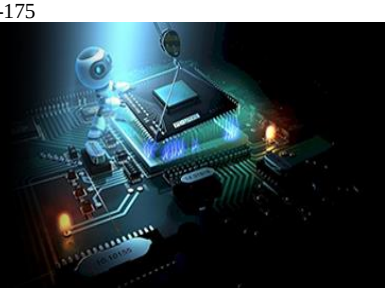
International Journal of Engineering in Computer Science



E-ISSN: 2663-3590
P-ISSN: 2663-3582
Impact Factor (RJIF): 5.52
www.computersciencejournals.com/ijecs
IJECS 2025; 7(2): 166-175
Received: 22-07-2025
Accepted: 26-08-2025

Huda Yass Khudhair
Computer Engineering Techniques Department, Engineering Technical College, University of Imam Jaafar Al-Sadiq, Thi-Qar Branch, Thi-Qar, Iraq

Ahmed M
Betti, Computer Engineering Techniques Department, Engineering Technical College, University of Imam Jaafar Al-Sadiq, Thi-Qar Branch, Thi-Qar, Iraq

Corresponding Author:
Huda Yass Khudhair
Computer Engineering Techniques Department, Engineering Technical College, University of Imam Jaafar Al-Sadiq, Thi-Qar Branch, Thi-Qar, Iraq

Improve K-Means algorithm to increase the accuracy of data classification in the internet of Things using ACO and PSO algorithms

Huda Yass Khudhair and Ahmed M Betti

DOI: https://www.doi.org/10.33545/26633582.2025.v7.i2b.213

Abstract
This study aims to analyze the data which received in the community as much as possible given the growing prevalence of the Internet of Things. It was essential to employ practical and effective methodologies. It acted truly deals with the Internet-based interconnection and communication of various items. The IoT's biggest problem was managing the data sent to it because of the diversity and volume of data that it received. Utilizing data mining and its numerous algorithms, including information clustering, was one of the most effective ways to handle this problem. Information clustering was one of the most useful techniques for categorizing IoT data. One of the key data mining techniques wais The Kmeans algorithm, which had drawbacks if additional algorithms were offered, was one of the most crucial ways for data clustering, a task for which several approaches had been presented. In the field of data mining, which, in addition to the use of IoT data mining, would be beneficial in other areas connected to data mining, might be of great assistance in resolving or minimizing its issues. Additionally, the outcomes of using the suggested technique and evaluating it against existing algorithms to improve clustering performance are shown.

Keywords: Performance improvement, K-Means algorithm, increased accuracy, data classification, IoT

1.1 Introduction
The Internet of Things, in its most basic form, just links different objects together and enables communication between them. The Internet of Things has several challenges, but the most important of them is managing the volume and variety of incoming data. Management is the information that is given to it. The use of data mining and its multiple algorithms, among which information clustering methods are among the most useful approaches in classifying huge data, is one of the most critical ways to address the aforementioned problem. They will use the Internet of Things.
The idea of "Internet of Things" is quite new. ICT or information and communication technology can be defined as an advanced and developing field of study that enables widespread communication and the transmission and reception of data through networks such as the Internet or intranet to be [21]. Due to the huge amount of data collected by the Internet of Things, it is necessary to manage and organize it in to use the stored information. In other words, without data mining the data collected in the Internet of Things, we cannot use them to their full potential. Clustering is one of the most popular data mining techniques among many available methods. Clustering is a subfield of unsupervised learning that automatically classifies data into groups of individuals that share common characteristics. They are referred to as clusters. Therefore, an object cluster is a set of entities that are related to each other, but not connected to other object clusters. For example, a distance measure may be applied to clustering. Depending on this type of clustering, items that are closer to each other may be considered as a cluster. For similarity, several factors may be considered. Distance is its name. Several algorithms have been proposed to reduce the drawbacks and increase the benefits of this method as a result of its widespread use in recent years. The following researches are the most important [23].
K-means technique, which has wide applications in various industries, is one of the most important clustering algorithms. This method includes problems such as being sensitive to

noisy and outlier data and dependence of the final output on the initialization, despite advantages such as simplicity of implementation and high speed. In recent years, these problems have been solved by combining k-means with different methods. Harmonic algorithms, PSO, PSOKHM and ACO are among the most important examples of hybrid algorithms. While each of these algorithms has succeeded in solving some of the drawbacks of the k-means algorithm, an ideal approach that can solve all these drawbacks by combining multiple algorithms has not yet been established. In this research, a hybrid approach to address k-means algorithm problems using bee swarm (PSO) and ant colony (ACO) algorithms is investigated. Iris, Wine and Glass datasets are also used to evaluate the proposed method used

Background of the study

Due to the huge amount of data collected by the Internet of Things, it is necessary to manage and organize it in order to use the stored information. In other words, without data mining the data collected in the Internet of Things, we cannot use them to their full potential [2]. Clustering is one of the most popular data mining techniques among many available methods. Clustering, an automated method in which samples are classified into groups with similar members (called clusters), is a subset of unsupervised learning. Consequently, an object cluster is a collection of items where each item is unique but has common characteristics. Different standards of comparison can be considered. For example, distance measures may be used for clustering, and based on this form of clustering, items that are closer to each other may be considered as a cluster. Its name is distance [9].

