Optimizing IoT Service Matching Using Simulated Annealing Enhanced K-means Clustering

Weiyong Gao[1*], Xia Chen[2]
[1]College of IoT and Communications, Henan Institute of Economics and Trade, Zhengzhou 450003, China
[2]College of Commerce and Logistics, Henan Institute of Economics and Trade, Zhengzhou 450003, China
E-mail: gaoweiyong2024@163.com, chenxia5202023@163.com
*Corresponding author

Keywords: internet of things technology, simulated annealing algorithm, mean clustering, service matching, clustering effect

Received: February 19, 2025

With the wide application of IoT technology, traditional IoT service models face challenges in terms of service matching efficiency and computational burden. Although the existing K-means clustering algorithm is widely used, it is sensitive to the initial centre of mass selection and is prone to fall into local optimal solutions, which affects the accuracy of the clustering results and the service matching efficiency. For this reason, the study proposes a K-means clustering method combined with a simulated annealing algorithm. By simulating the physical annealing process, the local optimal problem is effectively avoided and the global optimization ability of clustering is improved. The experimental results show that the Silhouette score of SA-K-means clustering is 0.82, and the Davis Boulding index is only 0.41. Under high concurrency, the proposed algorithm achieves a data reception accuracy of 97.8%, a response time of 1276.18ms, and a service success rate and reliability of 95.12% and 96.75%. In addition, the median inverse generative distance of SA-K-means in different complexity scenarios is 0.0061 and 0.0065, which is closest to the Pareto optimal frontier. This study provides a new theoretical approach for IoT service matching, enriches the research content, and provides a reference for other fields that require efficient clustering processing.

Povzetek: Članek predstavi SA-K-means algoritem za izboljšano ujemanje IoT storitev, ki združuje simulirano ohlajanje in K-means ter dosega boljšo natančnost, odzivni čas in robustnost.

1 Introduction

With the rapid development of technology and the wide adoption of smart devices, the Internet of Things (IoT) has become a key technology in modern life. IoT services bring about a smarter, safer, and sustainable lifestyle through seamless connectivity between devices, people, and systems [1-2]. However, despite the great potential of IoT technologies, traditional IoT service models still face many challenges in terms of service matching efficiency and computational burden. Matching services in the IoT present a significant challenge, as current algorithms struggle to keep pace with the escalating requirements of dynamic and expansive IoT systems [3]. Traditional approaches, like K-means clustering, are extensively utilized; however, they are highly susceptible to the initial selection of centroids and often succumb to local optimal solutions [4-5]. Especially when dealing with complex and large-scale data in IoT applications, this drawback can seriously affect the accuracy of clustering results and the efficiency of service matching. To solve these problems, this study proposes to combine the simulated annealing (SA) algorithm with the K-means clustering method to construct a new SA-K-means model. The SA algorithm, as a global optimization technique, is inspired by the physical annealing process and has an enhanced global search capability, which is able to effectively avoid the common local minima problem of traditional clustering methods. Therefore, the study applies the proposed application to IoT services and constructs models. The research aims to improve the efficiency of service matching in IoT and provide users with high-performance and smooth IoT service technologies while quickly finding the optimal solutions they need.

The innovation of the research lies in combining the SA algorithm with the traditional K-means clustering algorithm to form a new SA-K-means model. The SA algorithm effectively avoids the problem that the traditional K-means algorithm is prone to fall into the local optimum by simulating the physical annealing process, and thus enhances the global optimization ability of clustering. By applying the optimized SA-K-means algorithm to the field of IoT service matching, an efficient IoT service model is constructed, which can effectively cope with the challenges of service matching in large-scale and complex data environments, and provide users with a more efficient and accurate IoT service experience. The model proposed by the research can better adapt to the dynamically changing demands in the IoT environment, providing strong technical support for the efficient operation of IoT services.

The SA-K-means algorithm proposed in the study provides a new theoretical approach to the clustering problem in IoT service matching and enriches the research in this field. Based on this, the constructed IoT service model excels in service matching efficiency, data reception accuracy and response time, and is able to quickly converge and adapt to complex scenarios, which improves the overall performance of the IoT system. Through a large number of experiments, it is verified that the model can still maintain efficient service success rate and reliability under high concurrency and large-scale data environments, which provides a strong practical basis for the actual deployment and application of IoT service.

2 Related work

As an innovative technology, IoT services can achieve automatic irrigation and fertilization in agricultural production. When people drive vehicles, they can provide clear and precise route planning through satellite positioning. IoT services are applicable across a broad spectrum of many fields and have profound significance and influence [6]. Many researchers around the world have conducted research in the field of IoT services. For example, Kasilingam D et al. analysed the application of IoT in marketing through least squares structural equation modelling. The study used online questionnaire data to explore the impact of IoT technology on market acceptance of different consumer groups [7]. Khan A A et al. proposed an industrial IoT framework that supports blockchain technology to address challenges and limitations related to information storage, node transactions, privacy, and security protection in existing IoT technologies. Simulation experiments showed that the framework could provide a secure and trustworthy execution environment [8]. Ushakov D et al. collected data from several public transport sectors and analysed the application of IoT technology in global transport systems. By analysing the functions of different public transport sectors, the study examined how IoT can optimize the operations of airlines and other transport systems by collecting real-time data (e.g., weather, traffic conditions, flight schedules, etc.) [9]. Wei L et al. analyzed the definition, aggregation, and computation of trust in response to security incidents, data breaches, and service fraud issues in the development of IoT technology. The objective was to elucidate the various dimensions of trust issues and management within the realm of the IoT, while also piquing the interest of prospective scholars to conduct further research in this domain [10].

K-means is a commonly-used distance-based algorithm, which has the advantages of compact optimization within clusters, simple and efficient search for optimal solutions, and is commonly used in IoT service clustering processing [11]. Many researchers have carried out discussions and study on the K-means. For example, Ikotun A M et al. discussed that K-means, as the mainstream clustering algorithm, still faces many challenges. The capability of this algorithm is easily affected by the selection of the initial cluster, and its greediness has fallen into the minimum local convergence. They also introduced research work to improve the performance and robustness of K-means, and discussed the future research prospects of recommendation [12]. Zubair M et al. raised an effective method for finding the optimal initial centroid to reduce the number of iterations and execution time of the K-means. After analyzing examples from 8 dimensions, the results showed that this method was superior to the traditional K-means and the random clustering center initialization method in terms of iteration times and computation time [13]. In addition, the SA algorithm exhibits the characteristic of efficiently finding global optimal solutions in large spaces and large-scale data, and has the advantage of effectively solving various complex problems [14]. Many researchers have carried out research on the SA algorithm. For example, Shi K et al. raised an improved dynamic path planning SA algorithm to address the limitations of existing methods in terms of computational workload, making it difficult for robots to avoid obstacles in dynamic situations. Through experimental analysis, this method was determined to outperform other approaches and could provide the optimal solution in dynamic environments [15]. Fontes D B M M et al. raised a hybrid particle swarm optimization and SA algorithm, derived a fast lower bound process, and conducted extensive computational experiments on 73 benchmark test cases. The outcomes indicated that the method had superior performance, solved a variant of the job shop scheduling problem, and improved the overall performance of the manufacturing system [16]. Liu J et al. proposed to optimize the heat extraction part in energy tunnels by combining numerical simulation and SA to address the shortcomings of the design method for ground source heat pump systems. After experimental verification, the results showed that the normalized feasible range in the tunnel increased by 1099.6% and the inlet fluid temperature decreased by 64.3%, proving that this method has broader prospects [17].

From the research of domestic scholars mentioned above, most scholars only explore and analyze the vision and challenges of IoT services, but there are few substantial breakthroughs. For example, Ikotun A M et al. proposed a strategy of initial centre of mass optimization for K-means, but the method relies on heuristic search and cannot effectively avoid the local optimum problem. Zubair M et al. on the other hand, proposed an improved initialization method to reduce the number of iterations, but this is still not able to solve the problem of slow convergence and local optimum of K-means when dealing with high dimensional data. In addition, although the SA algorithm itself has strong global optimization capabilities, when combined with K-means, many studies focus on general optimization problems such as path planning or scheduling optimization, and lack the application of matching specific requirements for IoT services. The majority of current research endeavors concentrate on enhancing a solitary facet of K-means, while lacking a holistic approach to concurrently elevate both the efficiency and precision of IoT service matching. Existing methods perform poorly when dealing with large-scale and dynamically changing data, and are not effective in improving service response speed and accuracy. Therefore, the study proposes to optimize the clustering processing method of K-means using SA algorithm and construct an IoT service model, aiming to achieve efficient operation of IoT services and quickly provide accurate service content to users. A comparison of optimization studies of different K-means clustering algorithms is shown in Table 1.

