fuzzy C-Means Clustering	Ding, 2016
48	Lilin Jie	2017	37	Fuzzy C-Mean	Genetic	optimal selection	Jie et al., 2017
49	Siripen Wikaisuksakul	2014	111	Fuzzy C-Mean	Genetic	automatic data clustering	Wikaisuksakul, 2014
50	Saeed Mirghasemi	2019	25	Fuzzy C-Mean	PSO	noisy image segmentation	Mirghasemi et al., 2019
51	Dinh Sinh Mai	2021	61	Fuzzy C-Mean	PSO	Satellite Image Analysis	Mai et al., 2021
52	A.N. Benaichouche	2013	207	Fuzzy C-Mean	PSO	image segmentation	Benaichouche et al., 2013
53	Abdenour Mekhmoukh	2015	130	Fuzzy C-Mean	PSO	MR brain image segmentation	Mekhmoukh & Mokrani, 2015
54	Krishna Gopal Dhal	2021	71	Fuzzy C-Mean	FireFly	fuzzy image clustering	Dhal et al., 2021
55	Partha Ghosh	2018	54	Fuzzy C-Mean	FireFly	Segmentation of Brain Tissues	Ghosh et al., 2018
56	Rohulla Kosari Langari	2020	87	Fuzzy C-Mean	FireFly	privacy preserving in social networks	Langari et al., 2020
57	Saravanan Alagarsamy	2019	66	Fuzzy C-Mean	Bat	Multi-channeled MR brain image segmentation	Alagarsamy et al., 2019
58	Berat Dogan	2012	70	Fuzzy C-Mean	Ant Colony	continuous domains	Doğan & Korürek, 2012
59	Jaspreet Kaur	2018	-	Fuzzy C-Mean	Bee Colony	improved Clustering Algorithm	Kaur & Nazeer, 2018
60	Aboul Ella Hassanien	2015	73	Fuzzy C-Mean	Bee Colony	Retinal blood vessel localization	Hassanien et al., 2015

SUGGESTIONS and Conclusion

One of the future research directions could focus on analyzing the computational time complexity of all the algorithms discussed in this paper. Such analysis can be beneficial in addressing the problem of automatic clustering in large-scale datasets. A key challenge in clustering is that algorithms often do not know the number of clusters in advance, which creates difficulties during computation. Therefore, it is essential to develop methods that can either resolve or significantly reduce this limitation.

Clustering algorithms and metaheuristic algorithms can be applied across a variety of domains. One notable application is text mining, where combining these algorithms can result in more effective text clustering compared to traditional methods.

In this paper, we reviewed studies related to partitional clustering algorithms and metaheuristic algorithms. It was discussed that the combination of metaheuristic techniques and partitional clustering algorithms has been widely used to solve numerous problems. One of the main purposes of using metaheuristic algorithms is to overcome the issue of getting trapped in local optima, which is a common drawback of clustering algorithms. Metaheuristic approaches significantly mitigate this issue and help achieve a global optimal solution.

In recent years, many researchers have employed metaheuristic algorithms to enhance the accuracy of clustering. Given the introduction of several new metaheuristic algorithms in the past few years, it is highly likely that we will witness even broader applications of these techniques in various domains, particularly in the field of clustering, in the near future.

In addition to computational time complexity, future studies could explore the scalability of these clustering algorithms when applied to extremely large datasets. The ability of clustering algorithms to efficiently handle big data is crucial, especially in real-time applications such as recommendation systems or social network analysis. Researchers should also consider integrating advanced techniques such as deep learning to further enhance clustering accuracy and efficiency.

Beyond text mining, the combination of metaheuristic and partitional clustering algorithms shows promise in fields such as bioinformatics, image processing, and customer segmentation. In these domains, the need for effective and scalable clustering solutions is becoming increasingly important as the volume of data continues to grow.