Several algorithms have been proposed to reduce the drawbacks and increase the benefits of this method as a result of its widespread use in recent years. The study of K-means algorithm has created an optimal method [25]. Choosing a random sample of points that is equal to the number of required clusters is the first step in this method. Then, the data is assigned to one of these clusters based on their close relationship, which leads to the generation of new clusters. This procedure can be repeated to average the data, creating new centers for each iteration, and assigning the data to new clusters. This method is repeated until the center of the clusters does not change. Among its advantages is the simplicity and ease of implementation, high speed and suitable for large data sets, and its disadvantages include the need to determine the number of clusters at the beginning, sensitivity to noisy and outlier data, and the dependence of the results. The final value of the initial centers and the number of clusters is the algorithm getting stuck in the local optimum and premature convergence. By focusing on the distance between points, a new algorithm has been presented that solves the problem of the K-means algorithm to some extent [20]. It has been proven that this algorithm does not have the problem of initialization of the center, which in this case has a significant improvement in the quality of the clustering results compared to the k-means algorithm, but the tendency to converge to local optima is one of the disadvantages of this method. Using a method used in nature, an optimal algorithm for clustering was found [8]. Particle Swarm Optimization Algorithm This was proposed as a non-deterministic search method for functional optimization. This algorithm is inspired by the collective movement of birds looking for food. A group of birds are randomly looking for food in a space.

There is only one piece of food in the space in question. None of the birds know where the food is. One of the best strategies can be to follow the bird that has the shortest distance to the food. This strategy is actually the essence of the algorithm. Each solution called a particle in the PSO algorithm is equivalent to a bird in the bird collective movement algorithm. Each particle has a merit value which is calculated by the merit function. The closer the particle is to the target (food in the bird movement model) in the search space, the more merit it has. Each particle continues to move in the problem space by following the optimal particles in the current state. This algorithm does not have the problem of initialization of the center, but the most important disadvantage of this method is its slow speed. By combining PSO and KHM methods, an algorithm was created to achieve the optimal method in clustering [7]. Although the greedy search feature of the KHM algorithm requires less performance evaluation than the PSO approach, it often falls prey to the local optimization trap. The PSOKHM hybrid clustering approach aims to optimize the benefits of KHM and PSO. This KHM algorithm is repeated four times in each generation of 8 generations, and for this reason, the PSO algorithm is used and repeated eight times in each generation. By examining the process used to locate food for ants, one of the optimization techniques for clustering problems has been discovered and is based on the mechanism of ant colony disorder analysis. Its goal is to assign optimal values by minimizing objective functions [5]. This algorithm has a lower error rate than other methods, and it is also suitable for grouping data with high dimensions and mass into several clusters. Also, this algorithm can achieve global optimal solutions with the help of random global search.

The purpose of the study

The purpose of this research is to develop a new algorithm by analyzing the current algorithms, which adequately overcomes the shortcomings of the previous techniques, some of which are mentioned below.

1. Discover the strategy to select the first ideal cluster. Find a way to run the algorithm faster on big data
2. Presenting a final method to improve clustering accuracy.

Internet of Things

The basic idea behind the formation of the Internet of Things was to connect all objects around the world through the Internet platform, in which case objects can make independent decisions to do things. Although it seems impossible to connect all objects through the Internet, but with the growth of science and technology, this theory is being realized. With the increasing expansion of the Internet of Things at different levels, new challenges and issues related to it are also expanding. the most important of which is the presence of massive data received in it and how to convert these data into the outputs required by the user, this is where the need to use data mining to discover new knowledge is raised. This technology will provide possible solutions for discovering hidden data [6]. The data produced by IOT1 can be considered a type of Big Data that requires new tools and methods for data analysis. There are several ways to deal with the issue of data explosion in IOT, such as

limiting the received data or reducing the dimensions of the data with the use of distributed computing is presented, the most important of which is the use of data mining techniques. According to the characteristics of the data available in IOT, such as the variety and high volume, conventional data mining methods are not able to solve the problem, and we need changes and design of new data mining methods to extract new knowledge from the available information, various methods to reduce the complexity of data in IOT has been proposed, one of the most important of which is the use of clustering and classification of sequential models. In this study, an attempt has been made to examine the most important existing clustering algorithms and provide a workaround to improve this algorithm [11].

Description of the optimization problem

Optimization has been one of the important research fields in recent decades, the result of which was the design of different types of algorithms. Optimizing, changing the inputs and characteristics of a device in such a way as to provide the optimal output or result. The objective function, cost function, or fitness function under consideration is a process whose inputs are variable. Cost, profit and efficiency are other definitions of efficiency. Most optimization problems are thought to involve the minimization of a cost function. It is straightforward to show that any type of optimization problem may be expressed as a minimization problem. The optimization problem has the following typical form:

$$\max/\min \; z = f(X)$$

$$subject\ to: \begin{cases} g_i(X) \le 0 \; ; i = 1,2,\ldots,M \\ h_j(X) = 0 \; ; j = 1,2,\ldots,P \end{cases}$$

$$X = [x_1, x_2, \ldots, x_n] \qquad (1.3)$$

So that:, is the desired function that we want to minimize.
The constraint is called an inequality.
It is called an equality constraint.
By convention, standard form describes a minimization problem. A maximization problem can be obtained by making the objective function negative. In general, the objective functions are divided into three categories:

1. Inseparable objective function: A function is called inseparable if it cannot be written as a sum of several separate functions. Finding the global optimal point of non-separable objective functions is very difficult.
2. Multimodal objective function: A function is multimodal if it has 2 or more local optimal points. Finding the global optimal point of these functions is very difficult, and this complexity increases when the optimal points are spread over the entire search space.
3. Non-derivable objective function: An objective function is non-derivable if it is non-derivable in each of the points of its solution space.

Optimization problems are divided into different categories from different points of view. Figure (1-2) is shown some of these categories. None of these branches are completely separate from each other. For example, a dynamic optimization problem may be constrained or unconstrained. Furthermore, some variables may be continuous while others may be discrete.