Table 1: Comparison of different methods.

References	Methodology	Data sets	Evaluation indicators	Key findings	Weaknesses
Kasilingam D et al. [7]	Least squares structural equation modelling	Online survey data	Predictive capability	Models with high predictive capability to assist in the development of market strategies	Primary focus on market acceptance, no specific technology optimization involved
Khan A A et al. [8]	Blockchain-enabled industrial IoT framework	Industrial activity data	Security and trustworthiness	Providing a secure and trustworthy execution environment	Not optimized for service matching efficiency
Ushakov D et al. [9]	Functional analysis	Public transport agency data	Impact analysis	Analyses the impact of IoT on global transport systems	Lack of specific technical optimization and experimental validation
Wei L et al. [10]	Trust management analysis	Home-grown dataset	Trust dimensions	Provides a multi-dimensional understanding of IoT trust issues	Does not address service matching efficiency improvements
Ikotun A M et al. [12]	K-means improvement study	Homemade dataset	Clustering performance	Proposed a method to improve the performance of K-means	Not combined with a global optimization algorithm, may still fall into a local optimum
Zubair M et al. [13]	Method for optimizing initial clustering centers	Benchmark test dataset	Number of iterations and computation time	Outperforms traditional K-means and random initialization methods	Not validated for application in IoT services
Shi K et al. [15]	Improved SA algorithm for dynamic path planning	Robot dynamic environment data	Obstacle avoidance capability	Outperforms other methods in dynamic environments	Not applied to IoT service matching field
Fontes D B M M et al. [16]	Particle swarm optimization with SA hybrid algorithm	73 benchmark cases	Manufacturing system performance	Job shop scheduling problem solved and overall performance improved	Higher computational complexity
Liu J et al. [17]	Optimization approach for joint numerical simulation and SA	Ground source heat pump system data	Thermal extraction efficiency	Enhanced design performance of the thermal extraction part	Limited computational efficiency
This paper	SA-K-means algorithm	IoT service data	Clustering effect, response time, data receiving accuracy	Effective technical support for IoT services	-

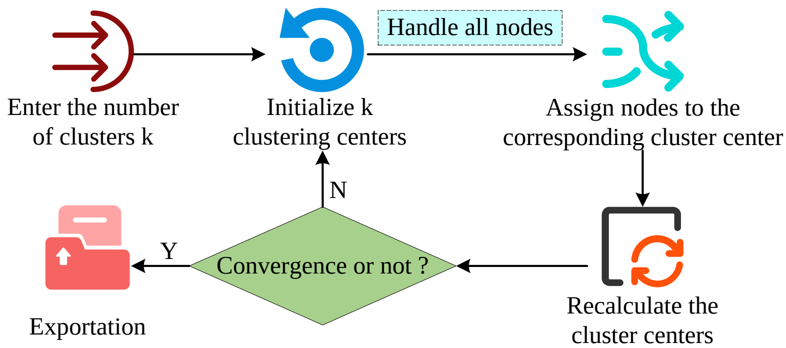


Figure 1: The running steps of K-means.

3 Improved clustering processing technology for IoT services

3.1 Improved K-means design incorporating SA algorithm

With the rapid development of the IoT, IoT services have penetrated into multiple fields such as industrial control, biotechnology, and information engineering. The demand for IoT services is increasing, and it has become increasingly urgent to quickly find corresponding service tasks based on user needs [18]. The service matching capability, as the core of IoT services, traditional keyword-based service matching methods are inefficient and no longer meet the large-scale demand for IoT services today [19]. The K-means is commonly used in target matching and classification tasks due to its advantages of fast computation and stability in large-scale data. Its core idea is to maximize the compactness within the clustering range [20]. Given the remarkable efficiency and robustness of the K-means algorithm, it is proposed to apply it in the context of IoT services. The running steps of K-means are shown in Figure 1.

As shown in Figure 1, the object is first given, then inputted into k clusters, and after multiple iterations, the initial cluster centers are determined. Then, all nodes are processed and assigned to their corresponding cluster centers. Once the allocation is finalized, the mean value of the existing cluster center is computed and designated as the new center. Subsequently, similarity measures are evaluated and categorized to see if the convergence condition is met. The K-means algorithm uses Euclidean distance to evaluate the similarity of each data point to the current clustering center. The data points are classified based on their similarity to the cluster centre and reassigned to the closest cluster centre. This classification process directly affects the subsequent clustering centre calculation and data point assignment. If it does not converge, the process returns to the initialization cluster center stage and repeats until convergence is achieved. Finally, the clustering result is output, which is the final output obtained by the IoT service. The expression for the square of the distance between all objects and the cluster center in the K-means is shown in equation (1).

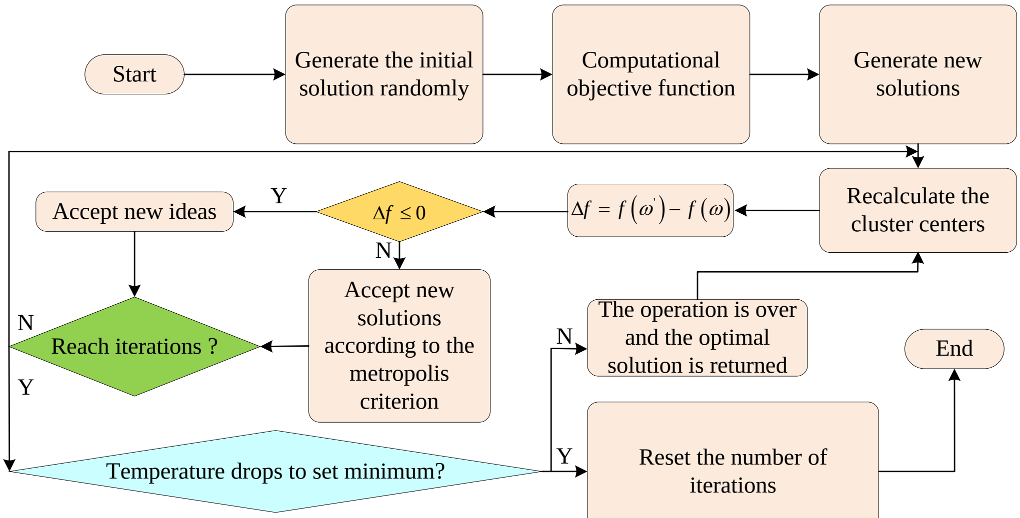


Figure 2: The execution phases of the SA algorithm.

$$J = \sum_{j=1}^{k} \sum_{x_i \in C_j} \left\| x_i - z_j \right\|^2 \quad (1)$$

In equation (1), k and x_i respectively represent the number of categories and sample objects, and z_j represents the clustering center points in C_j. The K-means has advantages such as simplicity, effectiveness, and strong global search capability when applied to IoT services. However, it relies too heavily on the initial clustering center, which can easily lead to local optimization of clustering results and reduce the efficiency of global search. The idea of SA algorithm is to simulate the process of physical annealing and cooling, and it has the characteristics of parallelism and gradual convergence for finding the global optimal solution [21]. Therefore, the study proposes to optimize and improve the K-means using the SA algorithm. The implementation process of the SA algorithm is in Figure 2.

In Figure 2, the execution phases of the SA algorithm first randomly generate an initial solution ω and calculate the objective function $f(\omega)$, and subsequently, a new solution and its corresponding objective function value are randomly generated, and the discrepancy Δf between this new solution and the initial one is computed. If the newly generated solution demonstrates superior performance compared to the existing initial solution, it shall be promptly implemented. Conversely, if its performance is inferior, its acceptance will be subject to the Metropolis criterion. Following a series of iterations, if the predetermined iteration count has not been reached, the process will revert to the step of generating a new solution. Once the iteration count is attained, an assessment will be made to determine whether the termination condition has been fulfilled. In the event that the stopping criterion has not been met, the temperature will be gradually decreased, and the iteration count will be reset. Subsequently, the process will loop back to the step of obtaining a new solution and continue repeating until both the iteration count and termination condition are satisfied, with the entire operation being carried out in a passive manner. Upon completion of the operation, the process will revert to the optimal solution obtained. The expression for the cooling probability of the energy difference is shown in equation (2).

$$p(dE) = e^{dE/kT} \quad (2)$$

In equation (2), dE represents the energy difference of temperature T. Due to the gradual cooling, therefore $0 < p(dE) < 1$. The expression for the probability of state transition is shown in equation (3).

$$p(\Delta f) = e^{\Delta f/T} \quad (3)$$

In equation (3), f is the state generation function. After gradual annealing, the global optimization ability of the SA algorithm also improved. The study will integrate the SA algorithm to optimize the K-means. The optimization process is in Figure 3.