A key factor for the future development of clustering algorithms will be their adaptability to various data structures and types. Research focusing on the hybridization of existing algorithms with new techniques could significantly broaden their applicability, enabling them to better handle heterogeneous, noisy, or incomplete datasets.

REFERENCES

Kenidra, B., Benmohammed, M., Beghriche, A., & Benmounah, Z. (2016, August). A partitional approach for genomic-data clustering combined with k-means algorithm. In 2016 IEEE Intl Conference on Computational Science and Engineering (CSE) and IEEE Intl Conference on Embedded and Ubiquitous Computing (EUC) and 15th Intl Symposium on Distributed Computing and Applications for Business Engineering (DCABES) (pp. 114-121). IEEE.

Velmurugan, T., & Santhanam, T. (2011). A survey of partition based clustering algorithms in data mining: An experimental approach. Information Technology Journal, 10(3), 478-484.

De Castro, L. A. F., & Daniel. (2016). Introdução à Mineração de Dados: Conceitos Básicos, Algoritmos e Aplicações.

Downs, G. A. B., & J. M. (2002). Clustering methods and their uses in computational chemistry. Reviews in Computational Chemistry, 18, 1-40.

Agrawal, R., Gehrke, J., Gunopulos, D., & Raghavan, P. (1998, June). Automatic subspace clustering of high dimensional data for data mining applications. In Proceedings of the 1998 ACM SIGMOD international conference on Management of data (pp. 94-105).

DeRosa, M. (2004). Data mining and data analysis for counterterrorism. Washington, DC: CSIS Press.

Abonyi, J., & Feil, B. (2007). Cluster analysis for data mining and system identification. Springer Science & Business Media.

Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey wolf optimizer. Advances in engineering software, 69, 46-61.

Agrawal, P., Abutarboush, H. F., Ganesh, T., & Mohamed, A. W. (2021). Metaheuristic algorithms on feature selection: A survey of one decade of research (2009-2019). IEEE Access, 9, 26766-26791.

Nanda, S. J., & Panda, G. (2014). A survey on nature inspired metaheuristic algorithms for partitional clustering. Swarm and Evolutionary Computation, 16, 1-18.

Sastry, K., Goldberg, D., & Kendall, G. (2005). Genetic algorithms. In E. K. Burke & G. Kendall (Eds.), Search Methodologies: Introductory Tutorials in Optimization and Decision Support Techniques (pp. 97-125). Springer.

Assunção, W. K. G., Pozo, A., & Vendramin, A. C. B. K. (2011). Application of bio-inspired metaheuristics in the data clustering problem. CLEI Electronic Journal, 14(3).

Maulik, U., & Bandyopadhyay, S. (2000). Genetic algorithm-based clustering technique. Pattern Recognition, 33(9), 1455-1465.

Russell, S. J., & Norvig, P. (2003). Artificial Intelligence: A Modern Approach (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Dorigo, M. (1992). Optimization, learning and natural algorithms (PhD thesis). Politecnico di Milano.

Dorigo, M. (1994). Learning by probabilistic Boolean networks. In Proceedings of the IEEE International Conference on Neural Networks (pp. 887-891).

Linyi, L. I., & Deren, L. I. (2010, June). Fuzzy classification of remote sensing images based on particle swarm optimization. In 2010 International Conference on Electrical and Control Engineering (pp. 1039-1042). IEEE.

Kennedy, J., & Eberhart, R. (2001). Swarm Intelligence. San Diego, CA: Academic Press.

Mahapatra, S., Jagadev, A. K., & Nai, B. (2011). Performance evaluation of PSO-based classifier for classification of multidimensional data with variation of PSO parameters in knowledge discovery database. International Journal of Advanced Science and Technology, 34, 93-102.

Veeramachaneni, K., Yan, W., Goebel, K., & Osadciw, L. (2007, April). Improving classifier fusion using particle swarm optimization. In Computational Intelligence in Multicriteria Decision Making, IEEE Symposium on (pp. 128-135). IEEE.

