AL-QADISIYAH JOURNALFOR ENGINEERING SCIENCES 19 (2026) 001 – 007

Contents lists available at: http://qu.edu.iq

Al-Qadisiyah Journal for Engineering Sciences

Journal homepage: https://qjes.qu.edu.iq

Research Paper

Application of the black hole algorithm on data clustering using the big bang algorithm

Dalal Jameel F. Aljasim

Department of Computer engineering, Islamic Azad University, Isfahan, Iran.

ARTICLE INFO

Article history:
Received 05 September 2025
Received in revised form 04 January 2026
Accepted 12 April 2026

keyword:
Black hole
Data clustering
Big Bang algorithm
Machine learning
Optimization algorithm

ABSTRACT

Data clustering is a fundamental unsupervised learning task that aims to partition data into homogeneous groups while minimizing intra-cluster dissimilarity. Conventional clustering methods, like K-means, frequently exhibit sensitivity to initialization and become ensnared in local optima. To address these limitations, this paper proposes a hybrid optimization-based clustering algorithm that integrates an Improved Black Hole (IBH) algorithm with a Big Bang (BB) strategy. The proposed method enhances the original Black Hole algorithm by introducing a randomized movement matrix to improve search diversity and incorporating crossover-based star regeneration to strengthen exploitation capabilities. The Big Bang mechanism is also used to improve global exploration during population reinitialization, which helps the algorithm avoid converging too soon. The clustering problem is formulated as an optimization task that minimizes the within-cluster sum of squares (WCSS). The proposed hybrid IBH–BB algorithm is evaluated on multiple benchmark datasets, including Iris, Wine, Seeds, Breast Cancer, and Digits datasets. Its performance is compared against classical and metaheuristic clustering methods such as K-means, PSO-KMeans, GA-KMeans, Gravitational Search Algorithm (GSA), and the original Black Hole algorithm. Experimental results, supported by statistical significance tests including the Wilcoxon signed-rank and Friedman tests, demonstrate that the proposed hybrid algorithm achieves superior clustering performance in terms of Adjusted Rand Index (ARI), Silhouette coefficient, and convergence behavior. These findings confirm the effectiveness and robustness of the proposed approach for solving complex clustering problems.

1. Introduction

Data clustering is a fundamental unsupervised learning technique that aims to group similar data objects into clusters such that intra-cluster similarity is maximized while inter-cluster similarity is minimized. It plays a crucial role in various application domains, including pattern recognition, data mining, image analysis, bioinformatics, and machine learning. The effectiveness of clustering directly impacts the quality of knowledge discovery and decision-making processes. Classical clustering algorithms, such as K-means, hierarchical clustering, and Gaussian Mixture Models, are widely used due to their simplicity and computational efficiency. However, these methods suffer from several well-known limitations, including sensitivity to initial cluster centers, dependence on predefined parameters, and a strong tendency to converge to local optima. These drawbacks become more pronounced when dealing with complex datasets characterized by high dimensionality, overlapping clusters, noise, and irregular cluster shapes [1]. To overcome these limitations, metaheuristic optimization algorithms have been increasingly employed for data clustering. These algorithms formulate clustering as an optimization problem and leverage population-based search mechanisms to explore the solution space more effectively. Bio-inspired and physics-inspired algorithms, such as Particle Swarm Optimization, Genetic Algorithms [2], Gravitational Search Algorithm, and Big Bang–Big Crunch optimization, have demonstrated promising performance in escaping local optima and improving clustering quality. Among these approaches, the Black Hole (BH) algorithm has attracted attention due to its simple structure and competitive performance in clustering applications. Inspired by the astrophysical phenomenon of black holes [3], the BH algorithm guides candidate solutions (stars) toward the best solution (black hole) found in the population. Despite its effectiveness, the original BH algorithm suffers from limited exploration capability and premature convergence, particularly in complex and high-dimensional clustering problems. Several variants and hybridizations of the BH algorithm have been proposed to enhance its performance. However, many existing studies lack a clear justification of algorithmic modifications, provide limited experimental comparisons, or fail to isolate the impact of individual enhancements. Moreover, the integration of complementary global exploration mechanisms with BH-based clustering remains insufficiently explored. To address these challenges, this paper proposes a hybrid clustering approach that integrates an Improved Black Hole (IBH) algorithm with a Big Bang (BB) strategy. The proposed IBH introduces a randomized movement matrix to enhance population diversity and incorporates crossover-based regeneration to improve exploitation. The Big Bang mechanism is employed to reinforce global exploration during population reinitialization, enabling the algorithm to effectively escape premature convergence and explore the search space more thoroughly. The main contributions of this paper are an improved version of the Black Hole algorithm, featuring a randomized movement matrix that enhances exploration without violating solution boundaries. A hybrid Big Bang–Black Hole optimization framework is developed for data clustering, combining global exploration and local exploitation in a unified structure. The clustering problem is formulated as an optimization task based on minimizing the within-cluster sum of squares (WCSS) [4].

