Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 4, pp. 2889–2902, DOI 10.12694/scpe.v25i4.2859

INTELLIGENT MATCHING METHOD FOR COLLEGE DORMITORY ROOMMATES: CHAMELEON ALGORITHM BASED ON OPTIMIZED PARTITIONING

CUIPING WANG*

Abstract. A chameleon algorithm based on optimized partitioning was studied to solve the intelligent matching problem of college dormitory roommates. Using quantitative research methods, data on personal preferences and lifestyle habits of college students were collected, and the K-center object chameleon algorithm was used for data analysis and roommate matching. Test the algorithm performance on the BBC dataset, compare clustering quality indicators such as entropy, purity, and RI value, and verify the effectiveness of the algorithm. This algorithm can accurately assign students to their respective dormitories, avoiding overlapping situations and achieving excellent matching results. In terms of matching accuracy and running time, the K-center object chameleon algorithm shows superior performance compared to other algorithms. In terms of clustering quality evaluation, comparisons were made from three dimensions: entropy value, purity, and RI value. The experimental results show that the closer the entropy value is to 0, the closer the purity and RI value are to 1, and the better the matching effect. This result further validates the effectiveness of the algorithm in the intelligent matching problem of college dormitory roommates. The matching accuracy of this algorithm on the BBC dataset reached 98.82%, showing better clustering quality than other algorithms in terms of entropy, purity, and RI values. The entropy value approached 0, while the purity and RI values approached 1, verifying the efficiency of matching quality. The chameleon algorithm based on optimized partitioning proposed in the study has shown excellent performance in intelligent matching of college dormitory roommates, with the characteristics of high-precision matching and fast running time. It has important practical significance for improving the quality of life and learning efficiency of college dormitory students, and provides new research methods and ideas for related fields.

Key words: Dormitory roommates; Intelligent matching; Optimize partitioning methods; Chameleon algorithm; Big data

1. Introduction. The process of roommate matching in university environments is closely related to the quality of life, physical and mental health, learning experience, and social development of students. An effective roommate matching system not only affects daily living conditions, but also affects the welfare of students [1-2]. Simple surveys and manual coordination are existing methods for roommate matching, but due to the complexity and subtle differences in personal preferences and personalities, these methods do not meet the requirements. In addition, these methods lack the dynamic adaptability required to adapt to diverse and constantly changing student situations. Therefore, these traditional methods are prone to lifestyle and interpersonal conflicts when achieving harmonious living arrangements. The problem with roommate matching lies in accurately assessing and integrating various factors, such as habits, study plans, rest patterns, and social tendencies. Insufficient consideration of these factors can lead to disharmony and dissatisfaction, which in turn can harm students' academic and social pursuits [3-4]. The new intelligent matching method draws on the characteristics of biological chameleons to achieve optimal dormitory roommate matching, with adaptability and flexibility. This chameleon algorithm is based on optimization partitioning, inputting students' personality characteristics and interests, and using machine learning and data mining techniques for intelligent analysis and matching [5-6]. In view of this, the study introduces a chameleon algorithm based on optimized partitioning for roommate matching, to improve the quality of life in college students' dormitories. It is expected to create a harmonious and friendly dormitory environment, optimizing the learning and living environment. The research will be divided into four parts. This algorithm has been validated in over 10 different regions and types of Chinese higher education institutions. The research is mainly applied to the higher education environment in China, and the research subjects are selected from 1000 mainland Chinese university students. Using stratified random sampling to obtain students from different grades, majors, genders, and cultural backgrounds. Firstly, divide universities into first tier, second tier, and junior colleges, and then randomly select them based on their

*School of Artificial Intelligence and Big Data, Zibo Vocational Institute, Zibo, 255000, China (wcp197841@126.com).

majors and gender ratios. The research contribution lies in the development of roommate matching algorithms that effectively improve the quality of life and learning experience of college students by taking into account various factors such as personal habits, study plans, rest patterns, and social tendencies. The first part is an overview of the chameleon algorithm based on optimized partitioning for the intelligent matching method of college dormitory roommates. The second part is the study of the chameleon algorithm based on optimized partitioning for the intelligent matching method of college dormitory roommates. The third part is the second part's experimental verification. The fourth part is a summary of the research content and points out the shortcomings.