Shahana, K., Ghosh, S., & Jeganathan, C. (2016, April). A survey of particle swarm optimization and random forest-based land cover classification. In 2016 International Conference on Computing, Communication and Automation (ICCCA) (pp. 241-245). IEEE.

Yang, X. S. (2008). Nature-Inspired Metaheuristic Algorithms. Luniver Press.

Yang, X. S. (2009). Firefly algorithms for multimodal optimization. SAGA, 2009.

Jaradat, A., & Banihamad, A. (2018). Community structure detection using firefly algorithm. International Journal of Applied Metaheuristic Computing, 9(4), 52-70.

Jaradat, A., & Diabat, W. (2019, April). Solving traveling salesman problem using firefly algorithm and k-means clustering. In 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT) (pp. 586-589). IEEE.

Tiwari, V. (2012). Face recognition based on cuckoo search algorithm. Image, 7(8), 9.

Rodrigues, D., Pereira, L. A. M., Almeida, T. N. S., Papa, J. P., Souza, A. N., Ramos, C. C. O., & Yang, X.-S. (2013). BCS: A binary cuckoo search algorithm for feature selection. In Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS) (pp. 465-468).

Nakamura, R. Y. M., Pereira, L. A. M., Costa, K. A., Rodrigues, D., Papa, J. P., & Yang, X.-S. (2012). BBA: A binary bat algorithm for feature selection. In Proceedings of the 25th SIBGRAPI Conference on Graphics, Patterns and Images (pp. 291-297).

Yang, X., & Gandomi, A. H. (2012). Bat algorithm: A novel approach for global engineering optimization. Engineering Computations, 29(5), 464-483.

Kaufman, L., & Rousseeuw, P. J. (1990). Partitioning Around Medoids (Program PAM). Wiley Series in Probability and Statistics. Hoboken, NJ: John Wiley & Sons, Inc. (pp. 68-125).

De León-Aldaco, S. E., Calleja, H., & Alquicira, J. A. (2015). Metaheuristic optimization methods applied to power converters: A review. IEEE Transactions on Power Electronics, 30(12), 6791-6803.

Chong, E. K. P., & Zak, S. H. (2001). Introduction to Optimization (2nd ed.). Wiley-Interscience.

Yang, X. S. (2010). Firefly algorithm, stochastic test functions and design optimization. International Journal of Bio-Inspired Computation, 2(3), 133-144.

Yang, X. S. (2009). Firefly algorithm for multimodal optimization. In Stochastic Algorithms: Foundations and Applications (SAGA) (pp. 169-178). Springer Lecture Notes in Computer Science.

Hassanzadeh, T., & Meybodi, M. R. (2012, May). A new hybrid approach for data clustering using firefly algorithm and K-means. In The 16th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP 2012) (pp. 007-011). IEEE.

Girsang, A. S., Yunanto, A., & Aslamiah, A. H. (2017, August). A hybrid cuckoo search and K-means for clustering problem. In 2017 International Conference on Electrical Engineering and Computer Science (ICECOS) (pp. 120-124). IEEE.

Richardson, P., & Elphick, J. (2011). Bats. The Natural History Museum.

Gan, J. E., & Lai, W. K. (2019, June). Automated grading of edible birds nest using hybrid bat algorithm clustering based on K-means. In 2019 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS) (pp. 73-78). IEEE.

Bhaduri, A., & Bhaduri, A. (2009, October). Color image segmentation using clonal selection-based shuffled frog leaping algorithm. In 2009 International Conference on Advances in Recent Technologies in Communication and Computing (pp. 517-520). IEEE.

Eusuff, M., Lansey, K., & Pasha, F. (2006). Shuffled frog-leaping algorithm: A memetic meta-heuristic for discrete optimization. Engineering Optimization, 38(2), 129-154.

Karaboga, D., & Akay, B. (2009). A comparative study of artificial bee colony algorithm. Applied Mathematics and Computation, 214(1), 108-132.