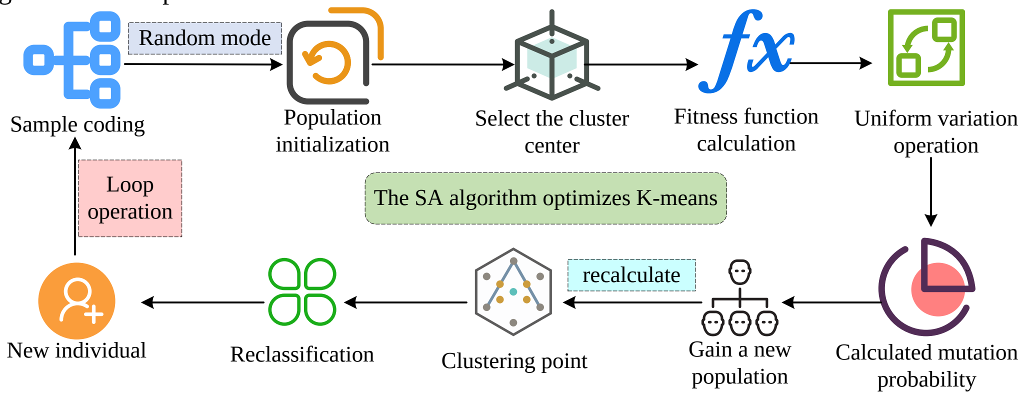


Figure 3: The SA algorithm optimizes the K-means process.

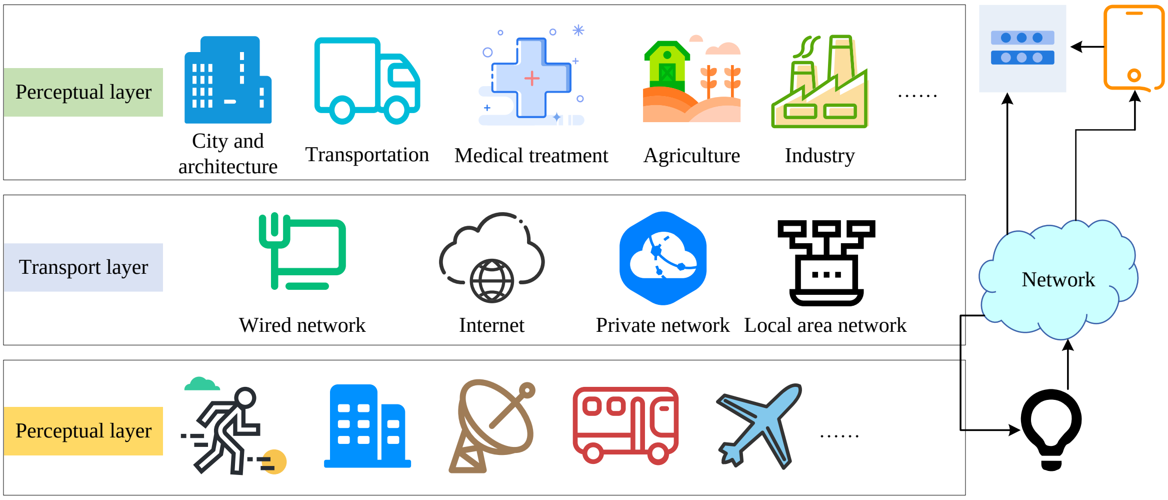


Figure 4: Architecture diagram of IoT application.

In Figure 3, the SA algorithm optimizes K-means by firstly generating new clustering solutions on the basis of the current clustering centre through the uniform variation operation to ensure that the solutions are uniformly distributed in the search space and avoid falling into local optimum. Secondly, the fitness value of each solution is calculated to evaluate its clustering effect, and the optimal clustering centre is selected based on the fitness value to optimize the clustering result. Then the mutation probability is calculated according to the mutation operation to decide whether to accept the new solution or not. When the mutation probability is high, the mutation of more solutions is allowed to increase the diversity. Ultimately, the clustering center is recomputed, and the clustering procedure is executed iteratively until the convergence criterion is met, thereby concluding the optimization process. Equation (4) calculates the difference between the maximum distance and the current distance to evaluate the fitness value and update the clustering centre.

$$P(S_i) = 1.5 \times J_{\max} - J(S_i), i = 1, 2, \ldots sizepop \quad (4)$$

In equation (4), $J_{\max}$ and $J(S_i)$ are the maximum value of all individual J values and the current individual J value, respectively. The J value represents the distance from the sample object to the cluster center point and *sizepop* represents the population size. The probability expression for selecting individuals to form a new population is shown in equation (5).

$$P_i = \frac{F(S_i)}{\sum_{i=1}^{k} F(S_i)}, i = 1, 2, \ldots sizepop \quad (5)$$

In equation (5), F is the area corresponding to the individual and P_i is the probability of selecting the individual i. The definition expression for the mutation probability when a new individual is obtained during the mutation operation process is shown in equation (6).

$$P_i = \frac{1.5 \times d_{\max}(x_i) - d(x_i - c_k) + 0.5}{\sum_{k=1}^{k} 1.5 \times d_{\max}(x_i) - d(x_i - c_k) + 0.5} \quad (6)$$

In equation (6), $d(x_i - c_k)$ represents the distance between the centroid c_k of the k th cluster and the sample x_i. c_k denotes the variation factor of the clustering center and x_i denotes the sample. The distance is Euclidean distance, expressed as equation (7).

$$d_{\max}(x_i) = \max_k \{d(x_i - c_k)\} \quad (7)$$

Afterwards, the clustering center points for each individual in the new population will be recalculated, as shown in equation (8).

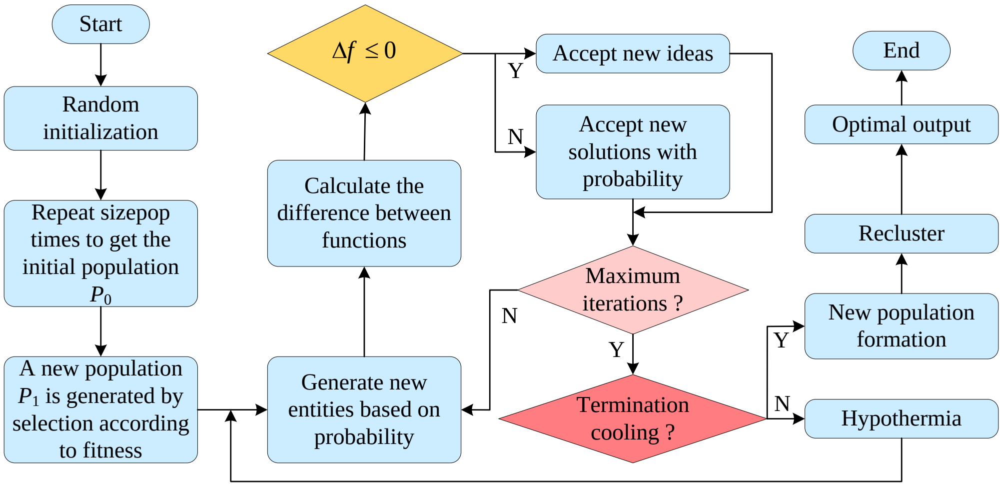


Figure 5: SA-K-means clustering process in IoT services.

$$z_j^* = \frac{1}{n_j} \sum_{x_m \in z_j} x_m, \ j = 1, 2, \ldots, k \tag{8}$$

After obtaining all individuals in the population, the above operation will be repeated to obtain the next round of population. After the above operations, the research will form a K-means clustering processing method optimized by the SA algorithm, namely the SA-K-means algorithm.

3.2 Construction of SA-K-means model in IoT services

In contemporary times, an expanding array of fields are becoming reliant on IoT services, and as demand continues to surge, IoT service technology is confronted with even more formidable challenges [22]. If the maturity of IoT service technology is not sufficient, it will lead to a decrease in user experience, and the low efficiency of service matching will also affect the efficiency of industrial production. The architecture of IoT applications is shown in Figure 4.

As shown in Figure 4, the IoT application framework is divided into application layer, transmission layer, and perception layer. The application layer includes cities and buildings, healthcare and security, transportation and logistics, agriculture and environment, and industrial production. The transmission layer encompasses telecommunication networks, the Internet, as well as local area networks, whereas the perception layer comprises terminal devices like buildings, vehicles, and satellites. Internet service pertains to the physical devices that connect to the IoT platform and relay data and control commands to mobile phones or displays via the Internet and home LAN. The IoT platform will perform device management and grouping, perform data parsing and storage operations, implement real-time monitoring, and finally transmit information to application entities [23]. To improve the efficiency of IoT service matching and accelerate information transmission speed, the SA-K-means is proposed to be applied to the IoT service matching stage for clustering processing in IoT services. The clustering process of SA-K-means in IoT services is shown in Figure 5.