2. Related Works. In the field of dormitory roommate matching research, a large amount of research has focused on traditional methods such as large-scale grouping and random allocation. J é rme et al. studied the fully distributed channel allocation problem in clustered wireless networks. The trial-and-error framework was expanded to prove that its direct application in random environments does not provide ideal solutions. A robust trial and error learning algorithm was proposed to restore good convergence by introducing thresholds. The random effects under Rayleigh fading were analyzed and the theoretical claims was confirmed. Online algorithms were developed to dynamically estimate the optimal threshold and adapt to instantaneous interference [7]. Zucker et al. derived the optimal power allocation for nodes and the optimal rate allocation for networks. A suboptimal scheme was proposed, with threshold scheduling for power allocation and linear adaptive scheme for rate allocation. The simulation outcomes denoted that the loss of the suboptimal scheme was less than 1.5% [8]. Kiesnetter et al. found that plants benefit from microbial composition, but changes in microbial composition also greatly affect plant performance and resource allocation. The random forest model identified fungal and bacterial families that are important for plant performance, and they significantly changed with fragmentation. Research supported the significant fragmentation effect of natural microbial communities and confirmed that fragmentation related changes in microbial communities can affect natural plant performance and investment [9]. Bistritz I's research team proposed a non cooperative energy allocation game and adopted the optimal reaction dynamics as a distributed algorithm. The results showed that the algorithm converged to a pure Nash equilibrium with a maximum of N steps within no more than N steps. However, these equilibria might be inefficient. The proposed algorithm avoided inefficient equilibrium and achieved asymptotic optimal performance in almost all games [10]. Lai et al. derived the optimal power allocation for nodes and the optimal rate allocation for networks. A suboptimal solution was proposed: threshold scheduling was used for power allocation, and linear adaptive scheme was used for rate allocation. The simulation findings indicated that the loss of the suboptimal scheme was less than 1.5% [11]. Henning Smith et al. proposed that with the aging population, barriers to accessing medical services in rural areas are particularly prominent, and it is necessary to evaluate the healthcare access of medical insurance beneficiaries in rural areas in order to develop effective strategies. The research results indicate that after adjusting for factors such as age, gender, and marital status, rural beneficiaries have significantly lower satisfaction with home visits and access to expert care compared to urban beneficiaries, and are more inclined to avoid seeking medical treatment or not informing others when sick. These differences indicate the quality and acceptability issues faced by rural beneficiaries in healthcare, which may affect their seeking behavior and ultimate health outcomes [12].

However, these methods fail to fully consider students' personalized needs and interests, often leading to tense relationships and frequent conflicts in the dormitory. The chameleon algorithm based on optimized partitioning provides a new solution for intelligent matching of college dormitory roommates. Liang Z et al. proposed an automatic clustering method that has clustering stability and can automatically detect correct structures. The experimental results showed that this method had better robustness in synthetic and real data, and was robust in parameter selection [13]. Li et al. put forward a new dynamic verifiable retrieval scheme for encrypted images to address the problems of low retrieval accuracy, low efficiency, and low validation efficiency in dynamic environments that exist in existing schemes. This scheme extracted image features by using a pre trained convolutional neural network model, designed a K-means clustering algorithm-based encrypted index, and used a chameleon hash-based dynamic verification tree to improve retrieval accuracy, efficiency, and dynamic verification efficiency [14]. Alexiadis et al. utilized a machine learning model based on the RASA framework for NLU and conversation flow classification. CIPA-Generation B can generate plans and execute them based on user requests, and has been deployed for testing in real-world scenarios in the energy and health fields. Experimental research has shown that the developed dialogue agent system has usefulness and acceptability in multi-intention recognition and dialogue processing [15]. Nie et al. proposed a personal carbon trading (PCT) subsidy policy. Based on the Stackelberg game model and actual data from Beijing, research has found that PCT is more efficient than purchasing subsidies and can achieve higher CO2 emission reduction effects. The achievements would help promote the application of PCT subsidy mechanism and assist the transportation industry in achieving carbon peak and carbon neutrality goals [16]. Feng et al. proposed a three-step mathematical programming method to alleviate the problem of excess or insufficient passenger stations on PTU, combining discrete wavelet transform and artificial neural network to predict demand, minimizing the weighted total travel cost and unmet user needs. The research results showed that the proposed demand prediction and optimization method was significantly superior to traditional methods and had robustness [17]. Zhu X et al. proposed the impact of interpersonal relationships in dormitories on emergency evacuation behavior among college students. The research results indicate that close social connections within dormitories promote the efficiency of emergency evacuation; In emergency situations, these relationships are partially transformed into a leader follower mode, reducing evacuation efficiency. In addition, research has found that gender and grade are important factors affecting evacuation behavior, and female and lower grade students have relatively higher evacuation efficiency [18].

In summary, the chameleon algorithm based on optimized partitioning provides a new perspective on the matching problem of college dormitory roommates, improving the accuracy and efficiency of matching. However, the practical application effectiveness and stability of this algorithm still need further verification and exploration. At the same time, how to apply this theory to the actual allocation of university dormitories is also a problem that needs to be solved. Despite some challenges, it is still hoped that the chameleon algorithm can bring more harmony and comfort to the dormitory life of college students in future applications, further improving their quality of life and learning efficiency.

3. Materials and methods. Firstly, it introduces the chameleon algorithm based on optimized partitioning. Then, the original chameleon algorithm is optimized and an optimized chameleon algorithm based on optimized partitioning is put forward to improve the accuracy and efficiency of matching. Finally, a model of intelligent matching method for college dormitory roommates incorporating optimized chameleon algorithm is established, further improving the application framework of this matching method. It is hoped to provide new basis for intelligent matching of college dormitory roommates and open up new perspectives for research in related fields.

3.1. Fusion method of chameleon and K-medoids algorithms. The chameleon algorithm based on optimized partitioning reflects data objects and their similarities through sparse graphs, ensuring effective expansion of large datasets. The algorithm combines preliminary data segmentation of the graph with condensed hierarchical clustering to form clusters of different shapes, sizes, and densities. Firstly, it constructs a sparse graph and divides it into sub clusters of approximate size, and then they are merged based on the proximity and interconnectivity of the clusters. Simultaneously, it needs to calculate the real-time similarity function between clusters to guarantee that it does not exceed the threshold. The chameleon algorithm process is shown in Figure 3.1.