Nesamalar, E. K., & Chandran, C. P. (2012, March). Genetic clustering with bee colony optimization for flexible protein-ligand docking. In International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME-2012) (pp. 82-87). IEEE.

Whelan, C., Harrell, G., & Wang, J. (2015). Understanding the K-medians problem. In Proceedings of the International Conference on Scientific Computing (CSC) (p. 219). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).

Kuo, R. J., Zheng, Y. R., & Nguyen, T. P. Q. (2021). Metaheuristic-based possibilistic fuzzy K-modes algorithms for categorical data clustering. Information Sciences, 557, 1-15.

Aydilek, I. B., & Arslan, A. (2013). A hybrid method for imputation of missing values using optimized fuzzy C-means with support vector regression and a genetic algorithm. Information Sciences, 233, 25-35.

Tagorda, I. P., Calata, L. E., Limjoco, W. J. R., & Dizon, C. C. (2020, November). Development of a vehicle routing system for delivery services. In 2020 IEEE Region 10 Conference (TENCON) (pp. 1187-1191). IEEE.

Pan, S. (2019, August). UAV delivery planning based on K-means++ clustering and genetic algorithm. In 2019 5th International Conference on Control Science and Systems Engineering (ICCSSE) (pp. 14-18). IEEE.

Yousif, B. B., Ata, M. M., Fawzy, N., & Obaya, M. (2020). Toward an optimized neutrosophic K-means with genetic algorithm for automatic vehicle license plate recognition (ONKM-AVLPR). IEEE Access, 8, 49285-49312.

Irfan, S., Dwivedi, G., & Ghosh, S. (2017, October). Optimization of K-means clustering using genetic algorithm. In 2017 International Conference on Computing and Communication Technologies for Smart Nation (IC3TSN) (pp. 156-161). IEEE.

Sukumar, J. A., Pranav, I., Neetish, M. M., & Narayanan, J. (2018, September). Network intrusion detection using improved genetic K-means algorithm. In 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI) (pp. 2441-2446). IEEE.

Sunantasaengtong, P., & Chivapreecha, S. (2014, October). Mixed K-means and GA-based weighted distance fingerprint algorithm for indoor localization system. In TENCON 2014-2014 IEEE Region 10 Conference (pp. 1-5). IEEE.

Saxena, A., Pathik, N., & Gupta, R. (2018, July). Improved variance K-means algorithm using multi-objective genetic algorithm for validate cluster generation. In 2018 International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering (ICRIEECE) (pp. 1782-1787). IEEE.

Assir, A., Harmouchi, M., Lyazidi, A., Rattal, M., & Mouhsen, A. (2020, April). Fully automatic computer-aided detection of breast cancer based on genetic algorithm optimization. In 2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET) (pp. 1-6). IEEE.

Bhushan, S., Pal, R., & Antoshchuk, S. G. (2018, August). Energy efficient clustering protocol for heterogeneous wireless sensor network: A hybrid approach using GA and K-means. In 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP) (pp. 381-385). IEEE.

Shi-Wei, L., & Xiao-Dong, Q. (2010, August). Date clustering using principal component analysis and particle swarm optimization. In 2010 5th International Conference on Computer Science & Education (pp. 493-497). IEEE.

Feng, Z. K., Niu, W. J., Zhang, R., Wang, S., & Cheng, C. T. (2019). Operation rule derivation of hydropower reservoir by K-means clustering method and extreme learning machine based on particle swarm optimization. Journal of Hydrology, 576, 229-238.

Munlin, M. A., & Anantathanavit, M. (2015, June). Hybrid K-means and particle swarm optimization for symmetric traveling salesman problem. In 2015 IEEE 10th Conference on Industrial Electronics and Applications (ICIEA) (pp. 671-676). IEEE.

Chen, X., Miao, P., & Bu, Q. (2019, July). Image segmentation algorithm based on particle swarm optimization with K-means optimization. In 2019 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS) (pp. 156-159). IEEE.