As shown in Figure 5, the SA-K-means clustering processing steps in IoT services first randomly initialize and generate an individual based on the received information, repeat sizepop times to obtain the initialized population, then calculate the fitness value of each individual, and select and generate a new population based on the fitness value. Next, the current solution is obtained based on the function criterion function, and then a new solution is obtained based on probability. The discrepancy between the new and current solutions is determined. If the new solution surpasses the current one, it is adopted. If it is inferior, acceptance is granted based on a probabilistic assessment. Afterwards, it will be observed whether the max quantity of iterations has been reached. If not, the step of calculating the probability for a new solution will be returned to until the max quantity of iterations is reached. Then it will be checked if the termination cooling has been reached. If not, the cooling operation will be performed and the step of calculating the probability will be returned to until the termination cooling is reached. After the second new population has been generated, the current cluster center is calculated and the re-clustering is performed. After the third new population is generated, the individual with the highest fitness value is selected from the population as the best clustering result, which is the best service matching content for the IoT. In this case, the initial clustering centre is determined mainly based on the received IoT service data. The initial clustering centre is selected by calculating the distance between the data points, as shown in equation (9).

$$K = \sqrt{\frac{M}{2}} + 2 \tag{9}$$

In equation (9), K is the number of clusters and M is the number of matching results. At each iteration, the clustering centre is recalculated based on the members of the current cluster to ensure that its position is constantly optimized. The specific calculation formula is shown in equation (10).

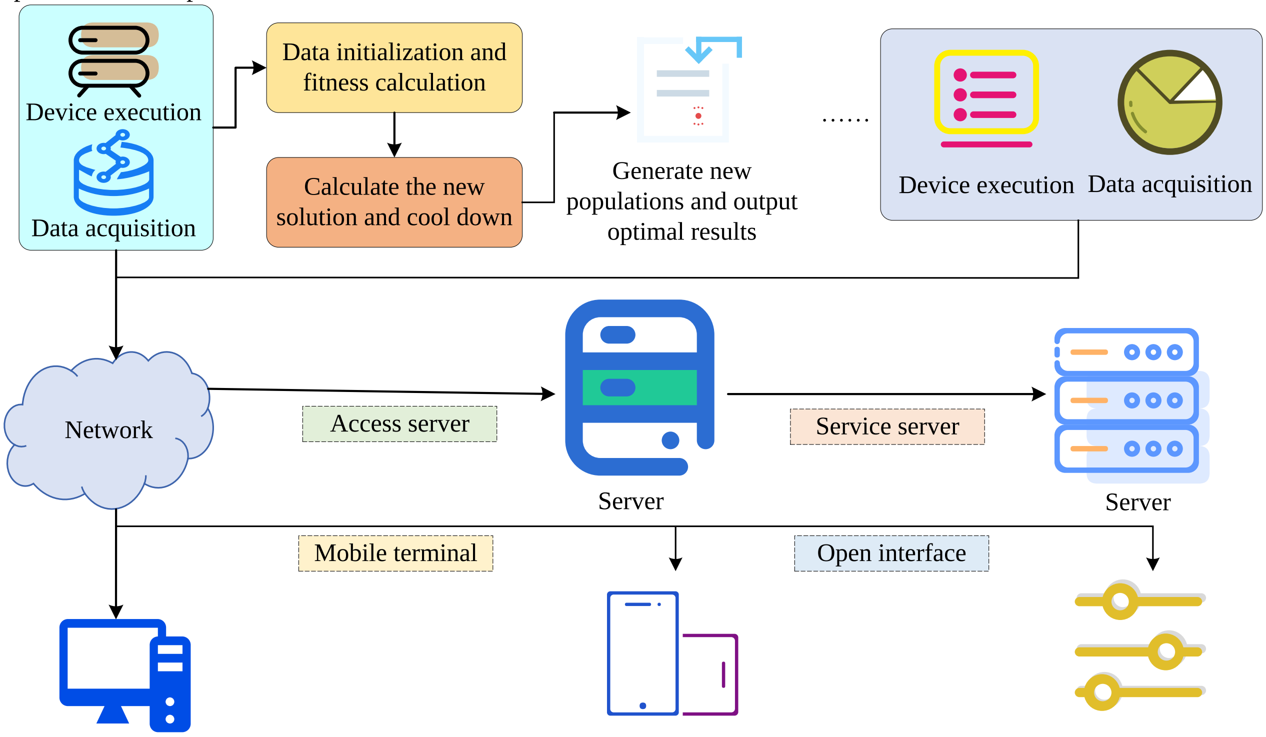


Figure 6: IoT service model based on SA-K-means.

$$Z_j\left(I+1\right)=\frac{1}{n}\sum\nolimits_{i=1}^{n}X_i\left(J\right), j=1,2,\ldots k \qquad (10)$$

In equation (10), $I+1$ represents increasing the number of iterations I by 1 and recalculating the cluster centers to obtain a better solution. By gradually decreasing the temperature, the algorithm gradually stabilizes, thus reducing the probability of accepting a bad solution. When the termination of cooling has not been reached, the cooling expression is shown in equation (11).

$$t_{i+1}=\frac{t_i}{\log_{10}1+\alpha} \qquad (11)$$

In equation (11), t_i and t_{i+1} respectively represent the current temperature and the temperature at the next moment. For rapid cooling, its expression is shown in equation (12).

$$t_{i+1}=\alpha t_i\left(i\geq 0,\alpha\in\left(0,1\right)\right) \qquad (12)$$

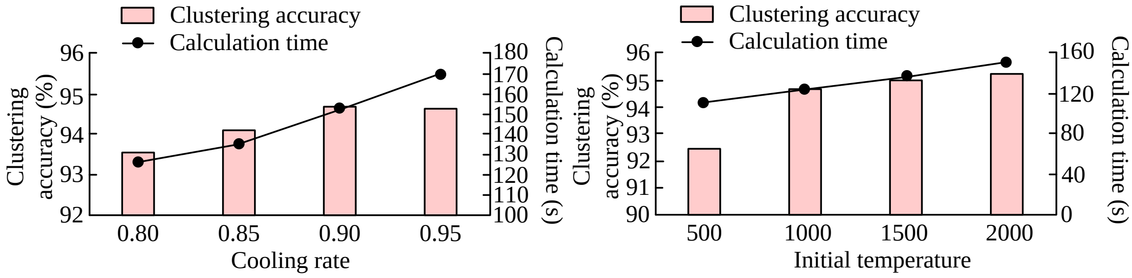
After the cooling operation, when generating the latest population, the individual with the highest fitness value in the population is selected as the final optimal clustering result of the SA-K-means, which is the optimal matching result in IoT services. The study applied the SA-K-means to IoT services and constructed an SA-K-means model. The IoT service model based on the SA-K-means is shown in Figure 6.

In Figure 6, the data acquisition module is responsible for collecting real-time data from different end devices such as sensors and monitoring devices, and the data is transmitted to the IoT platform for processing through the IoT network. Conversely, the device execution module is responsible for directing the devices to carry out relevant tasks in accordance with the data processed by the platform. This includes modifying device parameters, initiating or halting specific functions, and other similar operations. The SA-K-means model in IoT services first collects data from different executing devices, performs data initialization and recalculates cluster centers, and iteratively obtains the final optimal result. Subsequently, the results are conveyed via the network to access devices and business devices, and further to mobile terminals and open interfaces, among others. This enables the application of the ultimate desired outcomes in various domains such as healthcare, industrial production, agriculture, and personal life, thereby enhancing the efficiency of IoT services and elevating the user experience.

To improve the performance of SA-K-means algorithm in IoT services, the study selects the key parameters such as initial temperature, cooling factor and probability of variation through empirical optimization methods. The initial temperature determines the probability of accepting a worse solution at the beginning of the algorithm, and the cooling coefficient controls the rate of temperature decrease, thus affecting the global and local nature of the search. The variance probability is then adjusted through multiple simulated annealing experiments to ensure diversity in the solution space and to avoid locally optimal solutions. The study begins by setting a higher initial temperature based on preliminary experiments and continuously adjusting the rate of temperature decrease and the variance probability to obtain the best clustering results.

Table 2: Index analysis experimental equipment and configuration table.