Clustering

Unsupervised learning and supervised learning are two basic categories into which learning is often classified. In supervised learning, the categories are defined from the beginning and each part of the training data is categorized. That is, they say that in addition to the educational materials, there is a supervisor who informs the student [10]. However, with unsupervised learning, the learner only has access to the training data. Hence, it is up to them to look for a specific structure there. One of the subfields of unsupervised learning is clustering. Clusters are automatically created during this procedure, which divides the samples into groups whose members are similar to each other [13]. Therefore, a cluster is a group of things that are related to each other but distinct from the objects in other clusters. For example, the distance measure may be used for clustering, and objects that are closer to each other can be seen as a cluster based on this type of clustering. Different criteria can be considered for similarity. Also known as separation. As an illustration, the input samples in Figure 2-1 are separated into four clusters, as seen in the image on the right. Each input instance in this example belongs to one of the clusters. No instance is a member of more than one cluster.

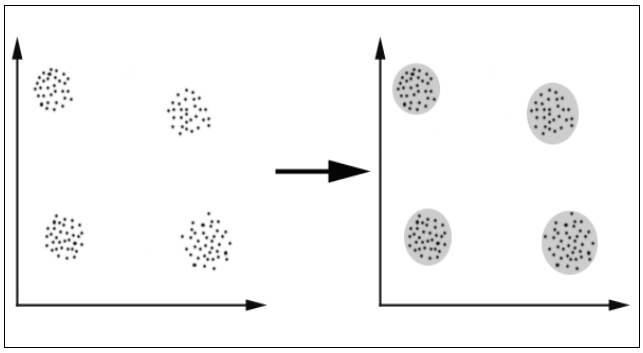


Fig 2.2: Clustering of input samples

As mentioned earlier, clustering identifies comparable groups of data without any prior information. In fact, clustering algorithms often consider a number of initial representatives for input samples, then determine which cluster the sample belongs to base on the sample's similarity to these representatives. After this step, new representatives are calculated for each cluster and samples are compared to these representatives once again to determine which cluster they belong to. Unless the cluster representatives do not change. There are several clustering techniques that have been introduced up to this point that can be used depending on the application. However, despite the wide variety of clustering techniques, there is still no approach that can accurately recognize many types of clusters. As a result, it is up to the user to choose the right technique for their needs

1.7 The difference between clustering and classification

Each piece of data is classified into one of a number of classes. Systems that rely on categorization really have two sets of inputs. The system is trained using these data, or in

other words, parameters that are added to the training set along with the data that is divided into separate categories by default. It assigns categories to itself. The system receives inputs to select the category after the training phase and forms another category [14].

Classification is not the same as clustering. The input samples are labeled in their classification, but have no primary label in the clustering. And indeed, it is through the use of clustering techniques that related data are found and implicitly categorized. In fact, it is possible to perform a clustering operation on the samples before classifying the data, and after that, the centers of the resulting clusters are calculated, labeled, and the classification process ends. Done for fresh input samples.

1.8 clustering applications: We often have no knowledge of the data, so the decision maker must make as many assumptions about the data as possible. The clustering process is considered as a reasonable option to identify meaningful correlations between data within this limitation and may be referred to as fundamental data understanding. Currently, clustering is used in many different disciplines, including those listed below.

1. Mechanical engineering, electrical engineering, machine learning and computational vision are all examples of engineering. Engineering clustering is used in various fields such as biometric identification, voice recognition, radar signal analysis, data summarization and noise removal.
2. Computer science: we see many uses of clustering in image processing (segmentation of satellite or medical images), spatial database analysis, information retrieval and web mining (classification of documents or classification of customers to sites). We have been.
3. Medical and biological sciences including pathology, microbiology, paleontology, psychology, genetics and biology (classification of organisms based on their characteristics). Basic applications of clustering are covered in these topics, which also serve as one of the key application areas for clustering methods. Applications that are very important include defining taxonomy, identifying protein and gene function, diagnosing and treating disease, etc.
4. Geography, geology and mobile reception are earth sciences and astronomy. Clustering may be used to group stars and planets, examine rivers and mountains, divide regions and cities, examine the structure of the globe, and classify houses according to their style and location. It is better to use seismic surveys that identify accident-prone places based on previous findings.
5. Social sciences: education, bibliotherapy (classification of books), psychology, anthropology, sociology, etc., behavioral pattern analysis, language evolution structure, social network analysis, archaeological diagnosis, artificial classification and psychological research Crime from other fields of attractive programs
6. Economy: trade and marketing. Cluster analysis is useful for a variety of tasks, including finding shopping trends, grouping businesses, and examining storage habits.

1.9 Types of clusters

- Clusters with clear boundaries between points: the points inside this cluster are more comparable to each other than the points outside it.
- Center-based clusters: This cluster has a group of points that are significantly closer to its center than the centers of other clusters.
- Clusters based on proximity: the set of points inside this cluster are more similar to one or more other points inside the cluster than the points outside.

1.9.1 Types **of clustering methods**

Clustering methods are introduced in different types and in different classifications. These methods can be divided into several aspects [12].

1. Exclusive clustering and overlapping clustering

In the exclusive method after clustering, each data belongs to exactly one cluster, like the k-means clustering method. But in clustering with the overlapping method, after clustering, each data is assigned a degree of belonging to each cluster, that is, one data can belong to several clusters with different ratios. An example of this method is fuzzy clustering.

2. The techniques categories

Different types and classes of clustering algorithms have been introduced. These techniques can be divided into several categories [1].

1. Exclusive clustering and overlapping

Similar to the k-means clustering method, each piece of data belongs to only one cluster in the exclusive method after clustering. When clustering is done using the overlap approach, one data may belong to many clusters with different values, however, each data after clustering shows a degree of belonging to each cluster. Fuzzy clustering is a clear example of this approach.