Wu, L., Peng, Y., Fan, J., Wang, Y., & Huang, G. (2021). A novel kernel extreme learning machine model coupled with K-means clustering and firefly algorithm for estimating monthly reference evapotranspiration in parallel computation. Agricultural Water Management, 245, 106624.

Kuo, R. J., & Li, P. S. (2016). Taiwanese export trade forecasting using firefly algorithm based K-means algorithm and SVR with wavelet transform. Computers & Industrial Engineering, 99, 153-161.

Xie, H., Zhang, L., Lim, C. P., Yu, Y., Liu, C., Liu, H., & Walters, J. (2019). Improving K-means clustering with enhanced firefly algorithms. Applied Soft Computing, 75, 16-30.

Girsang, A. S., Yunanto, A., & Aslamiah, A. H. (2017, August). A hybrid cuckoo search and K-means for clustering problem. In 2017 International Conference on Electrical Engineering and Computer Science (ICECOS) (pp. 120-124). IEEE.

Arjmand, A., Meshgini, S., Afrouzian, R., & Farzamnia, A. (2019, October). Breast tumor segmentation using K-means clustering and cuckoo search optimization. In 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE) (pp. 305-308). IEEE.

Ye, S., Huang, X., Teng, Y., & Li, Y. (2018, March). K-means clustering algorithm based on improved cuckoo search algorithm and its application. In 2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA) (pp. 422-426). IEEE.

Senthilnath, J., Kulkarni, S., Benediktsson, J. A., & Yang, X. S. (2016). A novel approach for multispectral satellite image classification based on the bat algorithm. IEEE Geoscience and Remote Sensing Letters, 13(4), 599-603.

Gan, J. E., & Lai, W. K. (2019, June). Automated grading of edible bird’s nest using hybrid bat algorithm clustering based on K-means. In 2019 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS) (pp. 73-78). IEEE.

Alhenaki, L., & Hosny, M. (2019, November). A genetic-frog leaping algorithm for large dataset document clustering. In 2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA) (pp. 1-4). IEEE.

Chen, J., Tian, M., Qi, X., Wang, W., & Liu, Y. (2019). A solution to reconstruct cross-cut shredded text documents based on constrained seed K-means algorithm and ant colony algorithm. Expert Systems with Applications, 127, 35-46.

Jian, W., Zhi-Hua, H., & Zhi-Yong, Z. (2014, June). Clustering analysis of sports performance based on ant colony algorithm. In 2014 Fifth International Conference on Intelligent Systems Design and Engineering Applications (pp. 288-291). IEEE.

Aparna, K., & Nair, M. K. (2014, February). Enhancement of K-Means algorithm using ACO as an optimization technique on high dimensional data. In 2014 International Conference on Electronics and Communication Systems (ICECS) (pp. 1-5). IEEE.

Zhao, B., Zhu, Z., Mao, E., & Song, Z. (2007, August). Image segmentation based on ant colony optimization and K-means clustering. In 2007 IEEE International Conference on Automation and Logistics (pp. 459-463). IEEE.

Yang, X., & Hui, Z. (2015, August). Intrusion detection alarm filtering technology based on ant colony clustering algorithm. In 2015 Sixth International Conference on Intelligent Systems Design and Engineering Applications (ISDEA) (pp. 470-473). IEEE.

Reddy, T. N., & Supreethi, K. P. (2017, July). Optimization of K-means algorithm: Ant colony optimization. In 2017 International Conference on Computing Methodologies and Communication (ICCMC) (pp. 530-535). IEEE.

Okur, A. G. H. İ., Görgel, Ö. Ü. P., & Sertbas, A. (2018, April). Segmentation on digital mammogram images using hybrid K-means clustering and ant colony algorithms. In 2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT) (pp. 1-4). IEEE.

Ruihong, Z., & Zhihua, H. (2020). Collaborative filtering recommendation algorithm based on bee colony K-means clustering model. Microprocessors and Microsystems, 103424.