Experimental tool	Disposition
CPU	Intel(R) Core (TM) i5-11300H @3.10 GHz 3.11 GHz
Graphics card	NVIDIA GeForce RTX 3090
Running memory	16GB
Storage space	1TB NVMe SSD
Simulation software	MATLAB
Input device	Touch screen, wireless keyboard, mouse
Operating system	Windows 10(64bit)
Language environment	Python
Learning framework	TensorFlow



(a) Effect of cooling rate on clustering performance (b) Effect of initial temperature on clustering performance

Figure 7: Effect of different values of SA parameters on the clustering performance of the algorithm.

4 Performance analysis of SA-K-means clustering processing technology integrated in IoT services

4.1 Performance analysis of SA-K-means

To confirm the superior capability of the SA-K-means in IoT services, a performance comparison analysis was conducted between the Density-Based Spatial Clustering of Applications with Noise Algorithm (DBSCAN), Mean Shift Clustering Algorithm (MS), and Multi-layer Perceptron Algorithm (MLP). The experimental environment was set up, with the main configurations shown in Table 2.

Under the experimental environment in Table 2, the SA-K-means, DBSCAN, MLP, and MS algorithms were tested and trained using a leaf dataset to compare their fitness values, response time, and data reception accuracy. The leaf dataset (https://archive.ics.uci.edu/ml/datasets/Leaf) is a common dataset used for clustering and classification tasks with good diversity and structure. Although this dataset itself is not directly from a real IoT scenario, it is representative in terms of data characteristics and processing complexity, and thus can effectively validate the performance of this model. The effect of different values of SA parameters on the performance of the algorithm was first analyzed, as shown in Figure 7.

Figure 7(a) shows the effect of the value of the cooling rate on the clustering performance of the algorithm when the initial temperature was fixed at bit 1000. As the cooling rate increased, the clustering accuracy first increased and then tended to stabilize. When the cooling rate was 0.90, the clustering accuracy reached the highest value of 94.72%. Figure 7(b) shows the effect of different initial temperatures on the clustering performance of the algorithm when the cooling rate was fixed at 0.90. Higher initial temperatures allowed the algorithm to explore the solution space more extensively in the early stages, thus improving the clustering accuracy. However, a high initial temperature also led to an increase in computation time. The clustering effect is shown in Figure 8.

As shown in Figure 8(a), the MS algorithm divided a large number of target samples into other samples, and its clustering effect was the worst. As shown in Figure 8 (b), the clustering effect of the MS algorithm was slightly better than that of the MLP algorithm, but there were still some target samples that were misclassified. As shown in Figure 8 (c), the DBSCAN algorithm divided a small portion of the target samples into other samples, resulting in higher clustering performance. From Figure 8 (d), SA-K-means recognized a very small portion of target samples as other samples, and its recognition performance was the best among the four algorithms. The experimental outcomes showed that SA-K-means had the best clustering effect and the most accurate recognition of samples. The clustering performance of the four algorithms is shown in Table 3.

In Table 3, the closer the Silhouette score and Adjusted Rand Index (ARI) are to 1 means the better the clustering, while the smaller the Davis Boulding index means the better the clustering. It can be seen that SA-K-means increased the Silhouette score by an average of 28.13% over the other three algorithms. The Davis Boulding index was used to evaluate the separation and compactness of clustering, and lower values indicate better clustering results. The low index of SA-K-means indicated that its clustering results had good separation and compactness, and it could efficiently divide the data points into different clusters. Comparing the ARI of the

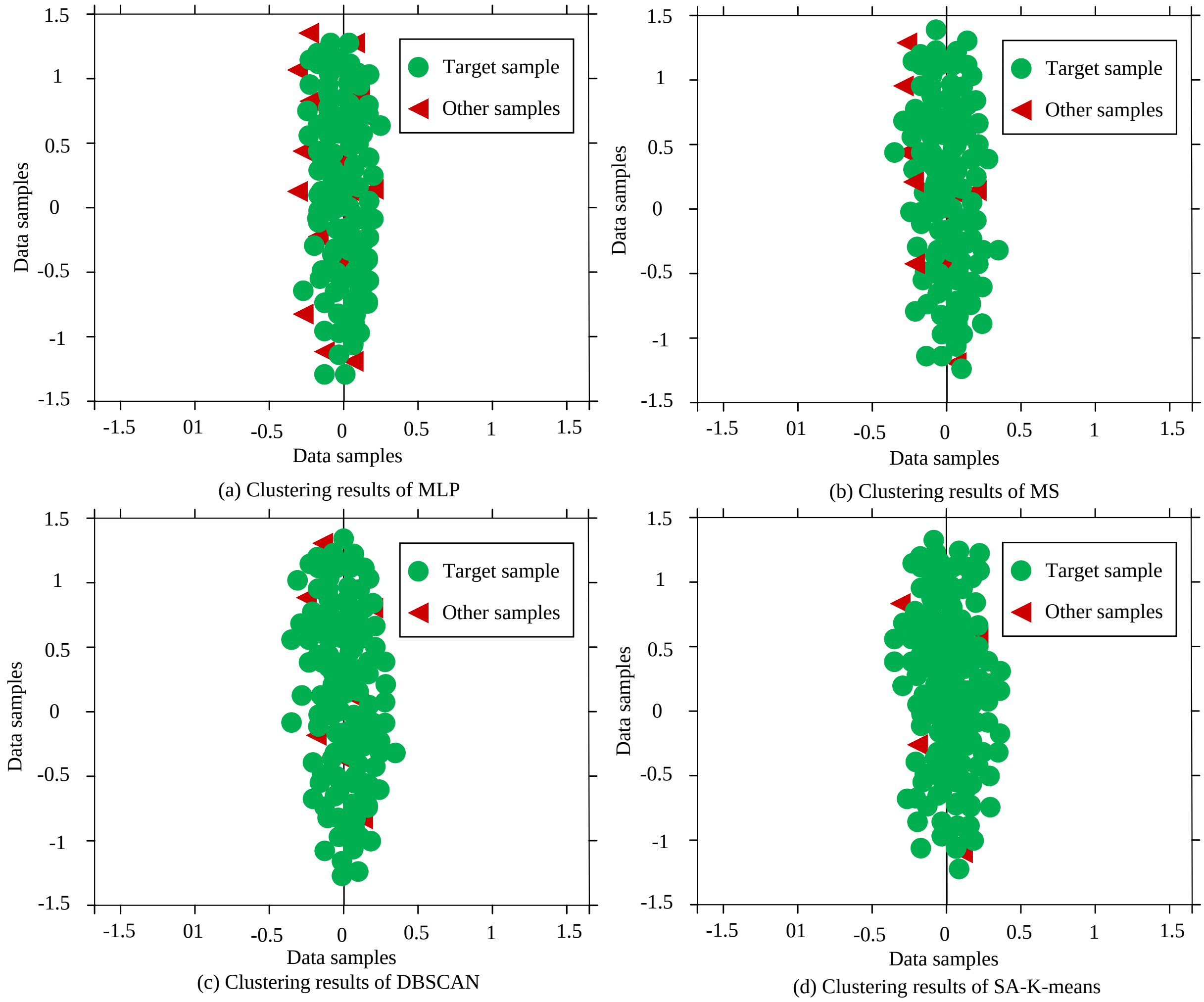


Figure 8: Comparison of sample division of four algorithms.

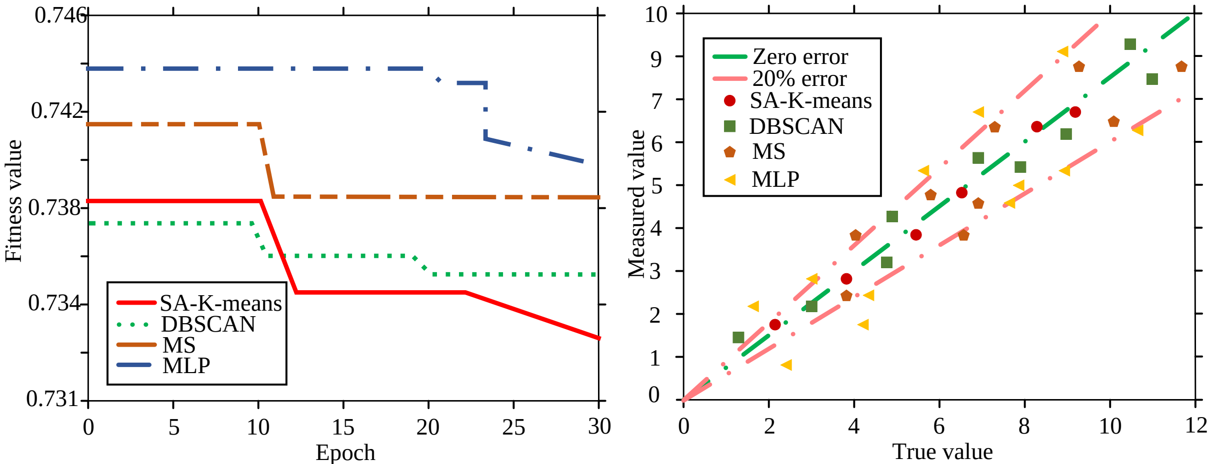
Table 3: Comparison of clustering performance of 4 algorithms.