In Figure 3.1, the chameleon clustering algorithm can be divided into clusters of different shapes, sizes, and densities, but the effects of noise and outliers cannot be eliminated. The K-medoids clustering algorithm, as an improved version of the K-means algorithm, is based on partitioning methods for clustering. It randomly selects a set of data objects as the initial center points, and assigns each data object to the cluster represented by the closest center point. Then, it will continue to iterate and determine the new center point of each cluster until the clustering quality no longer improves. This method can process noise and outliers in stages in the graph to avoid affecting the clustering effect. The research chose the Chameleon algorithm because it can handle large datasets and effectively form clusters of different shapes, sizes, and densities. It can adaptively merge clusters based on proximity and interconnectivity, making it suitable for complex data structures often encountered in the real world. The K-means algorithm usually uses Euclidean distance to calculate the distance between data objects, and the calculation method is shown in equation (3.1).

$$D = \sqrt{(x_1 - y_1)^2 - (x_2 - y_2)^2} \tag{3.1}$$

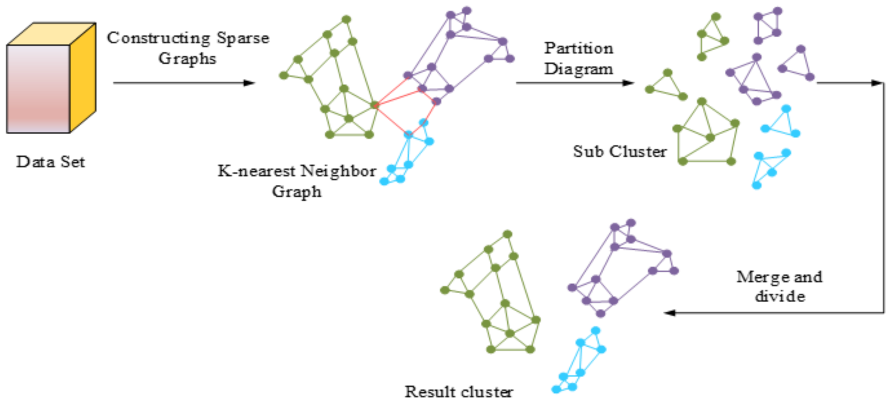


Fig. 3.1: Chameleon algorithm process

In equation (3.1), $(x_1 - y_1)$ and $(x_2 - y_2)$ are the coordinates of two points in the class cluster, and D is the Euclidean distance value. In cluster analysis, Euclidean distance measures the similarity of data points, and the closer the distance, the higher the similarity. By calculating the contour coefficient to evaluate the clustering effect, it reflects the compactness within the cluster and is an external indicator, as defined in equation (3.2).

$$S_i = \frac{b_i - a_i}{\max\left(a_i, b_i\right)} \tag{3.2}$$

In equation (3.2), a_i means the average distance between each data object in the cluster and other objects in the cluster, and b_i refers to the mini average distance between the data object in the cluster and all other clusters. The improved chameleon algorithm first uses K-Means to partition the initial dataset into subcategories with high cohesion, and then performs K-nearest neighbor graph partitioning to reduce the impact of outliers and noise. The aggregation algorithm of the chameleon algorithm is used to analyze small clusters and obtain the final clustering results. On the basis of not changing the chameleon algorithm process, the K-medoids algorithm is inserted to reduce the impact of outliers, improve intra-cluster similarity and inter-cluster difference, thereby improving algorithm efficiency and results. The chameleon algorithm based on optimized partitioning also adopts a similar approach, but further optimization and improvement have been made. The flowchart of the chameleon algorithm based on optimized partitioning is denoted in Figure 3.2.

In Figure 3.2, the K-medoids algorithm requires repeated iterations to determine the center point, resulting in an increase in clustering convergence time and limited processing ability for large datasets, making it more suitable for clustering analysis of small datasets. However, its sensitivity to abnormal data endows it with good anti-interference ability at outliers. In the chameleon algorithm based on optimized partitioning, these issues have been improved and optimized to handle large datasets more effectively while maintaining anti-interference ability against abnormal data.

3.2. Detailed explanation and process research of optimizing chameleon algorithm based on K-medoids. The optimized chameleon algorithm introduces the K-medoids clustering algorithm to replace the original graph partitioning algorithm, resulting in changes in all variables during the process of building a dynamic model. Due to changes in the weighted edges between nodes in the K-nearest neighbor graph obtained after passivation of the dataset, the calculation methods and requirements for the internal proximity and interconnection of clusters, as well as the relative proximity and interconnection between clusters, need to be redefined [19]. The schematic diagram of the algorithm clustering process is shown in Figure 3.3.

Firstly, it is necessary to determine k clustering centers and divide sub samples based on the determined center points, and continuously optimize the center points during the iteration process. Improve the K-means algorithm in terms of its ability to handle noise and outliers. It selects actual data points as the clustering