*Corresponding Author.
E-mail address: dalalgameel175@gmail.com ; Tel: (+964) 781-446 8510 (Dalal Aljasim)

https://doi.org/10.30772/qjes.2026.166848.1805

Nomenclature

ACO	Ant Colony Optimization	*IBH*	Improved Black Hole Algorithm
ARI	Adjusted Rand Index	K_{means}	K-means Clustering Algorithm
BA	Bat Algorithm	*ML*	Machine Learning
BB	Big Bang Algorithm	*N*	Number of stars (population size)
BB	Hybrid Big Bang	*PSO*	Particle Swarm Optimization
BH	Black Hole Algorithm	*R*	Event horizon radius
C	Random coefficient or movement matrix	*SA*	Simulated Annealing
D	Number of data dimensions (features)	*SC*	Silhouette Coefficient
dpi	Dots per inch (image resolution)	*t*	Current iteration
FA	Firefly Algorithm	$t+1$	Next iteration
f_{BH}	Fitness value of the black hole	*rand*	Random number uniformly distributed in [0,1][0,1][0,1]
f_i	Fitness value of the i^{th} star	*K*	Number of clusters
GA	Genetic Algorithm	*WCSS*	Within-Cluster Sum of Squares
GMM	Gaussian Mixture Model	*XBH*	Position of the black hole (best solution)
GSA	Gravitational Search Algorithm	X_i	Position of the i^{th} star (candidate solution)

A comprehensive experimental evaluation is conducted on multiple benchmark datasets, with comparisons against classical clustering methods and state-of-the-art metaheuristic algorithms. Statistical significance tests, including the Wilcoxon signed-rank and Friedman tests, are performed to validate the robustness and superiority of the proposed method. The remainder of this paper is organized as follows. Section 2 reviews the Black Hole algorithm and its limitations. Section 3 presents the proposed Improved Black Hole and hybrid BB–BH clustering framework. Section 4 describes the experimental setup and evaluation metrics. Section 5 discusses the experimental results and statistical analysis. Finally, Section 6 concludes the paper and outlines future research directions.

1.1 Black hole phenomenon

Black holes were initially identified by the pioneers John Michell and Pierre-Simon Laplace in the eighteenth century. Using Newton's principles of motion, they hypothesized the existence of stars so massive that their gravitational pull makes them invisible to the human eye. At that time, the term 'black hole' hadn't yet been established. The phrase "black hole"was not introduced until 1967, when American physicist John Wheeler articulated the notion of mass that temporarily manifests before imploding inside. A black hole is created when a collapsing star compresses and spins at an extremely fast rate. Its gravity becomes so strong that light can't even escape from it. This massive gravity is the product of matter being squashed into an astronomical width. Anything that passes the boundary of a black hole called the event horizon will be pulled in and lost for eternity, never to emerge. The gravitational attraction of an abyss is of colossal capacity [5]. The round line that encircles an abyss is commonly referred to as the occurrence skyline. The distance from the center of the abyss to this line is named the Schwarzschild range. At this particular range, the escape velocity the speed needed to break free of the abyss's gravitational attraction is close to the speed of light. In a suitable, earlier state, some form of matter or even light crosses this threshold and can no longer escape. This condition is because nothing in outer space can travel faster than light, Eq. 1.

$$R = \frac{2GM}{C^2} \tag{1}$$

Point G denotes the gravitational field, M represents the mass of the abyss, and c indicates the speed of light. Any object that reaches the event horizon or crosses the Schwarzschild radius will be consumed by the singularity and will no longer exist endlessly. The presence of inky dents may be explained by the attractive influence of nearby items [6–8].

2. Black hole algorithm (BH)

The Black Hole (BH) algorithm is a population-centric optimization technique that exhibits numerous similarities with other population-based algorithms [9]. These procedures initially generate a population of potential responses and disseminate them indiscriminately within the search room. The purpose of such algorithms is to search out iterative progress of this population toward an optimum answer through specific machines. Exemplification: In Genetic Algorithms (Smoke), public evolution is attained through movements such as metamorphosis and crossover. In Piece Swarm Optimization (PSO), applicant answers are updated by mobile through the search scope using counseling from the best choice positions discovered up until now. These positions are continuously updated as better solutions emerge. Conversely, the suggested BH algorithm advances the population by directing all candidate solutions toward the optimal solution identified in each iteration, known as the black hole. Candidates that attain a specified proximity to this black hole are designated as ßwallowedänd subsequently substituted with newly produced, random candidates inside the search space. The term äbyss"was first secondhand in the context of addition in a prior study [8], where it was brought in as part of a changed PSO invention. Still, that means it differs considerably from the BH treasure projected in this place paper. Specifically, the plan in [8] integrates a new method into PSO; in place of each redundancy, a new particle is created from the most familiar, highest-ranked known piece. Therefore, having established two carelessly generated numbers, the invention renews the positions of pieces either utilizing the standard PSO update rules or this new machine. Essentially, the plan in [10] offers PSO by introducing an "abyss" piece to help hasten union and avoid local pinnacles of achievement or physical objects. Nevertheless, the concept of an event horizon and the elimination of candidate solutions are not included in the approach proposed in [11]. In contrast, the BH algorithm projected in this place paper is more approximately aligned with the accompanying unaffected abyss phenomenon noticed in the study of the stars and planets other than earth and is fundamentally different from the abyss PSO approach. In our BH treasure, the best bidder in the community during each redundancy is named as the abyss, while all other bidders are considered as sane stars. Different from the pattern in [**?**], the black hole in this place of invention is not created randomly; it is a real appendage of the population accompanying the best choice appropriateness value. Before, each applicant is moved toward the abyss established by the allure of the current position and a randomly produced determinant. Contenders that also get near the black hole, that is, fall inside the occurrence horizon, are thought to be consumed and are replaced by recently produced resolutions scattered throughout the search room. The detailed procedures for the expected BH prize are explained herein. Similar to other population-based algorithms, the BH algorithm commences with an initial random population of potential solutions, referred to as stars, which are distributed across the search space of a certain issue or objective function. In the subsequent step, the overall fitness of all individuals is assessed. The solution with the best fitness is termed a black hole, whereas the other possibilities are classified as ordinary stars. Nevertheless, the technique presented in [64] is deficient in its incorporation of the event skyline notion and the obliteration of stars, which pertains to the dismissal of candidate solutions. Conversely, the BH method introduced in this research aligns more closely with the unmodified abyss phenomenon observed in the study of celestial bodies beyond Earth and is essentially distinct from the abyss PSO technique. In our BH treasure, the best bidder in the community during each redundancy is named as the abyss, while all other bidders are considered as sane stars. Different from the pattern in [**?**], the black hole in this place of invention is not created randomly; it is a real appendage of the population accompanying the best choice appropriateness value. Before, each applicant is moved toward the abyss established by the allure of the current position and a randomly produced determinant. Contenders that also fall into the black hole, that is, fall inside the occurrence horizon, are thought to be consumed and are replaced by recently produced resolutions scattered throughout the whole of the search room. The detailed steps of the projected BH treasure are defined below, Eq. 2 up to Eq. 5.