2. Two types of clustering are hierarchical and flat

Based on their totality, the final clusters are given a hierarchical structure in the hierarchical technique. Each final cluster created using the planar approach has some level of generality, such as k-means.

Considering that hierarchical clustering methods produce more and more accurate information, they are recommended for detailed data analysis, but since they have high computational complexity, for large data sets, the method flat clustering is suggested.

1.9.2 Hierarchical clustering

As mentioned earlier, the final clusters in the hierarchical clustering approach are given a hierarchical structure depending on their generality, usually in the form of a tree. Dendogram is the name of this hierarchical tree. Each leaf node is considered as an individual data point, while the root node of the tree diagram represents the complete data set. Consequently, the height of a tree diagram often represents the distance between a data point and a cluster, and the central nodes of the range of objects that are close to each other. Cutting the tree diagram at different levels will reveal the final clustering findings. In particular, this diagram provides an insightful description of the clustering structure of the data when there are real hierarchical links in the data. This clustering method is divided into two categories based on the hierarchical structure produced by them [4] top-down or split-and-down rise or thicken.

1.9.3 Divisive hierarchical clustering: This approach divides data that are less comparable to each other into different clusters after treating them all as a single cluster through an iterative process. This process continues until reaching clusters with one member.

1.9.4 Condensing hierarchical clustering: Each piece of data is initially treated as a single cluster and then, through an iterative process, the most similar clusters are joined to form one or a certain number of clusters. The terms single join, full join, and average join can be used to describe several types of dense hierarchical clustering techniques. A key distinction between both approaches relates to how similarity between clusters is determined.

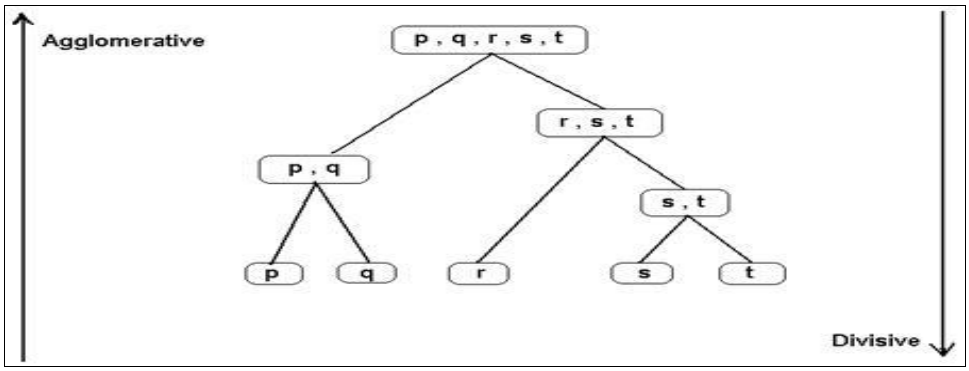


Fig 3.2: The difference between dividing and condensing methods

1.10 k-means algorithm

Despite its ease of use, several clustering techniques, including fuzzy clustering, consider this strategy to be a basic strategy [18].

This approach is flat and exclusive [22], this method has been expressed in several different ways, but they all share an iterative process that seeks to estimate the following for a predetermined number of clusters:

- Finding the average points belonging to each cluster and using those points as cluster centers.
- Each data sample assigned to a cluster must be the closest to the center of the cluster.

In a simple version of this method, points are initially randomly selected according to the required number of clusters. Then the data is attributed to one of these clusters according to the degree of proximity (similarity) and new clusters are obtained in this way. By repeating the same procedure, new centers can be calculated for each repetition by averaging the data and assigning the data to the new clusters. This process continues until there is no change in the data. More precisely:

- **First step:** K sites are first randomly selected to serve as cluster centers. Since different centers produce varied results, these centers should be selected carefully. Therefore, it is recommended that the centers are as far apart as possible.
- **Second step:** each data sample is assigned to a cluster whose center is closest to that data sample.
- **The third step:** this step is to create a new point as the center of each cluster, which is the average score of the points that belong to each cluster after assigning all the data to one of the clusters. Repeat steps 2 and 3 as necessary until the cluster centers remain unchanged. The application of the clustering algorithm for a data set with two data groups is shown in the image below. An asterisk and a circle are used to indicate different groups of data (2-4). A point is selected as the center of the first stage of clusters and is displayed in red (2-5). The second step involves placing each data sample into one of these two clusters and calculating a new central coefficient for each central cluster, as shown in Sections (2-6). As far as there is no more change, this process is done.