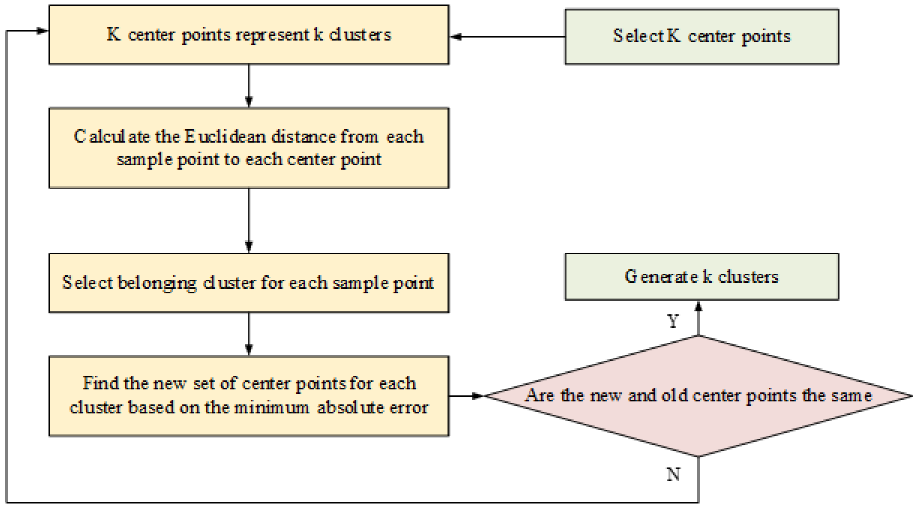


Fig. 3.2: Flow chart of chameleon algorithm based on optimized partitioning method

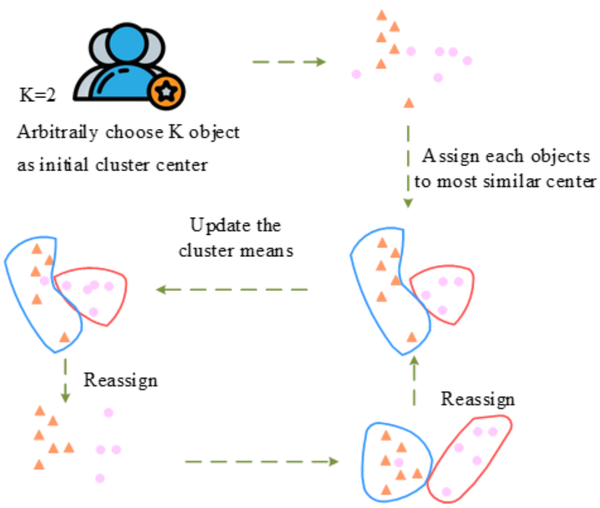


Fig. 3.3: Schematic diagram of algorithm clustering process

center instead of the average value of points in the cluster (such as k-means), providing robustness to outliers. The set of sub samples divided is represented as $(B_1, B_2, \ldots B_k)$, and the purpose of the K-medoids algorithm is to minimize the value of equation (3.3).

$$E = \sum_{i=1}^{k} \sum_{x \in B_i} \|x - \mu_i\|_2^2 \tag{3.3}$$

In equation (3.3), μ_i is the centroid of the ith subsample set B_i. In the K-means algorithm, the input sample is $A = \{A_1, A_2, \ldots, A_m\}$, and any k samples are obtained as the initial centroid, represented as $\{\mu_1, \mu_2, \ldots, \mu_k\}$. In a certain subset, the distance between the ith sample x_i and the jth centroid μ_j are calculated using equation

(3.4).

$$d_{ij} = \|x_i - \mu_j\|^2 \tag{3.4}$$

If μ_j can minimize the d_{ij} distance, it will choose μ_j as the new clustering center and update the sub samples. The K-medoids algorithm is used to optimize the chameleon algorithm to improve text feature clustering performance. However, due to the increase in data volume and random selection of initial clustering centers, this algorithm faces efficiency and accuracy challenges when processing large-scale high-dimensional data. The research divides data points into static and dynamic groups: static groups maintain existing clustering structures, while dynamic groups explore new clustering structures. This strategy improves the stability and flexibility of clustering algorithms, achieving higher clustering efficiency and accuracy when processing large-scale high-dimensional data. The separation behavior S_i, alignment behavior A_i, aggregation behavior C_i, foraging behavior F_i, and enemy avoidance behavior E_i of this algorithm can be expressed as equation (3.5).

$$\begin{cases} S_i = -\sum_{j=1}^{n} X - X_j \\ A_i = \frac{\sum_{j=1}^{n} V_j}{n} \\ C_i = \frac{\sum_{j=1}^{n} X_j}{n} - X \\ F_i = X^+ - X \\ E_i = X^- + X \end{cases} \tag{3.5}$$

In equation (3.5), X represents the position of the individual. X_j is the position of the j adjacent individual. n is the number of adjacent individuals. V_j is the velocity of the j adjacent individual. X^+ is the location of food. X^- represents the location of the natural enemy. After continuing the iteration, the step vector $\triangle X_{t+1}$ and position vector X_{t+1} of the individual can be calculated according to equation (3.6).

$$\begin{cases} \triangle X_{t+1} = (sS_i + aA_i + cC_i + fF_i + eE_i) + w\triangle X_t \\ X_{t+1} = X_t + \triangle X_{t+1} \end{cases} \tag{3.6}$$

In equation (3.6), s, a, c, f, e represent the weights of individual separation behavior, alignment behavior, aggregation behavior, foraging behavior, and enemy avoidance behavior, respectively. w is the inertia weight. Through continuous iteration, the optimal solution is found, which is the optimal initial clustering center of the K-means algorithm, thereby improving the clustering effect. However, the initial population diversity of DA is poor, and its global optimization ability is weak. The final solution obtained from its solution may not be the optimal solution, which affects the optimization effect of the K-means algorithm. If the dimension of each individual is d, during population initialization, an d dimensional individual X_{id} is generated, and then the individual is mapped to another space to obtain a new individual CX_{id}, as shown in equation (3.7).