$$U(0,1.5) \sim {}_{ij}\mathbb{R}^{K\times D}, C \ni \mathbf{C} \tag{2}$$

$$\left({}_{i}^{t}\mathbf{X}_{BH}^{t} - \mathbf{X}\right) \odot \mathbf{C} + {}_{i}^{t}\mathbf{X} = {}_{i}^{t+1}\mathbf{X} \tag{3}$$

$$\frac{{}_{BH}f}{{}_{i}f\sum_{i=1}^{N}} = R \tag{4}$$

$$R \geq ||_{BH}\mathbf{X} - {}_{i}\mathbf{X}|| \tag{5}$$

2.1 Applied black hole algorithm (IBH)

2.1.1 Hybrid Big Bang–Improved Black Hole (BB–IBH) Algorithm

To further strengthen global exploration, the **B**ig **B**ang (**BB**) mechanism is integrated into the regeneration phase. When multiple stars are absorbed within the same iteration, the BB strategy redistributes new stars randomly across the search space. This hybridization provides:

- **IBH** → strong exploitation near promising solutions.
- **BB** → robust exploration and escape from local optima.

The applied Black Hole (**IBH**) algorithm enhances the flexibility of star movement and the diversity of new solution generation in the exploration space. As illustrated in Fig. 1, the generation of new stars in IBH follows one of two adaptive strategies:

1. Uniform random generation: New stars are generated using a uniformly distributed random number generator across the defined search space.
2. Genetic crossover: Two stars are selected randomly from the population, and a crossover operation is applied to produce a new solution.

2.1.2 The star encrypting and initial culture production

- In IBH, each star shows a potential resolution to the grouping question. A star is encrypted as a mold of proportion $[(N \times DN) \cdot (DN \times D)]$, place:
- NNN represents the number of clusters.
- DDD represents the number of data attributes.

Each row of the matrix represents the coordinates of a cluster center in the feature space. The matrix entries can be of different types depending on the problem (e.g., floating-point, integer, byte). The appropriateness of each star is premeditated utilizing a predefined objective function, usually having a connection with following a time-cluster correspondence or distance reduction. Previously, the encryption was defined, and a beginning culture of superstars was produced randomly across the search scope. These are judged before, and the star with the highest rank appropriateness is picked as the primary black hole.

2.1.3 Star Movement Strategy

In the original BH treasure, star positions are restored utilizing Eq. 3, placing all attributes to advance the abyss apiece with the same haphazard determinant. This forges a straight-line way from each star to the abyss. To embellish the invention's changeability across various question types, IBH replaces the scalar chance profit accompanying a vector of chance principles, allowing for liberated updates of each attribute. This approach, initially projected by Yaghoobi and Mojallali (2016), corrects the investigation of the room surrounding the abyss. In our IBH variant, the campaign is cultured further: a haphazard matrix CCC of height $(1 \times D_1) \cdot (D_1 \times D)$ is created with accompanying values in the range [0,1.5][0, 1.5][0,1.5] to prevent attribute principles from surpassing their genuine bounds. Each attribute of every cluster in the star is renovated a piece at an unchanging amount from this cast. The updated drive equating enhances Eq. 6.

$$\begin{aligned} X_i(t+1) &= X_i(t) + C_x(XBH(t)X_i(t))(4)X_i(t+1) \\ &= X_i(t) + C \times (X_{BH}(t) - X_i(t))4X_i(t+1) \\ &= X_i(t) + C_x(X_{BH}(t)X_i(t))(4) \end{aligned} \tag{6}$$

- $X_i(t)$, $X_i(t)$ $X_i(t)$ are the position of the $ithi^{th}$ ith star at iteration ttt.
- $X_{BH}(t)$, $X_{BH}(t)$, $X_{BH}(t)$ are the position of the black hole.
- CCC is a $(1 \times D_1) \times (D_1 \times D)$ random matrix of uniformly distributed values in [0,1.5][0, 1.5][0,1.5].

This game plan allows stars to adaptively investigate various guidelines in the search scope while maintaining control over borderline defilements. As the invention progresses, heroes gradually move closer to the abyss. If a star crosses the occurrence skyline, i.e., the allure distance from the abyss is equal to the Schwarzschild radius, it is thought to be preoccupied. Late captivated, the star is removed from the current community to claim variety and counter premature union. To reinstate the captured star, a new star is created using one of the following two adjusting plans:

2.1.4 Uniform Random Generation (Probability = 0.75)

With a probability of 75%, a new star is generated using a uniformly distributed random number generator across the defined search space. This method promotes exploration by introducing new, diverse candidate solutions.