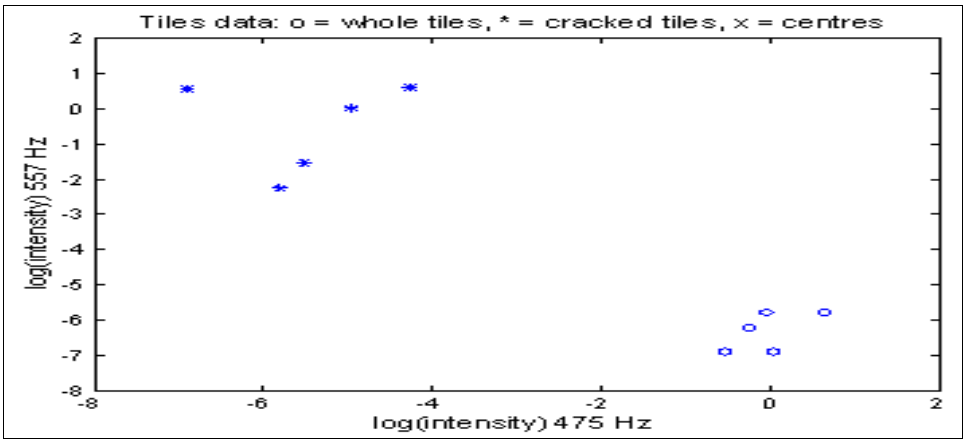


Fig 5.2: The first stage of clustering

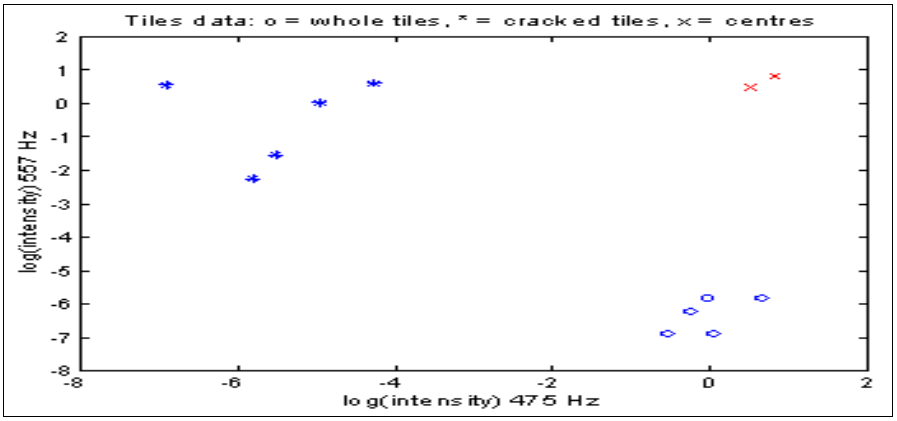


Fig 6.2: The second stage of clustering

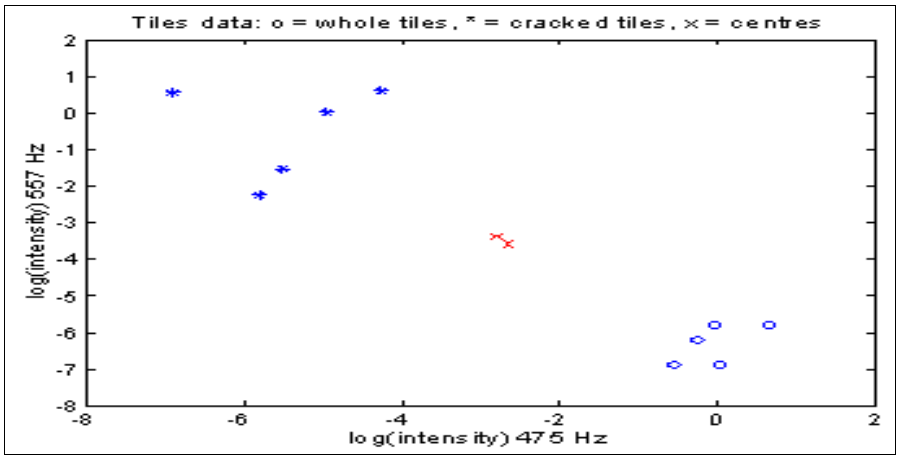


Fig 7.2: The third stage of clustering

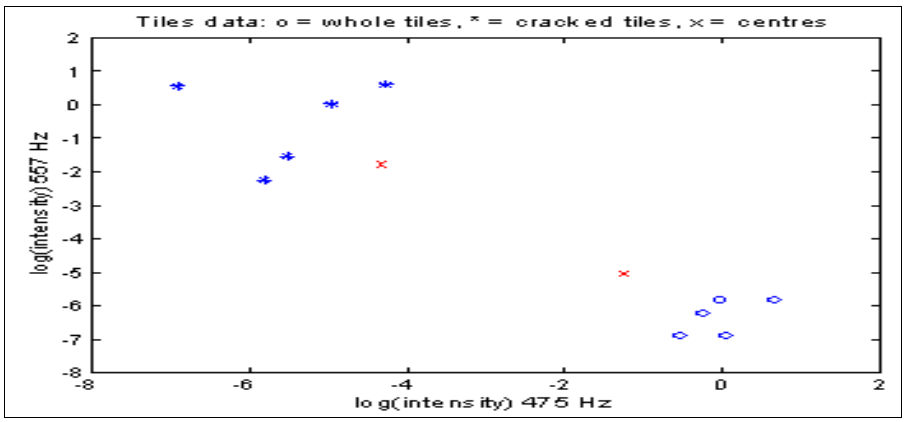


Fig 8.2: The fourth stage of clustering

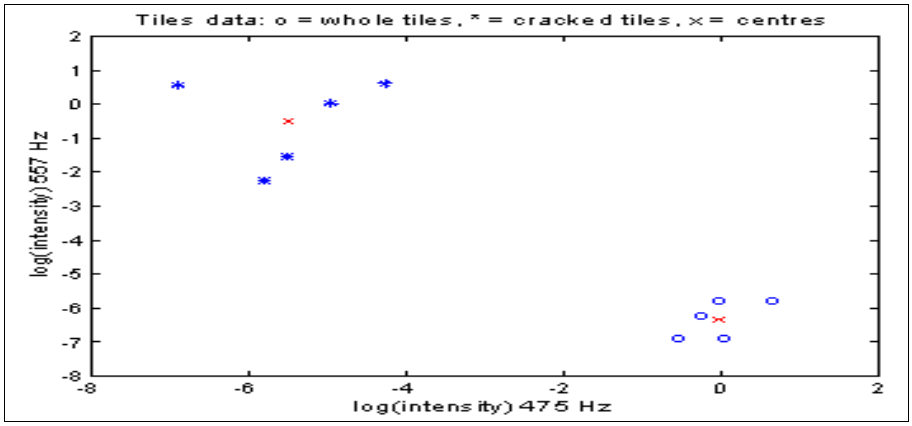


Fig 9.2: The fifth stage of clustering

The summary of the advantages and disadvantages of the mentioned algorithms is as described in the following table:

Article	Disadvantages	Advantages
Using the K-Means Algorithm for Clustering	The need to determine the number of clusters at the beginning, being sensitive to noisy and outlier data, the dependence of the final results on the quantification of the initial centers and the number of clusters.	Simplicity and ease of implementation, high speed and suitable for large data sets
K-harmonic means using the tabu-search method to cluster data	Tendency to converge to local optima	It does not have the center initialization problem
Genetically Improved PSO Algorithm for Efficient Data Clustering	Slow algorithm execution speed	It does not have the center initialization problem
K-harmonic means and particle swarm optimization are used to create an effective hybrid data clustering technique.	Trapped in the local optimum	Convergence speed higher than PSO algorithm
Using the Artificial Bee Colony (ABC) Algorithm, a Novel Clustering Approach		It has a lower error rate, it is highly suitable for grouping data with high dimensions and mass of several clusters

1.11 The working method of the proposed hybrid algorithm: As you can see in Figure (3-1), in the PSO algorithm, all the birds are flying towards the bird that has the least distance from the garden. In the middle, it is possible that one of the birds finds a better position than the best bird while following the best bird, which means that its distance from the best bird with food decreases. In this case, the birds follow the new bird. In fact, the best bird is the bird that has the least distance to the food. In Figure 2-4, among the birds following the best bird, one of the birds finds a smaller distance to the new food, which is larger than the previous food and closer to the birds, so this bird has a better position than the previous bird. As a result, it causes birds to flock to it.

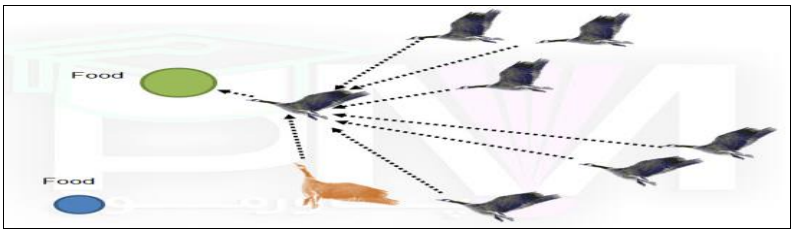


Fig 3.1: Flight towards the bird that has the shortest distance to the food

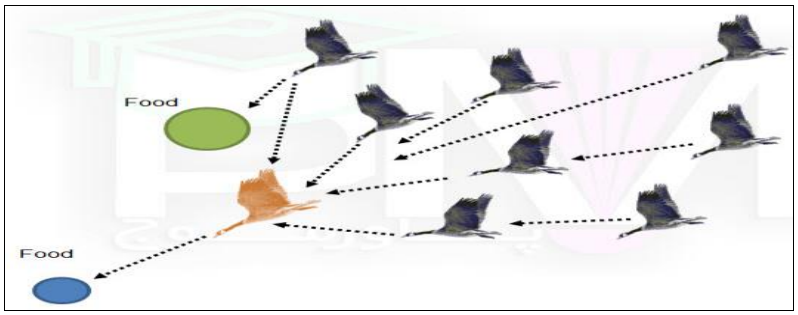


Fig 2.3: Flight to the best bird

As mentioned before, the particles in the PSO algorithm are constantly looking for a particle with the best position, and in the ACO algorithm, as in Figure 3-4, the particles easily choose their optimal path based on the intensity of the pheromone. We have given PSO properties as a pheromone so that with this change, the movement towards the desired particle can take place according to the path with higher pheromone intensity, and considering that the path with higher pheromone intensity is necessarily a more optimal path, therefore the movement of PSO particles will be optimized towards the new particle.

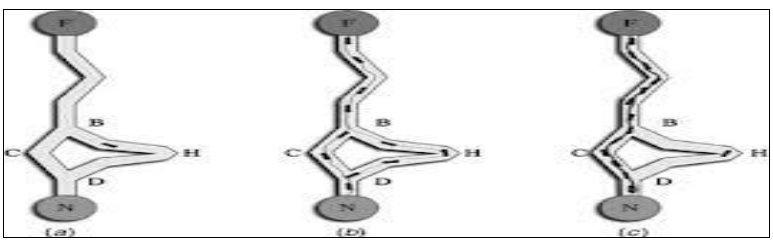


Fig 3.3: The movement of ants on the path with more furomen

1.11.1 Bee mating algorithm

Honey bee mating behavior is used as an extensive optimization strategy that is based on insect behavior and is really inspired by the actual mating behavior of bees. In addition to the queen, a hive contains tens of thousands of worker bees and hundreds of male bees. The main responsibilities of the queen are egg production and reproduction. The hive father is a male bee. The young are raised by worker bees who sometimes lay eggs. Male bees are produced by infertile eggs, while queen and worker bees are produced by live eggs.

During mating, the queen leaves the hive at a certain speed. Male bees chase the queen, and those that reach her succeed in mating with her, but die after mating. The queen flies until the capacity of the sperm chamber is sufficient. After it reaches the required size, it returns to the hive if it still has energy. The following formula simulates the probability that each male bee will be able to reproduce:

$$Prob(Q,D) = e^{-\frac{\Delta(f)}{s(t)}}$$

The possibility of adding the sperm of the male bee to the volume of the queen's sperm chamber or the probability of a successful mating. The absolute value is the difference between the objective function of the male bee and the objective function of the queen. The queen moves at the current speed. This indicates that the speed of the queen for mating at the beginning of the flight is high, and if the fitness performance of the male bee is also suitable and close to the value of the fitness performance of the queen, the probability of successful mating is high. The vitality and speed of the queen gradually decreases. The following formula illustrates this problem.