$$\begin{cases} X_{id} = X_{\min d} + rand\,(X_{\max d} - X_{\min d}) \\ CX_{id} = X_{\min d} + X_{\max d} - X_{id} \end{cases} \tag{3.7}$$

In equation (3.7), $X_{\max d}, X_{\min d}$ are the upper and lower limit of the value, respectively. It merges the mapped population with the original population, and then selects the population with a better fitness value of 1/2 as the initial population to complete the optimization of the initial population. The inertia weight w of the algorithm generally adopts a linear reduction method, and the convergence speed is also slowly decreasing, as shown in equation (3.8).

$$w = 0.9 - t\,(0.5/t_{\max}) \tag{3.8}$$

In equation (3.8), $t_{\max}$ is the set iteration upper limit. However, the w reduction speed under this inertia weight strategy does not match the convergence speed of the algorithm, resulting in a decrease in the convergence performance of the algorithm. The study introduces an improved chameleon algorithm based on K-medoids. The main improvement is to use the K-medoids algorithm to replace the original Metis graph segmentation algorithm for sub cluster partitioning. Then, the optimized chameleon algorithm is used to dynamically merge the sub clusters. The workflow of this improvement plan is shown in Figure 3.4.

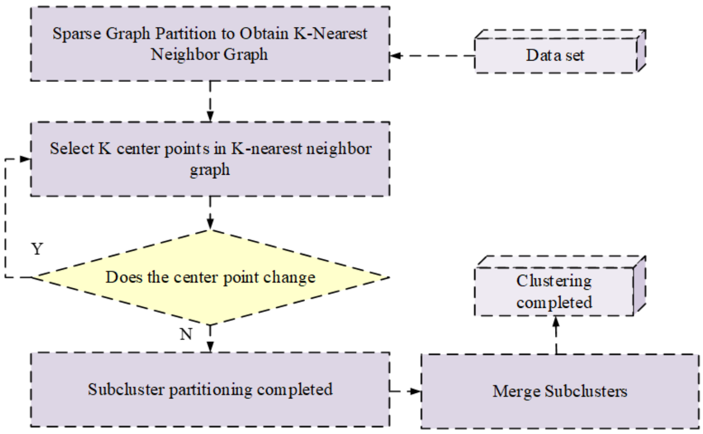


Fig. 3.4: Improved algorithm flowchart

In Figure 3.4, the algorithm first divides the dataset into K-nearest neighbor graphs, selects the initial cluster center, and divides the data into the nearest cluster. It optimizes the center point until the cluster stabilizes. After dividing the sub clusters, it needs to calculate the similarity value. If the maximum similarity value exceeds the threshold, it will merge the corresponding clusters. It will recurse to no clusters for merging, improving clustering accuracy and ensuring efficiency.

3.3. Intelligent matching model for college dormitory roommates incorporating optimized chameleon algorithm. The chameleon algorithm, which optimizes the partitioning method, exhibits advantages in processing large-scale high-dimensional data text clustering by adjusting parameters, improving similarity calculation, and adopting new clustering strategies for optimization. Data preprocessing includes cleaning and standardization, and for complex student information, data augmentation strategies such as feature crossing and encoding are adopted to improve the model's generalization ability. In the matching problem, a hierarchical matching strategy is used to enhance robustness, and possible matching is achieved through sorting to meet diverse needs [20-22]. Therefore, the label smoothing method in this article is to change the one hot label, and the specific method is shown in equation (3.9).

$$Q_i \begin{cases} 1 - \diamondsuit & ifi = true \\ \frac{\diamondsuit}{K-1} & else \end{cases} \tag{3.9}$$

In equation (3.9), $\diamondsuit$ is a predetermined hyperparameter much smaller than 1. i is the sample prediction label. K is the number of categories under that classification. It changes the format of one hot encoding, so that the optimization goal in SoftMax loss is not just 0 and 1. After label smoothing, the model output is no longer higher than better, thereby enhancing the robustness of the model. The parameters of the model are updated as indicated in equation (3.10).

$$\omega_t = \omega_{t-1} - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \diamondsuit}} \tag{3.10}$$

In equation (3.10), ω_t and ω_{t-1} are the updated parameter values. α is the updated learning rate. $\hat{m}_t$ is the exponentially weighted moving average of the gradient. $\hat{v}_t$ is the exponentially weighted average of the gradient squared. The relationship between the update equation and the learning rate and hyperparameters is shown in equation (3.11).

$$\omega_t = \omega_{t-1} - \alpha \frac{\beta_I m_{t-1} + (1 - \beta_I)\, g_t}{\sqrt{\beta_2 v_{t-1} + (1 - \beta_2)\, g_t^2}} \tag{3.11}$$