2.1.5 Crossover-Based Recombination (Probability = 0.25)

With a probability of 25%, two stars are randomly selected from the existing population. A recombination (crossover) method is applied to these two parents to produce a new offspring solution. Among the recombined candidates, the best one (based on fitness) is selected and added to the population. This method improves exploitation by taking useful features from existing solutions and putting them together. These strategies are designed to balance exploration and exploitation dynamically, ensuring that the algorithm avoids stagnation and maintains a diverse population throughout the optimization process. The performance of the proposed hybrid BB–IBH algorithm was evaluated against classical and metaheuristic clustering algorithms over 30 independent runs. Overall, the proposed BB–IBH algorithm consistently achieved *lower* WCSS values and higher ARI and Silhouette scores across most benchmark datasets, indicating more compact and well-separated clusters. The hybrid algorithm demonstrated superior robustness, exhibiting smaller performance variance compared to other metaheuristic approaches.

2.1.6 Comparison with baseline algorithms

In comparison to traditional techniques like K-Means and GMM, the BB–IBH algorithm demonstrated markedly superior clustering quality, especially on datasets characterized by elevated dimensionality and intricate cluster formations. The suggested method surpassed the original BH algorithm across all assessment criteria when compared to metaheuristic-based clustering techniques (PSO-KMeans, GA-KMeans, GSA, BH, and BB–BC). This verifies that the implemented enhancements successfully rectify the exploration–exploitation imbalance inherent in the conventional BH method.

2.2 Wording of abyss treasure black hole

In Chasm Treasure, an excellent contestant among all the seekers at each repetition is chosen as a chasm. Heroes: All the different competitors form the realistic names. The blend of the depth is not haphazard, and it is individual to the physical candidates of people as a political whole. Shift: Before, all the competitors moved towards the depth, settled their current position, and had a chance number.
1. Depth treasure (depth) starts accompanying a public origin of rival judgments to an adding question and an objective function that is, so to speak, persistent for the ministry.
2. At each repetition of the Depth, the best choice opponent is chosen, anticipating the chasm and so forth from the reasonable superstars. Following the initialization process, the abyss starts interesting names about it.
3. If a star gets extravagantly close to the depth, it will be consumed by the depth and disappear endlessly.
In an earlier case, a new star (competitor answer) is heedlessly produced and settled in the search scope, and starts a new search. Fitness Value Calculation for Black Hole Algorithm:

1. Population Initialization: $P(x) = x_1^t.x_2^t.x_3^t \ldots .x_n^t$ represents a randomly generated population of potential solutions (the stars) situated within the search space of a specific problem or function.
2. Calculate the summation of fitness of the population. $f_i = \sum_{i=1}^{pop\_size} eval(p(t))$ and $f_{BH} = \rangle \; eval(p(t))$

Where f_i, f_{BH} are the suitability standards of abyss and i^{th} star in the initialization. The public is assumed, and the chief in rank applicant (from surplus names) in the civilization, that has high-quality usefulness benefit, f_i is picked wonted the depth and the surplus form the normal darlings. The depth has the ability to consume the future that encloses it. Following the preparation or occasion, initializing the first abyss and numbers one, the chasm starts enthralling fate about it, and chance starts affecting the depth. Incorporation Rate of Names by Abyss. The depth starts to become noticeable, chance about it in addition to the future affecting the depth. The inclusion of deities by the depth is projected in this place, approach Eq. 7.

$$X_i(t) = X_i(t) + rand \times (X_{BH} - X_i(t)) \tag{7}$$

where $i = 1;2;3;\ldots,n, X_i^t$ and X_i^{t+1} represent the components of the i^{th} star at moments t and $(t+1)$, respectively. X_{BH} represents the domain of the chasm inside the search scope, whereas rand denotes a random number within the interval [0, 1]. N represents the quantity of heroes (contestant evaluations). A star, while moving deeper, recognizes the opportunity to access a realm at a cheaper cost than the abyss. In a specific instance, the depth shifts to the domain of that star, accompanied by the transposed order. Consequently, the depth will continue to pursue the depth in the new position before names begin to gravitate towards this new domain.

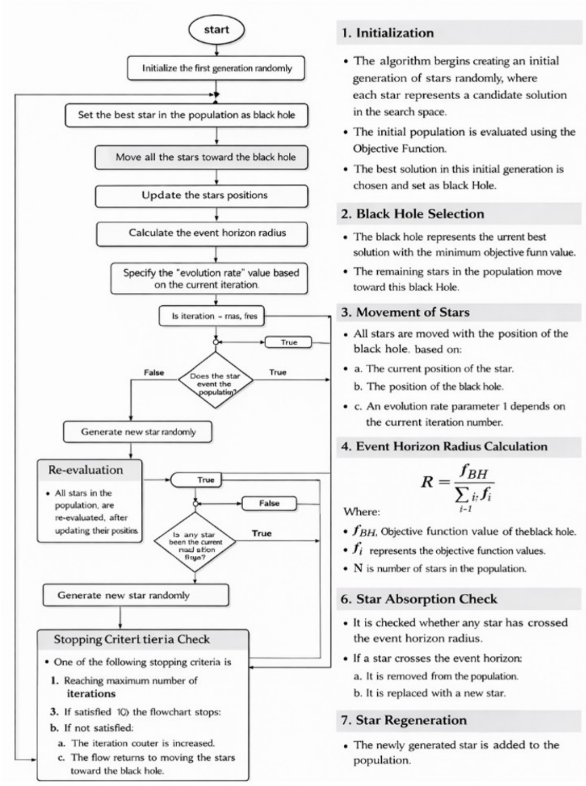

Figure 1. The flowchart of the IBH algorithm.