$$S(t+1) = \alpha \times s(t)$$
$$E(t+1) = E(t) - \gamma$$

1.12 Forbidden search pattern algorithm

Glover's presentation of this algorithm completed it in its current form, which may rival other methods in use in terms of versatility. In this approach, a flexible memory and (conscious) search for answers are needed to answer the problem intelligently. In addition, it offers an opportunity for efficient and inexpensive deployment in the solution space. According to the data collection, local optima are selected. The forbidden search pattern method is notable because it is distinctive and uses memoryless designs that rely mainly on quasi-random operations and perform some sort of sampling. The well-known genetic algorithm and steel plating processes, which are somehow derived from the physical and biological sciences, are examples of memoryless methods. The basic concept really comes from the fact that poor decision making may lead to a better solution than no decision at all. A wise stochastic decision of a poor decision can serve as a more useful hint on how to improve the outcome in a memory-based system. Movement in the solution space in this technique is not limited to the best solutions, but to the discovered neighborhoods to avoid entering the local optimal solutions and finding the global optimal solution. In terms of minimizing the objective function, the best option is selected, then this answer - which has better conditions than other answers among the neighbors of the starting point - is considered as the starting point of the next phase, and this process is repeated. It should be noted that the selected point may not always be superior to the start of the stage. The selected move is also added to the list of forbidden moves to avoid repeated responses and algorithmic closed loop. The result of this method is an almost ideal solution. It should be noted that any illegal action is allowed again after many iterations to prevent the algorithm from reaching the relative optimal position. According to the nature of the problem, the search space changes. By repeating these two processes, the algorithm finally converges to the absolute optimal solution [3].

1.13 Steel plating algorithm

Crick Patrick developed the cooling simulation method in 1983 by simulating between reducing the objective function of a problem and cooling the item until it reaches the fundamental energy state. The term "annealing" literally translates to "heating an object," but it refers to a physical method that involves heating an object to its melting point and then gradually cooling it to reduce its energy. Simulated also refers to simulating and recreating the behavior of a process according to a number of known or conjectured premises.

First, the objective function is calculated at a randomly selected point from the search space. It should be noted that the first solution, which plays a key role and is different depending on the type of issue, is one of the basic pillars of this method. The method of community development is another critical factor that will be very useful in determining the best course of action. After choosing these two parameters, the starting temperature is given to the system, which corresponds to the kinetic energy [19]. The value of the initial temperature is chosen arbitrarily, but depends on how the function behaves at the starting point. For example, if the function has little variation, a lower temperature is chosen to be less mobile, and if the function has a lot of variation (steep slope), a higher temperature is chosen to be more like the steady state. The minimum neighborhood can be left and moved. It would be better if the following move involves selecting a point from the nearby area. The problem determines how to choose using this strategy [17]. The choice is made to move to a new point: if the answer is better, the move will continue. If not, it moves to a new point. This probability is chosen according to the temperature of the system, the change in the size of the objective function and the distance between the origin and destination points. Two alternative functions are often used for the aforementioned probability and they are as follows:

$$P = 1/(1+\exp(\frac{-|F_2 - F_1|}{KT})) \quad (4.3)$$

1.14 Examining the results of the proposed algorithm

In this section, the results obtained from the proposed algorithm on well-known data will be examined. Then, the final answer of previous algorithms is compared with the algorithm discussed in this research. In general, there are different data sets to perform the clustering process, each of which has its own characteristics. Therefore, to achieve the existing goals, Iris, Wine, Glass datasets are selected. In the following, this set of data will be introduced and the obtained results will be displayed separately.

1.14.1 Introducing the data used
As mentioned, 3 famous data sets will be used to evaluate the performance of clustering systems; which will be explained briefly. It is worth mentioning that this data set is used in most clustering and classification problems.

1.14.2 Iris dataset
Fisher presented three examples of lily flowers as part of this dataset in 1936 to illustrate the use of linear separable approaches. Consequently, the Fisher-Lily dataset is another name for it. On the other hand, it is sometimes referred to as Anderson's iris data set because Edgar Anderson also collected this set as a result of the high geographical diversity of the Gaspé Peninsula.

Fig 4.1: Samples of iris flowers from Iris dataset

1.15 Wine dataset
This collection was taken from the MCI laboratory; It is obtained from the chemical reactions of wine. This dataset is the result of collecting 3 wine samples in different places in Italy. In this set, the parameters are (N=178, d=13, k=3). Out of these 178 cases, 106 cases were selected for training, 36 cases for validation and the remaining 36 cases for testing. It should be noted that all 13 features are continuous and none of them have a zero value. Also, the data distribution in each cluster is 59, 71, 48. 13 characteristics of this data set are as follows: alcohol, malic acid, ash, alkalinity of ash, magnesium, phonic acid, flavonoid, non-flavonoid phenol, peranisocyanins, color intensity, shape, OD280/OD315 of diluted wine and praline [15].

1.16 Glass dataset
This dataset, collected by a German research center, contains information on glass types. It contains a total of 214 samples, which are divided into 6 classes. Classes 1 to 6 have 70, 17, 76, 13 and 29 samples, respectively. Each class has 10 attributes with continuous values.

Simulation results of the proposed hybrid algorithm
The results of the implementation of the proposed algorithm by MATLAB software are shown by comparing the accuracy of particle clustering in the new algorithm compared to other algorithms. Table 1-4, which is related to the samples that are not included in their respective cluster.