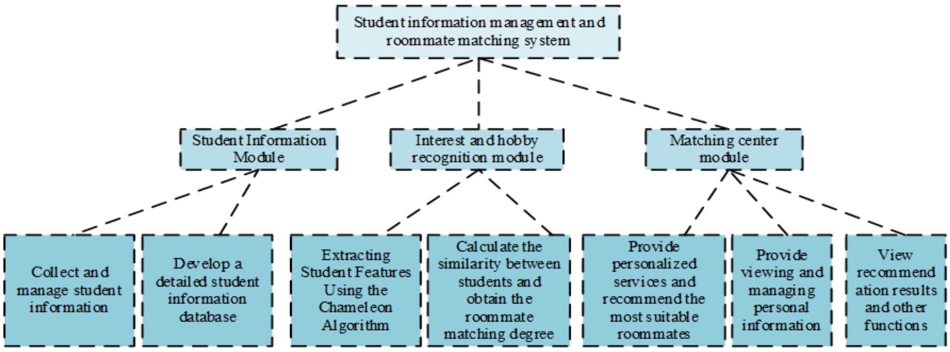


Fig. 3.5: System functional architecture diagram

In equation (3.11), β_I and β_2 are hyperparameters. To raise the accuracy and generalization ability of the chameleon algorithm in intelligent matching of college dormitory roommates, and reduce training time, the algorithm is pre trained using existing student information datasets, with training parameters as the initial parameters of the model. Then, the matching method is adjusted to adapt the algorithm to the multiple matching problem, and the appropriate similarity calculation method and clustering strategy are selected to optimize the matching effect. The intelligent matching system includes three modules: student information collection, interest and hobby recognition, and matching center. The overall system framework is shown in Figure 3.5.

In Figure 3.5, the student information module is responsible for collecting and managing student information such as interests, personality traits, etc., to form a detailed student information database. The interest and hobby recognition module utilizes the chameleon algorithm to extract student features, calculate similarity, and obtain roommate matching. The matching center module recommends the most suitable roommates and provides personalized services based on student information and recognition results. At the same time, this module also allows students to view and manage personal information and recommendation results, making it easy to use.

4. Result and Discussion. The intelligent matching of college dormitory roommates adopted the optimized chameleon algorithm, which significantly improved the matching accuracy by adjusting parameters and improving similarity calculation. By deeply mining student information and extracting features, the model's generalization ability was enhanced to adapt to different types of student groups. This method was stable and scalable, and could handle large-scale student information. It is expected that this method will provide a new direction for future intelligent matching of roommates.

4.1. Analysis of intelligent matching model for university dormitory roommates based on optimized chameleon algorithm. The application effect of the chameleon algorithm based on optimized partitioning method in intelligent matching of college dormitory roommates demonstrated the superiority of this method. Table 4.1 denotes the parameter settings of the model.

In the optimized the chameleon algorithm, the algorithm parameter was set to 0.5 to strengthen the feature weight balance. To prevent differences in the range of data values and the influence of data dimensions on the results, feature value standardization was carried out. Similarity was calculated using cosine similarity to ensure accurate matching of data in high-dimensional space. Based on the 2022 data of college student dormitories, a new algorithm was used to match student dormitories using Python 3.7 and TensorFlow 2.3.1 frameworks. The student clustering effect diagram is denoted in Figure 4.1.

In Figure 4.1, while retaining as much information as possible, the student's multiple feature values were compressed into two main dimensions, namely roommate similarity and personal satisfaction. It analyzed the matching effect by constructing scatter plots for these two dimensions. Based on roommate similarity, the

Table 4.1: Various clustering centers

	Category 1	Category 2	Category 3	Category 4	Category 5	Category 6
TnR	0.129	0.101	0.047	0.568	0.079	0.170
PR	0.104	0.168	0.290	0.381	0.071	0.051
WIN	0.135	0.036	0.025	0.034	0.038	0.075
LOSE	0.143	0.085	0.046	0.032	0.063	0.041
HT	0.079	0.040	0.028	0.024	0.058	0.063
TsR	0.124	0.095	0.051	0.062	0.056	0.079

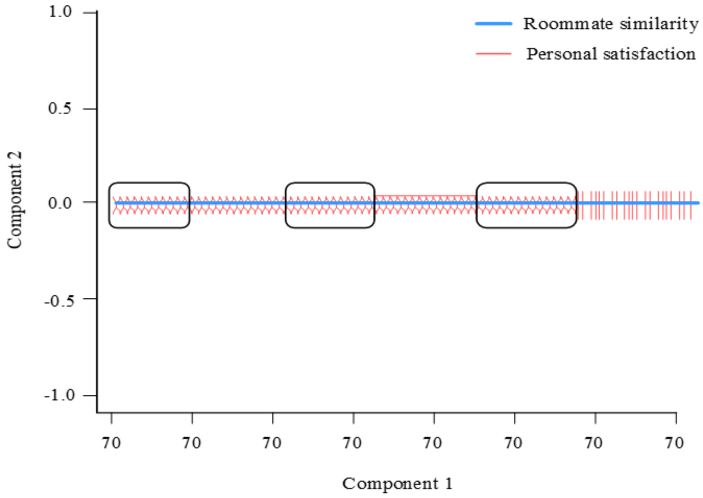


Fig. 4.1: Student clustering rendering

analysis content showed that students were clearly distinguished in this dimension, where different symbols represent different matching results. From the number of distinctions, students were accurately assigned to their respective dormitories without crossing, indicating a good matching effect. The clustering centers of six categories were divided by the clustering algorithm, as denoted in Table 4.2.

In Table 4.2, in the intelligent pairing of dormitory roommates, students were divided into six categories: freshmen, active participants, academic, sports enthusiasts, artistic, and work oriented. New students need more guidance, active participants bring vitality to the dormitory, academic students create a learning atmosphere, sports enthusiasts promote healthy living, artistic students enrich dormitory culture, and work oriented students provide management assistance. By optimizing the chameleon algorithm, efficient pairing between various types of students has been achieved, meeting the choices of students with different personalities and needs.