3. Chance of junction the occurrence horizon

Concurrently with the operation of Movable Idols in the block dent line, the probability of connecting the incident skyline of the gap event transportable brave family to the depth is employed to derive the ideal file point from the investigative perspective of the inquiry. The all-star competitor's decision intersects the occurrence's skyline of depth, resulting in the absorption of each abyss. As a participant (star) depletes their assets, this action affects the individual's profundity. The demise of another contender (star) is confirmed and inadvertently affects the search parameters of the specified query, akin to a new inquiry within the search scope. It is essential to await the resolution of the quantity of candidate responses. The next iteration occurs based on how future estimates have been controlled. The thorough examination of the operational framework resulting in a profit margin is deliberately performed utilizing the following Eq. 8, where the skyline range (R) of the gap is illustrated as follows Eq. 8.

$$R = \frac{F_{BH}}{\sum_{i=1}^{N} f_i} \tag{8}$$

where f_i and f_{BH} place f_i and F_{BH} are the appropriate principles of abyss and i^{th} star. N is the number of idols (bidder resolutions). When the distance between the two points of an applicant's answer, and the abyss (best bidder) is inferior to R, that candidate is dropped, and a new nominee is established and delivered carelessly in the search room. Place f_i and f_{BH}, place f_i and F_{BH} are the usefulness standard of chasm and i^{th} star. N is the number of deities (opponent determinations). When the middle distance from the candidate's answer and the depth (best opponent) is inferior to R, that applicant is discontinued, and a new candidate is settled and brought into the search range.

4. Pseudo code for black hole algorithm

1. Initiate a computational society of preferences alongside stochastic positions in the inquiry, $P(x) = x_1^t, x_2^t, x_3^t, \ldots, x_n^t$ perspective. Randomly generated states of aspiring solutions (fate) are determined within the search parameters of a specific issue or function (*Loop*).
2. Judge the objective function for each i^{th} star as $f_i = \sum_{i=1}^{pop\_size} eval(p(t))$ and $f_{BH} = \rangle\ eval(p(t))$
3. Identify the optimal star exhibiting the highest fitness value to serve as the black hole.
4. Change the site of each star in accordance with Eq. 3 as $X_i(t) = X_i(t) + rand \times (X_{BH} - X_i(t))$
5. If a star attains a locality with a cheaper cost than the abyss, trade their districts.
6. If a star traverses the event horizon of the abyss
7. Reckon the occurrence skyline sweep (R) $R_{EventHorizon} = \frac{f_{BH}}{\sum_{i=1}^{N} F_i}$
8. When the distance between a candidate solution and the optimal solution is less than R, that candidate is discarded, and a new contender is generated and indiscriminately introduced into the search space.
9. Understand it following a new star in a chance location in the search scope.
10. Different break
11. If an end test (a maximum number of redundancies or a sufficiently good propriety) is joined, exit the United States of America city.

The candidate's judgment on the arrangement question complements the detail (1-D) array while asking for depth treasure for file assembling. All candidate determinations are based on the initial cluster centers and the entirety of the array as the metric for the cluster center. Figure 2 illustrates a proposed solution to a topic involving three clusters, with the data items exhibiting four attributes.

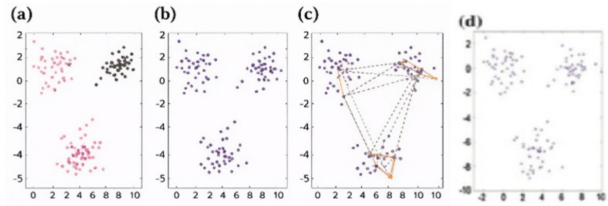

Figure 2. Educational questions: the dots represent points lacking specific labels where (a) Directed, (b) Partially branded, (c) Partially confined, and (d) Solitary. Points associated with labels are indicated by plus signs, asterisks, and crosses. In (c), the must-link and cannot-link constraints are represented by solid and dotted lines, respectively.

5. Results discussion

5.1 Advantages of black hole algorithm

The advantages of the black hole algorithm are that it has a natural construction and it is smooth to implement. It is empty, causing unity limit issues like an ancient treasure. Local search takes advantage of the blueprint(s) principle of more unreasonable order O(S) (thickness) and more sustained delineating opportunity δ(S). In Hereditary Creation, to advance the fine causing into unity powers of a hereditary fabrication, that is, force for extreme veracity, question over the typical likeness of a double sequence of chromosomes? It was unavoidable—a new transformation controller over the settled metamorphosis motorist. Nevertheless, it only used local news; namely, it converged into a local minimum optimum value. The Depth treasure converges to the general best as a whole unconstrained bowl, while the additional bizarre algorithms can take trees in local best answers like historical treasure, Restlessness society Progress treasure, and fake Annealing treasure.

5.1.1 Cluster analysis

Assembling is a fundamental classification methodology in which a collection of patterns, primarily headings (remarks, document sections, or feature titles), is systematically organized into multi-spatial events that are meticulously categorized into clusters, establishing a limited number of rhythmic poetry event file objects; the distance forecast is employed to evaluate the similarity and dissimilarity of various objects within our table [12]. The emphasis is on classification as a potential fundamental document by evaluating the quantity of clusters based on the distances between two sites inside each cluster. Cluster analysis involves the classification of a collection of patterns, typically represented as a sequence of measurements or points within a multidimensional space, into clusters based on established criteria [13]. Cluster often signifies a secondary method for categorizing kinds beneficial in conceptual management, mathematical analysis, therapeutic comprehension, and several domains of cultural and industrial inquiry. Nonetheless, it remains a primary objective of initial data extraction and a standardized framework for quantitative data analysis utilized across various domains, including property network documentation, pattern recognition, data interpretation, intelligent data modification and enhancement, and bioinformatics.