Table 4.1: Comparison of the proposed algorithm with other algorithms

DataSet	تعداد کلاس	K -means	K-means-aco	K-means-pso	K-maens-pso-aco
Iris	3	0.106667	0.106667	0.333333	0.106667
Wine	3	0.297753	0.297753	0.297753	0.297753
Glass	7	0.411215	0.448598	0.448598	0.448598

Proposal for future works
One of the most important clustering algorithms is the K-means algorithm, which has many applications in various fields. Despite the advantages such as simplicity of implementation and high speed, this algorithm has disadvantages such as being sensitive to noisy and outlier data and the dependence of the final result on the initialization [16]. These challenges have been overcome in recent years by using k-means in combination with other algorithms. Harmonic, PSO, PSOKHM and ACO algorithms can be mentioned among the most important combined algorithms, each of which They have been able to solve the disadvantages of the k-means algorithm to some extent, but an optimal method that can comprehensively solve all the disadvantages of the k-means algorithm by combining several algorithms has not yet been created. Considering the efforts made, the combination of fuzzy algorithms and ACO can be suggested as a future work [24].

Conclusion
Today, with the advancement of new technologies and its direct impact on human life, especially regarding the growth and use of the Internet of Things, we naturally need a smarter view regarding the optimal use of these opportunities. The science of data mining and its related topics are part of the sciences that will play a very colorful role in the future of the Internet of Things. The existing data mining algorithms each have specific advantages and disadvantages, for example, the Kmeans clustering algorithm, which is one of the most widely used algorithms. It is used in data mining, despite its easy implementation, it faces problems such as dependence on initial conditions, premature convergence, and getting stuck in local optima. Therefore, by using the ACO algorithm, which works based on the pheromone secreted by ants, and the PSO algorithm, which works based on the relationship of the mass of particles based on collective intelligence and wisdom, it was

tried to somehow increase the accuracy of clustering by using their characteristics. Raised a limit. By using the ACO foremans in the PSO algorithm, it was determined that the resulting clustering algorithm works with better accuracy on the data in the classes related to the cluster itself. Simulations have been run on several data sets, and most of the time, the new hybrid technique provides the answer with the highest level of accuracy. The problem of constant convergence to the local optimal response is also eliminated in this hybrid technique to an acceptable extent, besides the limitation of dependence on the initial conditions is significantly neutralized.

Acknowledge
I would like to express my sincere gratitude to Dr. Hussein Majeed Rustom for his invaluable guidance throughout this study. His insightful feedback and expertise were instrumental in shaping this research. I also extend my appreciation to the staff in Al-Sadiq University -Thi-Qar center whom assist me to complete this study. Finally, I appreciate the encouragement from my colleagues and family, whose support kept me motivated throughout the research process

Reference

1. Jain AK. Data clustering: a review. ACM Computing Surveys. 1999;29.
2. Yıldız AR. A novel particle swarm optimization approach for product design and manufacturing. Int J Adv Manuf Technol. 2009;44.
3. Singh A. An intelligent hybrid approach for hepatitis disease diagnosis: combining enhanced k-means clustering and improved ensemble learning. Expert Syst. 2021;44.
4. Xiao B. SMK-means: an improved mini batch k-means algorithm based on MapReduce with big data. Comput Mater Continua. 2018;222.
5. Tsai C-W. Data mining for internet of things: a survey. IEEE Commun Surv Tutorials. 2013;77.
6. Zhu C. Improved logistic regression model for diabetes prediction by integrating PCA and K-means techniques. Informatics Med Unlocked. 2019;66.
7. Karaboga D. A novel clustering approach: Artificial Bee Colony (ABC) algorithm. Appl Soft Comput. 2011;65.
8. Yang F. An efficient hybrid data clustering method based on K-harmonic means and Particle Swarm Optimization. Expert Syst Appl. 2009;36.
9. Tsekouras GE. A new approach for measuring the validity of the fuzzy c-means algorithm. Adv Eng Softw. 2004;44.
10. Papamichail GP. The k-means range algorithm for personalized data clustering in e-commerce. Eur J Oper Res. 2007;140.
11. Zare H. Determination of customer satisfaction using improved K-means algorithm. Soft Comput. 2020;33.
12. Wang J. Stochastic optimal competitive Hopfield network for partitional clustering. Expert Syst Appl. 2009;20.
13. Mittal K. Performance study of K-nearest neighbor classifier and K-means clustering for predicting the diagnostic accuracy. Int J Inf Technol. 2019;535.
14. Mittal K. Performance study of K-nearest neighbor classifier and K-means clustering for predicting the diagnostic accuracy. Int J Inf Technol. 2019;33.
15. Tian K. Segmentation of tomato leaf images based on adaptive clustering number of K-means algorithm. Comput Electron Agric. 2019;22.
16. Fernando Raguro MC. Extraction of LMS student engagement and behavioral patterns in online education using decision tree and K-means algorithm. 4th Asia Pacific Information Technology Conference. 2022;138.
17. Heath MT. Scientific computing: an introductory survey, revised second edition. 2018;33.
18. San OM. An alternative extension of the k-means algorithm for clustering categorical data. Int J Appl Math Comput Sci. 2004;33.
19. Cortez P. Modeling wine preferences by data mining from physicochemical properties. Decis Support Syst. 2009;33.
20. Abdel-Kader RF. Genetically improved PSO algorithm for efficient data clustering. 2nd Int Conf Mach Learn Comput. 2010;71.
21. Sagar S. Trust computational heuristic for social Internet of Things: a machine learning-based approach. ICC 2020-2020 IEEE Int Conf Commun (ICC). 2020;6.
22. Yu S-S. Two improved k-means algorithms. Appl Soft Comput. 2018;33.
23. Liang XW. LR-SMOTE an improved unbalanced data set oversampling based on K-means and SVM. Knowl Based Syst. 2020;1054.
24. Wang X-Y. Simulated annealing fuzzy clustering in cancer diagnosis. Informatica. 2005;44.
25. Güngör Z. K-harmonic means data clustering with tabu-search method. Appl Math Model. 2008;115.