4.2. Analysis of intelligent matching application for college dormitory roommates based on optimized chameleon algorithm. The experiment proved the application effect of the optimized chameleon algorithm in intelligent matching of college dormitory roommates. To evaluate performance in depth and compare it with other matching algorithms, it would explore and analyze them from aspects such as matching accuracy, computational complexity, and runtime, providing theoretical reference and practical guidance for the future. The clustering efficiency comparison of different clustering algorithms in the intelligent matching method for college dormitory roommates is indicated in Figure 4.2.

In Figure 4.2, the application of the Chameleon, SKMeans-Chameleon, K-medoid-BIRCH, and K-medoid-Chameleon algorithms in intelligent matching of college dormitory roommates was compared, and the number

Table 4.2: Various clustering centers

	Category 1	Category 2	Category 3	Category 4	Category 5	Category 6
TnR	0.129	0.101	0.047	0.568	0.079	0.170
PR	0.104	0.168	0.290	0.381	0.071	0.051
WIN	0.135	0.036	0.025	0.034	0.038	0.075
LOSE	0.143	0.085	0.046	0.032	0.063	0.041
HT	0.079	0.040	0.028	0.024	0.058	0.063
TsR	0.124	0.095	0.051	0.062	0.056	0.079

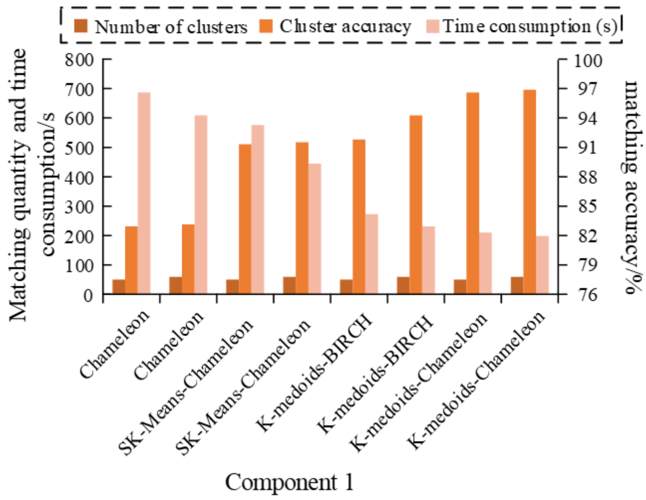


Fig. 4.2: Comparison of clustering efficiency of different clustering algorithms in intelligent matching methods for college dormitory roommates

of matches, matching accuracy, and running time were examined. The results showed that the K-medoids-Chameleon algorithm outperformed other algorithms in matching accuracy and runtime. Compared to the traditional Chameleon algorithm, its matching accuracy has been significantly improved and the running time has been significantly reduced. The clustering quality comparison of different clustering algorithms in the intelligent matching method for college dormitory roommates is expressed in Table 4.3.

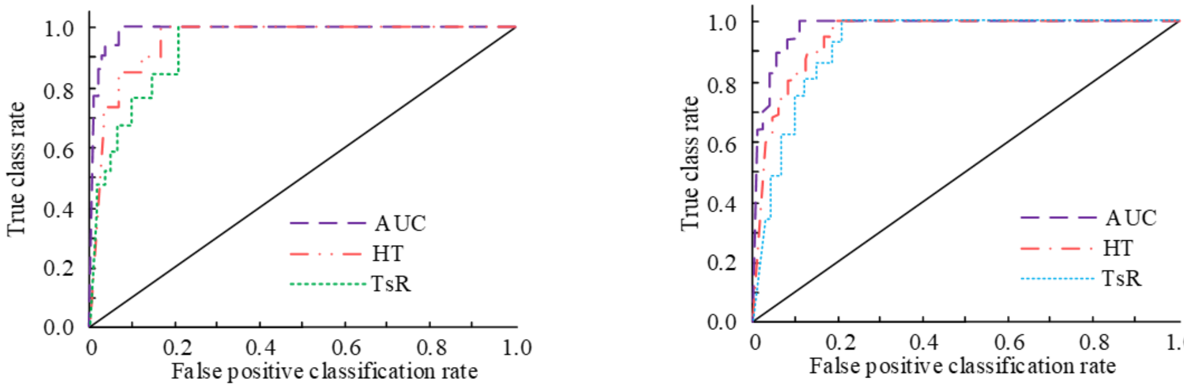
In Table 4.3, the Chameleon, SK-Means-Chameleon, K-medois-BIRCH, and K-medois-Chameleon algorithms were compared in terms of entropy, purity, and RI value for quality in intelligent matching of college dormitory roommates. The closer the entropy value was to 0, the closer the purity and RI values were to 1, and the better the matching effect. The K-medoids-Chameleon algorithm performed better than other algorithms in terms of minimum entropy, highest purity, and maximum RI value, verifying its effectiveness and superiority in roommate matching. It compared the AUC values of the chameleon algorithm based on optimized partitioning with other models on different datasets, as shown in Figure 4.3.

In Figure 4.3, the AUC value of the algorithm on the BBC dataset was 94.25%, which was higher than the other two models by 2.26% and 5.38%. On the Classic4 dataset, its AUC value was 93.13%, which was 2.08% and 5.04% higher than the other two models, respectively. These data fully demonstrated the superiority of this algorithm in intelligent matching of college dormitory roommates. The chameleon algorithm based on optimized partitioning was compared with other models in terms of accuracy, as shown in Figure 4.4.