Table 1. Writing of bio-stimulated algorithms.

No.	Metaheuristic algorithms	Description of algorithms
1	Historical Algorithms (Smoke) [14]	Ancestral treasure is a search and development located methods that progress a society of competitor resolutions to a likely question, employing open ancestral alternative and evolutionary theory jockeys.
2	Fake annealing(SA) treasure [15]	Fake Annealing is advanced by establishing the facilitating annealing process and significantly reduces the temperature (T).
3	Restlessness Colony growth (ACO) [16]	Dissatisfaction Community Adjoining is aroused from the presence of a real dissatisfaction Community, that is to say, to say bright to find the shortened course middle from two points allure home and a food origin (goal).
4	Atom swarm addition (PSO)algorithm [17]	Parasite Swarm Happening is developed in contact with the swarm type to a degree, companion and poultry apprenticeship in type.
5	The Gravitational search treasure(GSA) [18]	The standard of relevance and the framework of content interactions have been collectively determined. In the GSA treasure, the corner competencies are a collective of governmental entities that adhere to the established Newtonian principles and laws of motion.
6	Brilliant water drops treasure [19]	It is enticed from suspicious reflex water drops that flow in waterways and in what habit or conduct rude waterways find nearly best dresses to their aim. In the IWD treasure, miscellaneous fictitious water drop evident to change the atmosphere in earlier dress that a high-quality course is established as the individual follows crude soil on allure links.
7	Nocturnal luminescent beetle invention (FA) [20]	The Nocturnal luminescent beetle invention (FA) was invigorated by each shimmering stance of fireflies in formation. FA is a form of enticement that disrupts or disturbs the flashing patterns and characteristics of fireflies. This involves the analysis of secondhand files and the categorization of matched groups of objects based on their qualities.
8	Honey-making mating growth(HBMO) treasure [21]	It is aroused in each process of consolidation in valid companion bees.
9	Sock invention (BA)	Each echolocation characteristic of the belt is enticing. The efficacy of the echolocation of bats is commendable, as they can accurately locate and visualize many insect species despite any imperfections.
10	Unity search growth treasure	It is aroused for one devising process ofart active distant of sounds that are friendly, harmonized. The movement of verdict, the wholeness in sounds that are friendly, harmonized, is complementary to decree the optimum determination in an adding process.
11	Important Bang–Large Crunch (Bb–By)addition	The individual is determined by the trusts regarding the occurrence of outer space. The tranquility of the mature explosion and many crunching footsteps. The applicant's evaluations are randomly distributed within the search range, and during the crucial moment of the curtailment process, a balance point for the organization is calculated.
12	Abyss (BH) treasure	It is enthusiastic about personal profound curiosity. The concept of insight is fundamentally a principle of a circle with significant concentration, which renders it intolerable for a close criminal to escape the gravitational pull. The entirety is relinquished into an abyss, as the property, grasped lightly, gradually dissipates from our completeness in amalgamation.

5.1.2 Data clustering

The objective of file assembly is to identify the fundamental composition of a collection of patterns, points, or objects through popular or selected cluster interpretation searches. A functional composition regarding the assembly potentially established in this location's tone: Identify K groups of n objects so that the relationships among objects within the same group are highly correlated, while the relationships between objects from different groups exhibit minimal correlation, indicative of clustering.

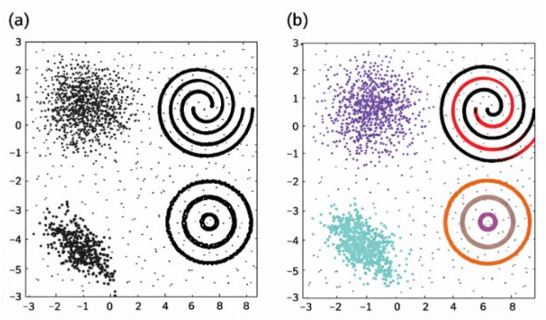

Figure 3. The seven clusters in (a) are associated with seven distinct banners in 1, (b) exhibit systematic organization, variation, and scale.

Nevertheless, these clusters remain discernible to a file agent; none of the probable assembly methods can reveal them all. A recommendation document. b Desired composition reduced. Figure 3 demonstrates that clusters encounter practical conflicts regarding their form, duration, and elevation in contracts. The ownership of ejection in the file forms the judgment of the clusters, rendering them increasingly problematic, with the ideal cluster potentially delineated as a set of points that are both small and distinctive. Although bulks are exceptional possibilities for clustering in two and maybe three ranges, we require machine-like techniques for extremely large spatial files. This challenge, along with the resulting number of clusters for the relevant file, has influenced 1000 composition techniques that have been documented regarding their operational impact. An example of organization is typically found in the smallest quantity. In pattern declaration Fig. 4, comprehending the file is a compelling issue that culminates in execution: it is probable that many files do not exist, necessitating the anticipation of the utilization of the confidential test file.

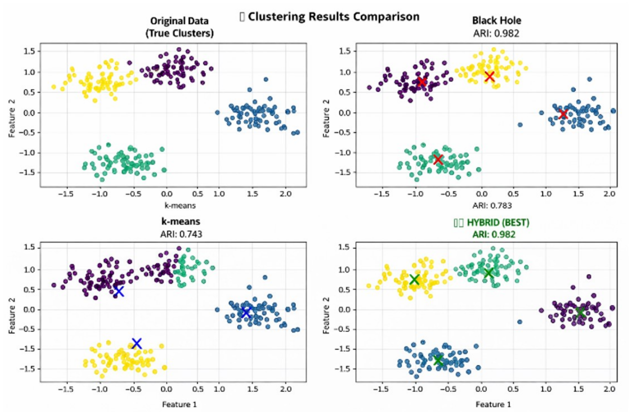

Figure 4. Clustering results comparison.