Algorithm	Silhouette score	Davis Boulding index	ARI
SA-K-means	0.82	0.41	0.93
DBSCAN	0.68	0.63	0.87
MS	0.59	0.74	0.82
MLP	0.65	0.69	0.85

four algorithms, it can be seen that SA-K-means was superior. This indicated that its clustering results were highly consistent with the real labels and could accurately identify the classes of data. To further confirm the adaptability and accuracy of the SA-K-means, four algorithms were tested, and their fitness values and error distribution results are shown in Figure 9.

According to Figure 9 (a), as the number of iterations increased, the fitness values of the four algorithms gradually decreased. Among them, the DBSCAN algorithm had the lowest initial fitness value of 0.7377. After 12 iterations, the fitness value of SA-K-means dropped to 0.7346, which was lower than that of DBSCAN. When the number of iterations reached 30, the fitness values of SA-K-means, DBSCAN, MLP, and MS algorithms were ranked from high to low as 0.7398, 0.7383, 0.7356, and 0.7327, respectively. According to Figure 9 (b), the measured values of the MS algorithm were distributed around the 20% error line, while the measured values of the MLP algorithm were distributed between the 0 error and 20% error lines, but relatively far from the 0-error line. The measured values of the DBSCAN algorithm were distributed around the zero-error line, but still at a considerable distance from it. Meanwhile, the measured values of SA-K-means closely followed the zero-error line, approaching a linear distribution. The experimental results showed that SA-K-means had a faster convergence speed, the best adaptability, and the smallest error, proving that it was more accurate in fitting the true values when measuring targets.

To further confirm the superiority of the SA-K-means model performance, the study conducted statistical significance tests on the DBSCAN, Mean Shift, MLP, and SA-K-means algorithms. Firstly, the clustering effectiveness of each pair of algorithms was compared



(a) Comparison of fitness values (b) Error contrast

Figure 9: Fitness values and error distribution results.

Table 4: Performance comparison of SA-K-means and other algorithms.

Algorithm	Clustering accuracy (%)	Response time (ms)	Data reception accuracy (%)	t-test p-value	ANOVA F-value (p-value)
SA-K-means	94.72	123.04	97.89	-	-
DBSCAN	88.43	179.34	93.67	0.04	10.34 (p< 0.01)
MS	90.56	158.97	95.21	0.03	12.56 (p< 0.01)
MLP	85.34	208.65	91.72	0.01	14.22 (p< 0.01)

Table 5: Scalability test results of SA-K-means algorithm.

Algorithm	Clustering accuracy (%)	Response time (ms)	Memory usage (MB)
SA-K-means	96.12	295.23	450.56
DBSCAN	90.35	230.12	298.47
MS	92.53	350.21	312.98
MLP	89.74	305.87	520.76

using t-test, which showed that SA-K-means significantly outperformed the other three algorithms in terms of clustering accuracy and response time (p< 0.05). In addition, for the performance differences of multiple algorithms, one-way analysis of variance (ANOVA) was applied to verify the significant differences between different algorithms in terms of clustering effect, data reception accuracy, and other indicators (p< 0.01). The specific results are shown in Table 4.

On this basis, the study further analyzed the computational complexity of the proposed algorithm in comparison with the benchmark method. The main computational steps of the K-means algorithm included the calculation of the distances between the data points and the clustering centers and the updating of the clustering centers. For each data point, the distance to each cluster centre was calculated, and assuming that there are n data points and k cluster centers, the computational complexity of K-means is O(n×k×t). Where, t is the number of iterations of the algorithm. While SA-K-means algorithm introduced SA for global optimization on the basis of K-means, its computational complexity can be O(n×k×tSA×m). Where, tSA is the number of iterations of SA and m denotes the factor of each simulated annealing. It can be seen that the computational complexity of the SA-K-means algorithm was higher than that of the traditional K-means, but its global optimization capability gave it a clear advantage in clustering effect. With the increase of data volume, the computational burden of SA-K-means was mainly reflected in the optimization process of simulated annealing, so the algorithm parameters needed to be reasonably adjusted according to the data size in practical applications to balance the computational efficiency and clustering accuracy. Meanwhile, the study further analyzed the scalability of SA-K-means in large-scale IoT data, which was tested on a dataset containing 100,000 data points. The specific results are shown in Table 5.

From Table 5, the response time of each algorithm increased significantly when the number of data extended bit 100000. Comparing the clustering effect of the four algorithms, the clustering accuracy of SA-K-means was 96.12%, which was significantly better than the other three algorithms. In terms of memory usage, SA-K-means memory usage was higher compared to DBSCAN and MS, which might be due to the fact that SA-K-means required a larger memory space for SA optimization. Overall, SA-K-means performed well in terms of clustering accuracy and it provided optimal clustering results when dealing with complex IoT service data.

4.2 Performance analysis of SA-K-means IoT service model

After verifying the performance superiority of the SA-K-means, the SA-K-means IoT service model was analyzed

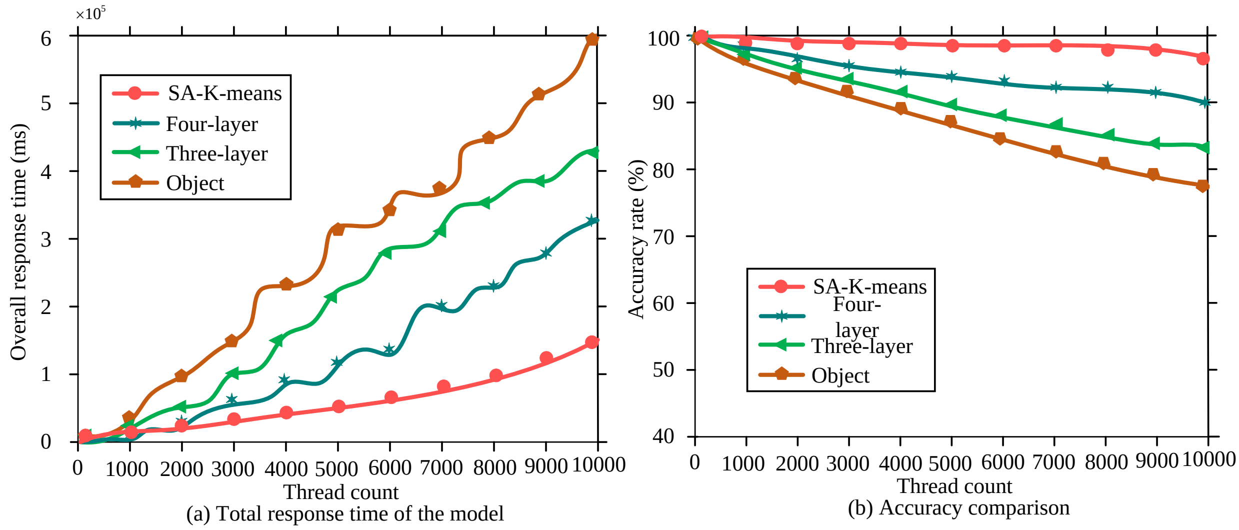


Figure 10: Comparison of response time and accuracy of the model.

for performance, and three existing IoT models, namely the Three-layer model, Four-layer model, and Object model, were selected for comparative experiments. Further simulation experiments were conducted using Microsoft's Microsoft Azure dataset in MATLAB to compare the service response time and accuracy of data reception, to assess the matching efficiency of the IoT service model. The results are shown in Figure 10.

In Figure 10 (a), when the concurrency was 2000, the response time of the SA-K-means model was almost the same as that of the Four-layer model. When the concurrency exceeded 3000, the response time of the SA-K-means model was significantly lower than that of the other three models. When the concurrency reached 10000, the total response time of the SA-K-means, Four-layer, Three-layer, and Object models, sorted from short to long, was 150943ms, 331524ms, 424635ms, and 600000ms, respectively. As shown in Figure 10 (b), when the concurrency was 1000, there was no significant difference in the accuracy of the four models. However, as the concurrency exceeded 3000, with the increase in concurrency, except for the SA-K-means model, the accuracy of data reception for the other three models decreased significantly. When the concurrency reached 10000, the data reception accuracy of the SA-K-means, Four-layer, Three-layer, and Object models, ranked from high to low, was 97.8%, 90.1%, 83.4%, and 78.8%, respectively. The experiment outcomes indicated that the SA-K-means model had the fastest processing speed for user messages, greatly reducing the response time of the IoT system. Its data reception accuracy was high, reducing the occurrence of system data loss. To further confirm the service composition effect and convergence of the SA-K-means model, four models were trained at different service composition degrees. Among them, 5200 represented the scenario with 5 service types and 200 candidate service instances, 5400 represented the scenario with 5 service types and 400 candidate service instances, and the Inverse Generative Distance (IGD) was a comprehensive performance evaluation index used to evaluate the convergence and diversity of the model. The frontier distribution and median values of IGD are shown in Figure 11.