Das, P., Das, D. K., & Dey, S. (2018). A modified bee colony optimization (MBCO) and its hybridization with K-means for an application to data clustering. Applied Soft Computing, 70, 590-603.

Balachandran, A., & Nazeer, K. A. (2018, September). An improved clustering algorithm based on K-means and artificial bee colony optimization for datasets that contain outliers. In 2018 International Conference on Computing, Power and Communication Technologies (GUCON) (pp. 1083-1088). IEEE.

Enríquez-Gaytán, J., Gómez-Castañeda, F., Moreno-Cadenas, J. A., & Flores-Nava, L. M. (2020, November). A clustering method based on the artificial bee colony algorithm for gas sensing. In 2020 17th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE) (pp. 1-4). IEEE.

Dang, X. T., Lai-Thuc, T., Nguyen, A. T., Vu-Huy, T., Tran, N. H. A., & Han, H. D. (2021, January). A genetic algorithm-based pilot assignment strategy for cell-free massive MIMO system. In 2020 IEEE Eighth International Conference on Communications and Electronics (ICCE) (pp. 93-98). IEEE.

dos Santos, W. G., Costa, W. S., Faber, M. J., Silva, J. A., Rocha, H. R., & Segatto, M. E. (2019, November). Sensor allocation in a hybrid star-mesh IoT network using genetic algorithm and K-medoids. In 2019 IEEE Latin-American Conference on Communications (LATINCOM) (pp. 1-6). IEEE.

Islam, M. T., Basak, P. K., Bhowmik, P., & Khan, M. (2019). Data clustering using hybrid genetic algorithm with K-means and K-medoids algorithms. In 2019 23rd International Computer Science and Engineering Conference (ICSEC) (pp. 123-128). IEEE.

Khatami, A., Mirghasemi, S., Khosravi, A., Lim, C. P., & Nahavandi, S. (2016). A new K-medoids clustering and swarm intelligence approach to fire flame detection. Expert Systems with Applications, 43, 68-80.

Zhang, X., Ding, W., Wang, J., Fan, Z., & Deng, G. (2008, November). Spatial clustering with obstacles constraints using PSO-DV and K-medoids. In 2008 3rd International Conference on Intelligent System and Knowledge Engineering (Vol. 1, pp. 246-251). IEEE.

Garg, S., Kaur, K., Batra, S., Kaddoum, G., Kumar, N., & Boukerche, A. (2020). A multi-stage anomaly detection scheme for augmenting the security in IoT-enabled applications. Future Generation Computer Systems, 104, 105-118.

Peng, L., Dong, G. Y., Dai, F. F., & Liu, G. P. (2014). A new clustering algorithm based on ACO and K-medoids optimization methods. IFAC Proceedings Volumes, 47(3), 9727-9731.

Das, S., Tazim, M. Z., Akter, J., & Khan, M. (2021, March). Influence of data pre-processing on genetic algorithm-based clustering. In 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 23-27). IEEE.

Nesamalar, E. K., & Chandran, C. P. (2012, March). Genetic clustering with bee colony optimization for flexible protein-ligand docking. In International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME-2012) (pp. 82-87). IEEE.

Kuo, R. J., & Nguyen, T. P. Q. (2019). Genetic intuitionistic weighted fuzzy K-modes algorithm for categorical data. Neurocomputing, 330, 116-126.

Soliman, O. S., Saleh, D. A., & Rashwan, S. (2012, May). A hybrid fuzzy particle swarm and fuzzy K-modes clustering algorithm. In 2012 8th International Conference on Informatics and Systems (INFOS) (pp. BIO-68). IEEE.

Salama, K. M., & Freitas, A. A. (2013, June). Clustering-based Bayesian multi-net classifier construction with ant colony optimization. In 2013 IEEE Congress on Evolutionary Computation (pp. 3079-3086). IEEE.