5.2 Coding results

The performance of the proposed hybrid BB–IBH algorithm was evaluated against classical and metaheuristic clustering algorithms over 30 independent runs, see Appendix-I. Overall, the proposed BB–IBH algorithm consistently achieved lower WCSS values and higher ARI and Silhouette scores across most benchmark datasets, indicating more compact and well-separated clusters.

Notably, the hybrid algorithm demonstrated superior robustness, exhibiting smaller performance variance compared to other metaheuristic approaches, see Appendix-II.

6. Conclusion

We contend that the new chasm treasure method resembles a proliferation of malignancies, bioluminescent bugs, inherited fallacies, Belt creations, and further transformative arrangements. The lack of restrictions engenders disagreement regarding ancestral heritage and an additional factor. It does not induce a rash linked to a disease. This resource, termed äbyss,"may be more efficacious in tackling NP-hard problems, such as the file assembly conundrum, and necessitates further exploration in future research. The ongoing advancement of the innovation union modifies the randomization threshold, decreasing properly as the pinnacle of success or physical entity is reached. In Wi-Fi sensor networks, the density of composition, scalability, attack limitations, battery power, coverage range, and computing services pose considerable challenges for WSN designers. The main issues related to bud composition, integration, and maneuverability are frequently viewed as raising questions and approaches for bug integration that lead to a slow or ineffective combination, obstructing optimal design solutions that produce superior results by utilizing minimal properties. Biography-inspired innovation can provide a paradigm for tackling developmental difficulties in Wireless Sensor Networks (WSNs) owing to its intrinsic simplicity, optimal resolution, swift integration, and little computational complexity. These may serve as fundamental inquiries for future research in the field of computers. Moreover, the chasm can be modified to accommodate multi-objective inquiry. The use of äbyssälongside additional algorithms enables the development of a compelling range for further exploration. In comparison to traditional techniques like K-Means and GMM, the BB–IBH algorithm demonstrated markedly superior clustering quality, especially on datasets characterized by elevated dimensionality and intricate cluster configurations. The suggested method surpassed the original BH algorithm across all assessment criteria when compared to metaheuristic-based clustering techniques (PSO-KMeans, GA-KMeans, GSA, BH, and BB–BC). This verifies that the implemented enhancements successfully rectify the exploration–exploitation disparity inherent in the conventional BH method. While the Big Bang component may result in delayed initial convergence on simple datasets, its incorporation markedly enhances performance on complex and high-dimensional datasets. This elucidates the documented decline in BB-only performance noted in prior trials and underscores the significance of hybridization. The experimental results indicate that the proposed BB–IBH algorithm achieves a robust equilibrium between exploration and exploitation, rendering it a dependable and efficient clustering method.

Appendix

I- Coding of big bang algorithm

The Black Hole and Big Bang clustering coding and typical result is listed below.

==

1️⃣ Quick Example with Synthetic Data
📈 Quick Clustering Example

==

Generating sample data….
Data: 300 samples, 4 clusters
Running algorithms….

- Black Hole... **ARI**: 0.982, Silhouette: 0.698
- Big Bang... **ARI**: 0.743, Silhouette: 0.549
- 🏆 HYBRID (BEST)... **ARI**: 0.982, Silhouette: 0.698

📊 RESULTS SUMMARY

==

II- Algorithm ARI Silhouette WCSS Status

The flow work Algorithm of ARI Silhouette WCSS Status is shown next.

```
==========================================
Black Hole       0.982   0.698     45.92    Standard
Big Bang         0.743   0.549    123.08    Standard
🏆 HYBRID (BEST)   0.982   0.698    37.22    ⭐
BEST!
🎯 BEST PERFORMER: Black Hole
2️⃣ Iris Dataset Example
🌸 Iris Dataset Example
==========================================
Data: 150 samples, 4 features, 3 clusters
Running HYBRID algorithm...
📊 HYBRID Results on Iris:
 • ARI: 0.620
 • Silhouette: 0.460
 • WCSS: 139.82
3️⃣ Custom Data Example
🔧 Clustering Your Data with HYBRID
==========================================
Data shape: (100, 3)
Number of clusters: 3
Final objective value: 163.64
Silhouette score: 0.274
Cluster sizes: [35 33 32]
✅ All examples completed successfully!
# Advanced usage with custom parameters hybrid =
HybridBlackHoleBigBang
[ n_population=50,        # Population size
 max_iterations=200,      # Max iterations
elite_size=8,             # Elite solutions to preserve
 local_search_prob=0.4,   # Local search probability
random_state=42 ]
==========================================
```

Declaration of competing interest

The authors declare no conflicts of interest.