According to Figure 11 (a), in the Pareto front graph with a complexity of 5200, the Object model had the worst convergence among the four models, while the SA-K-means model had the closest scatter distribution to the Pareto optimal front. As shown in Figure 11(b), even with a complexity of 5400, the SA-K-means model was still closest to the Pareto optimal frontier. As shown in Figure 11(c), in the scenario with a complexity of 5200, the SA-K-means model had the smallest upper and lower limits of IGD values, the smallest overall IGD value, and a median IGD value of 0.0061. As shown in Figure 11 (d), in the scenario with a complexity of 5400, the SA-K-means model still had the lowest IGD lower limit, with a median IGD value of 0.0065. It can also be seen from Figure 11 that the SA-K-means algorithm maintained superior convergence speed and matching accuracy as the service complexity increased, which was crucial for handling large-scale and high-complexity IoT services. In IoT applications, service combination capability directly affected user experience and system efficiency. The real-time and accuracy of services were crucial for optimizing production and ensuring security.SA-K-means improved the response speed and success rate of IoT services by optimizing service matching.

To further confirm the overall performance superiority of the SA-K-means model, this study conducted comparative experiments with the Four-layer, Three-layer, and Object models. Additionally, the study selected the Platform as a Service (PaaS) model and the Infrastructure as a Service (IaaS) model, and compared them with simulation experiments. In the scenario with a service quantity of 500, the comparison indicators included service quantity, service success rate, response time, reliability, and availability. The outcomes are in Table 6.

According to Table 6, the SA-K-means model had a service count of 485.6 and a service success rate of

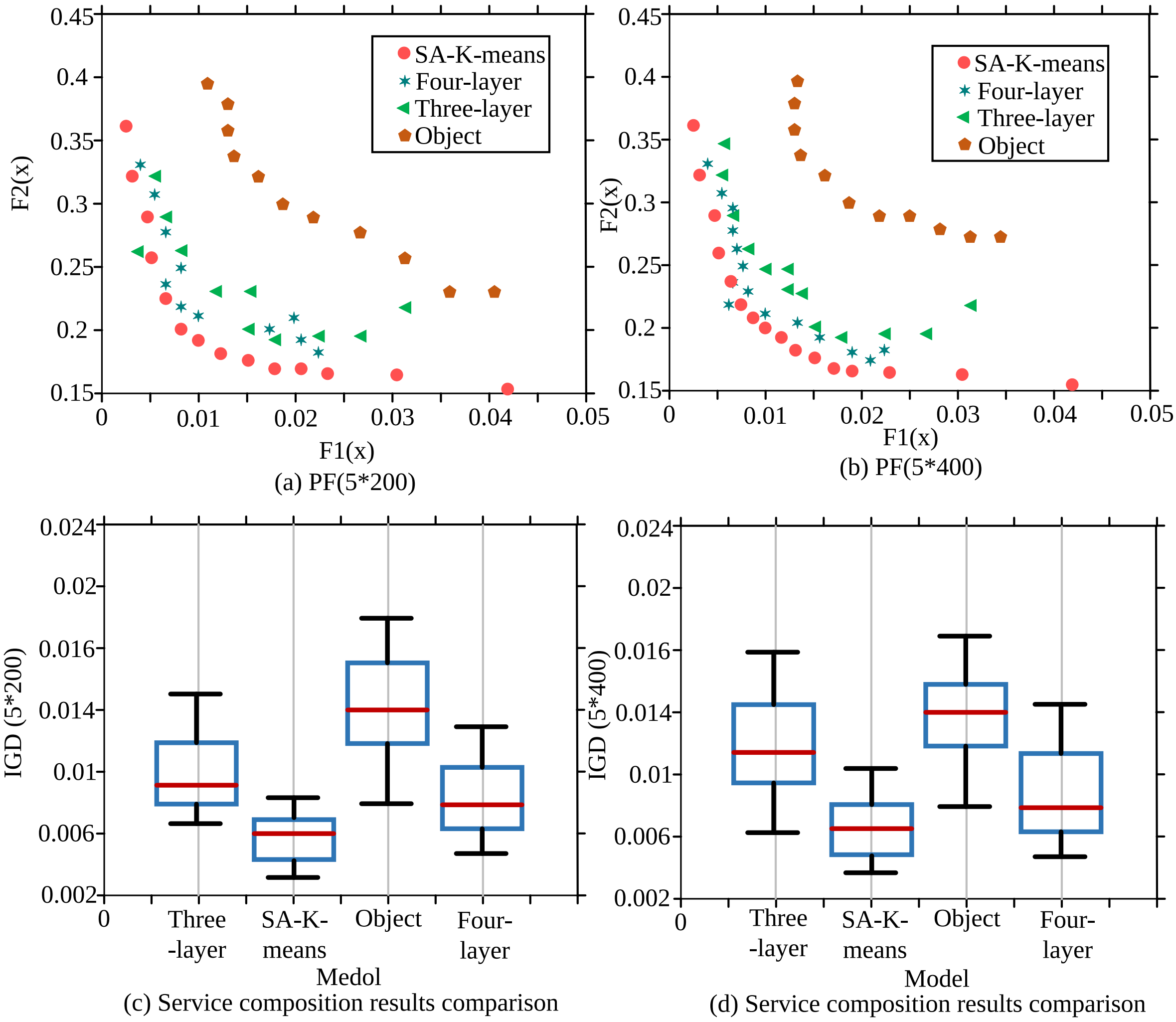


Figure 11: Convergence versus service composition results.

Table 6: Model index comparison.

Model	Number of services	Success rate	Response time	Reliability	Availability
SA-K-means	485.6	95.12%	1276.18 ms	96.75%	96.91%
Four-layer	450.1	87.02%	1475.54 ms	87.61%	88.49%
IaaS	423.6	84.72%	1863.47 ms	85.03%	96.47%
Three-layer	381.7	74.34%	2245.69 ms	75.42%	75.98%
Paas	356.2	65.24%	2795.18 ms	66.75%	67.43%
Object	327.2	60.44%	3152.72 ms	61.46%	64.53%

95.12%, both higher than other models. The total response time of the SA-K-means model was 1276.18 ms, which was the lowest among the comparison models. Compared with the Object model, which took the most time, the SA-K-means model had shortened the response time by 1876.54 ms. The reliability and availability of the SA-K-means model were 96.75% and 96.91%, respectively, both higher than the comparison models. Availability represented the percentage of the model's normal running time to the total time, while reliability represented the percentage of successful service matching. The higher the reliability, the fewer interruptions or failures occurred. The experiment outcomes indicated that the SA-K-means model had the fastest service matching speed and the highest success rate, could search for optimization as quickly as possible, and exhibited superior model performance. It could effectively provide reliable service content for users even in situations with large service volumes.

5 Discussion

The study proposed SA-K-means, a K-means clustering method optimized by combining SA algorithm, and verified its performance in IoT service matching. The experimental results showed that SA-K-means performed well in terms of clustering accuracy, response time, and data reception accuracy. Firstly, SA-K-means showed significant advantages in comparison with existing clustering techniques (K-means, DBSCAN, MS, and MLP). In terms of clustering accuracy, SA-K-means achieved 94.72%, which was a significant increase over the other three algorithms. This result indicated that SA-

K-means had higher accuracy in identifying and classifying data categories. In addition, SA-K-means showed good dynamic adaptability through the global optimization ability of the SA algorithm, which was able to converge quickly and adapt to scenarios of different complexity. These advantages made SA-K-means highly applicable in IoT service matching.

Secondly, the study used MATLAB as the main simulation tool, whose powerful matrix operations and data visualization functions facilitated the implementation and testing of the algorithm. However, MATLAB had limitations in terms of computational efficiency and resource consumption when dealing with large-scale data. In practical deployment, it was recommended to migrate the algorithms implemented in MATLAB to more efficient programming languages such as C++ or Python and optimize them in combination with high-performance hardware resources. In addition, considering the real-time requirements of real IoT systems, the SA-K-means model may need to have efficient real-time processing capabilities. By combining with a streaming data processing framework, fast response and analysis of real-time data could be achieved. Whereas high-performance computing and storage resources were the basis for processing large-scale data, the elastic scalability of cloud platforms provided flexibility and reliability to the system. Data security and privacy protection were important considerations for IoT systems, which must be safeguarded by encryption techniques and access control mechanisms.