Aydilek, I. B., & Arslan, A. (2013). A hybrid method for imputation of missing values using optimized fuzzy C-means with support vector regression and a genetic algorithm. Information Sciences, 233, 25-35.

Ding, Y., & Fu, X. (2016). Kernel-based fuzzy C-means clustering algorithm based on genetic algorithm. Neurocomputing, 188, 233-238.

Jie, L., Liu, W., Sun, Z., & Teng, S. (2017). Hybrid fuzzy clustering methods based on improved self-adaptive cellular genetic algorithm and optimal-selection-based fuzzy C-means. Neurocomputing, 249, 140-156.

Wikaisuksakul, S. (2014). A multi-objective genetic algorithm with fuzzy C-means for automatic data clustering. Applied Soft Computing, 24, 679-691.

Mirghasemi, S., Andreae, P., & Zhang, M. (2019). Domain-independent severely noisy image segmentation via adaptive wavelet shrinkage using particle swarm optimization and fuzzy C-means. Expert Systems with Applications, 133, 126-150.

Mai, D. S., Ngo, L. T., & Hagras, H. (2021). A hybrid interval type-2 semi-supervised possibilistic fuzzy C-means clustering and particle swarm optimization for satellite image analysis. Information Sciences, 548, 398-422.

Benaichouche, A. N., Oulhadj, H., & Siarry, P. (2013). Improved spatial fuzzy C-means clustering for image segmentation using PSO initialization, Mahalanobis distance and post-segmentation correction. Digital Signal Processing, 23(5), 1390-1400.

Mekhmoukh, A., & Mokrani, K. (2015). Improved fuzzy C-means based particle swarm optimization (PSO) initialization and outlier rejection with level set methods for MR brain image segmentation. Computer Methods and Programs in Biomedicine, 122(2), 266-281.

Dhal, K. G., Das, A., Ray, S., & Gálvez, J. (2021). Randomly attracted rough firefly algorithm for histogram-based fuzzy image clustering. Knowledge-Based Systems, 216, 106814.

Ghosh, P., Mali, K., & Das, S. K. (2018). Chaotic firefly algorithm-based fuzzy C-means algorithm for segmentation of brain tissues in magnetic resonance images. Journal of Visual Communication and Image Representation, 54, 63-79.

Langari, R. K., Sardar, S., Mousavi, S. A. A., & Radfar, R. (2020). Combined fuzzy clustering and firefly algorithm for privacy preserving in social networks. Expert Systems with Applications, 141, 112968.

Alagarsamy, S., Kamatchi, K., Govindaraj, V., Zhang, Y. D., & Thiyagarajan, A. (2019). Multi-channeled MR brain image segmentation: A new automated approach combining BAT and clustering technique for better identification of heterogeneous tumors. Biocybernetics and Biomedical Engineering, 39(4), 1005-1035.

Doğan, B., & Korürek, M. (2012). A new ECG beat clustering method based on kernelized fuzzy C-means and hybrid ant colony optimization for continuous domains. Applied Soft Computing, 12(11), 3442-3451.

Kaur, J., & Nazeer, K. A. (2018, September). An improved clustering algorithm based on fuzzy C-means and artificial bee colony optimization. In 2018 International Conference on Computing, Power and Communication Technologies (GUCON) (pp. 1089-1094). IEEE.

Hassanien, A. E., Emary, E., & Zawbaa, H. M. (2015). Retinal blood vessel localization approach based on bee colony swarm optimization, fuzzy C-means and pattern search. Journal of Visual Communication and Image Representation, 31, 186-196.

Maulik, U., & Bandyopadhyay, S. (2000). Genetic algorithm-based clustering technique. Pattern Recognition, 33(9), 1455-1465.

Xiao, Y., Wang, L., & Wang, H. (2008). An efficient PSO-based K-means algorithm for clustering. Journal of Software, 3(5), 91-97.

Omran, M. G. H., Engelbrecht, A. P., & Salman, A. (2007). An overview of clustering methods. Intelligent Data Analysis, 11(6), 583–605.