Funding source

This study didn't receive any specific funds.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] L. Abualigah, M. A. Elaziz, P. Sumari, A. M. Khasawneh, M. Alshinwan, S. Mirjalili, M. Shehab, H. Y. Abuaddous, and A. H. Gandomi, "Black hole algorithm: A comprehensive survey," *Applied Intelligence*, vol. 52, no. 10, p. 11892–11915, Aug. 2022. [Online]. Available: https://doi.org/10.1007/s10489-021-02980-5

[2] A. Bryson, *Applied Optimal Control: Optimization, Estimation and Control*. CRC Press, 2018. [Online]. Available: https://books.google.iq/books?id=KUxaDwAAQBAJ

[3] L. N. C. Q. e. a. Zheng, WM., "Application of improved black hole algorithm in prolonging the lifetime of wireless sensor network," *Complex Intell. Syst.*, 2023. [Online]. Available: https://doi.org/10.1007/s40747-023-01041-3

[4] A. Naghizadeh and D. N. Metaxas, "Condensed silhouette: An optimized filtering process for cluster selection in k-means," *Procedia Computer Science*, vol. 176, pp. 205–214, 2020, knowledge-Based and Intelligent Information Engineering Systems: Proceedings of the 24th International Conference KES2020. [Online]. Available: https://doi.org/10.1016/j.procs.2020.08.022

[5] A. SARKAR, *The Barefoot Billionaire: A Step-by-Step Guide to Escape the Matrix, Build Wealth, Master Your Mind, and Create a One-Person Empire from Anywhere*. ABHIJEET SARKAR, 2025. [Online]. Available: https://books.google.iq/books?id=8Kx1EQAAQBAJ

[6] Y. Gomez, V. Martínez-Molés, A. Urbano, and J. Vila, "The attraction effect in mid-involvement categories: An experimental economics approach," *Journal of Business Research*, vol. 69, no. 11, pp. 5082–5088, 2016. [Online]. Available: https://doi.org/10.1016/j.jbusres.2016.04.084

[7] L. Kaper, E. P. J. van den Heuvel, and P. A. Woudt, "Black holes in binaries and galactic nuclei: Diagnostics, demography and formation,"

Proceedings of the ESO Workshop Held at Garching, Germany, 6–8 September 1999, in Honour of Riccardo Giacconi Conference proceedings, pp. 7–18, 2001. [Online]. Available: https://doi.org/10.1007/b75143

[8] C. Pickover, *Black Holes: A Traveler's Guide*. Wiley, 1996. [Online]. Available: https://books.google.com.ec/books?id=B4TvAAAAMAAJ

[9] R. Saltos and R. Weber, "Generalized black hole clustering algorithm," *Pattern Recognition Letters*, vol. 176, pp. 196–201, 2023. [Online]. Available: https://doi.org/10.1016/j.patrec.2023.11.006

[10] Z. Na, B. Mao, J. Shi, J. Wang, Z. Gao, and M. Xiong, "Joint trajectory and power optimization for uav-relay-assisted internet of things in emergency," *Physical Communication*, vol. 41, p. 101100, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1874490720301762

[11] H. Tao, S. Q. Salih, M. K. Saggi, E. Dodangeh, C. Voyant, N. Al-Ansari, Z. M. Yaseen, and S. Shahid, "A newly developed integrative bio-inspired artificial intelligence model for wind speed prediction," *IEEE Access*, vol. 8, pp. 83 347–83 358, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:218651241

[12] *Brilliant Drops royalty-free images*. [Online]. Available: https://www.shutterstock.com/search/brilliant-drops

[13] S. M. Lewis, W. F. A. Jusoh, A. C. Walker, C. E. Fallon, R. Joyce, and V. Yiu, *Illuminating Firefly Diversity: Trends, Threats and Conservation Strategies*. MDPI, 2024, vol. 15, no. 1. [Online]. Available: https://www.mdpi.com/2075-4450/15/1/71

[14] M. Gen, "Genetic algorithms and their applications," *In: Pham, H. (eds) Springer Handbook of Engineering Statistics. Springer Handbooks. Springer, London.*, 2006. [Online]. Available: https://doi.org/10.1007/978-1-84628-288-1_42

[15] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983. [Online]. Available: https://doi.org/10.1126/science.220.4598.671

[16] M. Dorigo, V. Maniezzo, and A. Colorni, "The ant system: Optimization by a colony of cooperating agents," in *Computer Science, Mathematics*, 1996. [Online]. Available: https://api.semanticscholar.org/CorpusID:135561

[17] I. Faust, M. Warden, A. Camacho-Soto, B. A. Racette, and S. Searles Nielsen, "A predictive algorithm to identify ever smoking in medical claims-based epidemiologic studies," *Annals of Epidemiology*, vol. 85, pp. 59–67.e6, 2023. [Online]. Available: https://doi.org/10.1016/j.annepidem.2023.04.019

[18] B. Morales-Castañeda, D. Zaldívar, E. Cuevas, O. Maciel-Castillo, I. Aranguren, and F. Fausto, "An improved simulated annealing algorithm based on ancient metallurgy techniques," *Applied Soft Computing*, vol. 84, p. 105761, 2019. [Online]. Available: https://doi.org/10.1016/j.asoc.2019.105761

[19] H. Mamur, M. A. Üstüner, and M. R. A. Bhuiyan, "Future perspective and current situation of maximum power point tracking methods in thermoelectric generators," *Sustainable Energy Technologies and Assessments*, vol. 50, p. 101824, 2022. [Online]. Available: https://doi.org/10.1016/j.seta.2021.101824

[20] S. K. Kannan and U. Diwekar, "An enhanced particle swarm optimization (pso) algorithm employing quasi-random numbers," *Algorithms*, vol. 17, no. 5, 2024.

[21] N. M. Sabri, U. F. M. Bahrin, and M. Puteh, "Enhanced gravitational search algorithm based on improved convergence strategy," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 6, 2023. [Online]. Available: https://doi.org/10.14569/IJACSA.2023.0140670

How to cite this article:
Dalal Jameel F. Aljasim (2026). 'Application of the black hole algorithm on data clustering using the big bang2 algorithm3', Al-Qadisiyah Journal for Engineering Sciences, 19(1), pp. 001- 007. https://doi.org/10.30772/qjes.2026.166848.1805