Although SA-K-means performed well in terms of clustering accuracy, response time, and data reception accuracy, its performance was sensitive to the setting of hyper-parameters. In addition, in a large-scale real-time IoT environment, SA-K-means may face the problems of insufficient computational resources and response delay. It also faced challenges in terms of computational efficiency and memory consumption as the data volume grew. Future research will explore automated hyper-parameter tuning methods, introduce incremental learning and online learning mechanisms, as well as utilize distributed computing and parallel processing techniques, in order to address the above issues and further enhance the performance and application value of the SA-K-means algorithm in large-scale real-time IoT systems.

6 Conclusion

Aiming at the problem that the traditional K-means algorithm is easy to fall into local optimality and slow convergence when dealing with large-scale data, a SA-K-means clustering method combined with SA algorithm was proposed. The limitations of traditional K-means algorithm were avoided by simulating the global optimization ability of annealing algorithm. The experimental results showed that SA-K-means had an average increase of 28.13% in Silhouette score over the other three algorithms and its ARI was as high as 0.93. In the case of high concurrency, SA-K-means had a response time of 1276.18ms, which was an average increase in reliability of the other clustering algorithms by 28.57%.

Overall, the SA-K-means algorithm not only performed well in IoT service matching, but its optimized clustering capability and efficient data processing performance made it a potential for a wide range of applications in several other domains. In healthcare, SA-K-means could be used for patient data classification, medical resource optimization and disease prediction and prevention. In the field of smart cities, the algorithm could be applied to traffic flow management, energy management and environmental monitoring to optimize the efficiency of urban operation and improve the quality of life of residents. Future research will be devoted to further optimizing the algorithm and exploring its possibilities in multi-modal data fusion and cross-domain applications, in order to enhance its application value and performance in different fields.

References

[1] Ryan Wen Liu, Maohan Liang, Jiangtian Nie, Wei Yang Bryan Lim, Yang Zhang, and Mohsen Guizani. Deep learning-powered vessel trajectory prediction for improving smart traffic services in maritime Internet of Things. IEEE Transactions on Network Science and Engineering, 9(5):3080-3094, 2022. https://doi.org/10.1109/TNSE.2022.3140529

[2] Hoa Tran-Dang, Nicolas Krommenacker, Patrick Charpentier, and Dong-Seong Kim. The Internet of Things for logistics: Perspectives, application review, and challenges. IETE Technical Review, 39(1):93-121, 2022. https://doi.org/10.1080/02564602.2020.1827308

[3] Suaad Hadi Hassan Al-Taai, Huda Abbas Kanber, and Waleed Abood Mohammed al-Dulaimi. The importance of using the internet of things in education. International Journal of Emerging Technologies in Learning, 18(1):19-39, 2023. https://doi.org/10.3991/ijet.v18i01.35999

[4] Hassan Ismkhan, and Mohammad Izadi. K-means-G*: Accelerating k-means clustering algorithm utilizing primitive geometric concepts. Information Sciences, 618(5):298-316, 2022. https://doi.org/10.1016/j.ins.2022.11.001

[5] Tengjiao Guo, Guosheng Li, and Lei He. Risk assessment of typhoon storm surge based on a simulated annealing algorithm and the least squares method: A case study in Guangdong Province, China. Natural Hazards Research, 2(3):249-258, 2022. https://doi.org/10.1016/j.nhres.2022.08.005

[6] Ahmed Ahmim, Faiz Maazouzi, Marwa Ahmim, Sarra Namane, and Imed Ben Dhaou. Distributed denial of service attack detection for the Internet of Things using hybrid deep learning model. IEEE Access, 11(5):119862-119875, 2023. https://doi.org/10.1109/ACCESS.2023.3327620

[7] Dharun Kasilingam, and Rajneesh Krishna. Understanding the adoption and willingness to pay for internet of things services. International Journal

of Consumer Studies, 46(1):102-131, 2022. https://doi.org/10.1111/ijcs.12648
[8] Abdullah Ayub Khan, Asif Ali Laghari, Zaffar Ahmed Shaikh, Zdzislawa Dacko-Pikiewicz, and Sebastian Kot. Internet of Things (IoT) security with blockchain technology: A state-of-the-art review. IEEE Access, 10(3):122679-122695, 2022. https://doi.org/10.1109/ACCESS.2022.3223370
[9] Denis Ushakov, Egor Dudukalov, Elena Kozlova, and Khodor Shatila. The Internet of Things impact on smart public transportation. Transportation Research Procedia, 63(4):2392-2400, 2022. https://doi.org/10.1016/j.trpro.2022.06.275
[10] Lijun Wei, Yuhan Yang, Jing Wu, Chengnian Long, and Bo Li. Trust management for Internet of Things: A comprehensive study. IEEE Internet of Things Journal, 9(10):7664-7679, 2022. https://doi.org/10.1109/JIOT.2021.3139989
[11] Miin-Shen Yang, and Ishtiaq Hussain. Unsupervised multi-view K-means clustering algorithm. IEEE Access, 11(4):13574-13593, 2023. https://doi.org/10.1109/ACCESS.2023.3243133
[12] Abiodun M. Ikotun, Absalom E. Ezugwu, Laith Abualigah, Belal Abuhaija, and Jia Heming. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. Information Sciences, 622(3):178-210, 2023. https://doi.org/10.1016/j.ins.2022.11.139
[13] Md. Zubair, MD. Asif Iqbal, Avijeet Shil, M. J. M. Chowdhury, Mohammad Ali Moni, and Iqbal H. Sarker. An improved K-means clustering algorithm towards an efficient data-driven modeling. Annals of Data Science, 11(5):1525-1544, 2024. https://doi.org/10.1007/s40745-022-00428-2
[14] Kamran Zolfi, and Javid Jouzdani. A mathematical model and a simulated annealing algorithm for unequal multi-floor dynamic facility layout problem based on flexible Bay structure with elevator consideration. Journal of Facilities Management, 21(3):352-386, 2023. https://doi.org/10.1108/JFM-09-2021-0102
[15] Kun Shi, Zhengtian Wu, Baoping Jiang, and Hamid Reza Karimi. Dynamic path planning of mobile robot based on improved simulated annealing algorithm. Journal of the Franklin Institute, 360(6):4378-4398, 2023. https://doi.org/10.1016/j.jfranklin.2023.01.033
[16] Dalila B.M.M. Fontes, S. Mahdi Homayouni, and José F. Gonçalves. A hybrid particle swarm optimization and simulated annealing algorithm for the job shop scheduling problem with transport resources. European Journal of Operational Research, 306(3):1140-1157, 2023. https://doi.org/10.1016/j.ejor.2022.09.006
[17] Jiaxin Liu, and Chanjuan Han. Design and optimization of heat extraction section in energy tunnel using simulated annealing algorithm. Renewable Energy, 213(5):218-232, 2023. https://doi.org/10.1016/j.renene.2023.05.135
[18] Walid K. Ghamry, and Suzan Shukry. Multi-objective intelligent clustering routing schema for internet of things enabled wireless sensor networks using deep reinforcement learning. Cluster Computing, 27(4):4941-4961, 2024. https://doi.org/10.1007/s10586-023-04218-0.
[19] Aihua He. Study on internet of things anomaly data mining method based on improved differential evolution automatic clustering. International Journal of Intelligent Systems Technologies and Applications, 23(1-2):51-71, 2025. https://doi.org/10.1504/IJISTA.2025.145616
[20] Nawaf H.M.M. Shrifan, Muhammad F. Akbar, and Nor Ashidi Mat Isa. An adaptive outlier removal aided k-means clustering algorithm. Journal of King Saud University-Computer and Information Sciences, 34(8):6365-6376, 2022. https://doi.org/10.1016/j.jksuci.2021.07.003
[21] Zhihua Chen. Research on internet security situation awareness prediction technology based on improved RBF neural network algorithm. Journal of Computational and Cognitive Engineering, 1(3):103-108, 2022. https://doi.org/10.47852/bonviewJCCE149145205514
[22] B. Nagajayanthi. Decades of Internet of Things towards twenty-first century: A research-based introspective. Wireless Personal Communications, 123(4):3661-3697, 2022. https://doi.org/10.1007/s11277-021-09308-z
[23] Ramon Sanchez-Iborra, Abdeljalil Zoubir, Abderahmane Hamdouchi, Ali Idri, and Antonio Skarmeta. Intelligent and efficient IoT through the cooperation of TinyML and edge computing. Informatica, 34(1):147-168, 2023. https://doi.org/10.15388/22-INFOR505