In Figure 4.4, on the BBC dataset, the accuracy of this algorithm was 98.82%, which was 2.13% and 3.08% higher than the other two models, respectively. On the Classic4 dataset, its accuracy was 93.06%, which

Table 4.3: Comparison of clustering quality of different clustering algorithms in intelligent matching methods for college dormitory roommates

Clustering algorithm	Entropy	Purity	RI
Chameleon	0.31352	0.72453	0.54453
SK-Means-Chameleon	0.30143	0.73875	0.55098
K-medoids-BIRCH	0.29760	0.73986	0.55743
K-medoids-Chameleon	0.28607	0.75831	0.56083



(a) BBC

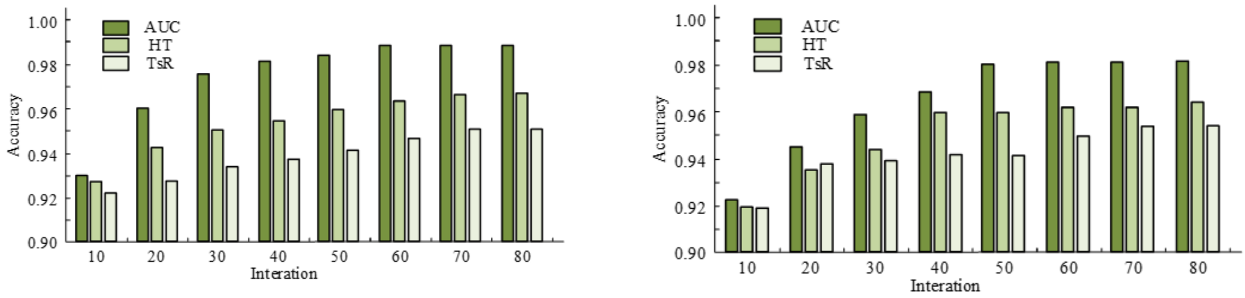
(b) Classic4

Fig. 4.3: AUC of several text clustering algorithms on different datasets

was higher than the other two models by 1.94% and 2.87%, respectively. These results further confirmed the efficiency and accuracy of the algorithm in roommate matching problems. The comparison between the chameleon algorithm based on optimized partitioning and other models in terms of modularity is expressed in Figure 4.5.

The chameleon algorithm based on optimized partitioning had the best flexibility when the number of modules was between 35 and 75. This might be due to the fact that in the college dormitory roommate intelligent matching scene, the data distribution within the number range of this module was more matched with the optimized division method of the chameleon algorithm, which reflected the flexibility of the algorithm. In cases where the number of modules was too low or too high, the effectiveness of the algorithm might decrease due to changes in data density or complexity. Compared with other models, this optimized partitioning chameleon algorithm had better modularity and flexibility when dealing with complex and ever-changing matching needs.

4.3. Discussion. The optimized chameleon algorithm provides an innovative method for intelligent matching of college dormitory roommates, integrating the personalities and needs of students, and improving the quality of life in dormitories. The algorithm demonstrates high matching accuracy, fast running time, and excellent clustering performance, indicating its social compatibility in the field of artificial intelligence assistance. Moreover, this high accuracy and performance translate to increased student satisfaction, as validated through satisfaction surveys post-matchmaking, aligning with the literature that correlates efficient roommate matching with enhanced student well-being and contentment. By adjusting parameters and standardizing data to achieve feature balance, the algorithm can accurately and efficiently allocate students to suitable dormitory environments. Compared with other clustering algorithms, the chameleon algorithm exhibits significant advantages in computational efficiency and matching quality, these advantages underscore the algorithm's effectiveness in creating harmonious living environments, as evidenced by longitudinal studies tracking student adaptation and



(a) BBC

(b) Classic4

Fig. 4.4: Accuracy of several text clustering algorithms on different datasets

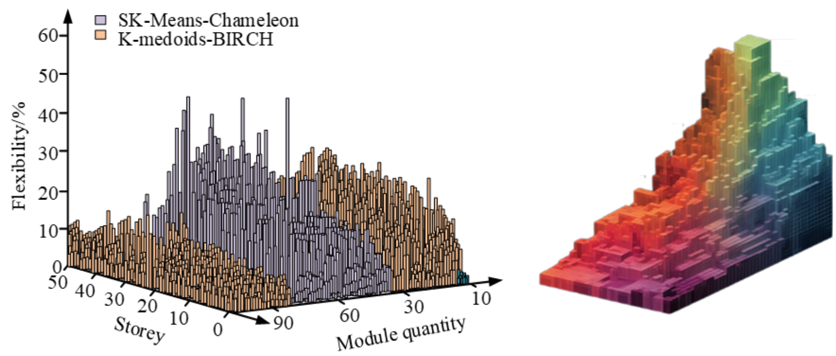


Fig. 4.5: Comparison of chameleon algorithm based on optimized partitioning with other models in terms of modularity

conflict resolution rates, thereby substantiating theories presented in the literature review regarding effective match outcomes. such as short running time, high entropy, purity, and RI index. The algorithm performs well on different datasets, demonstrating its wide applicability and reliability. The improvement results achieved through research are consistent with the trajectory of existing literature, which indicates that artificial intelligence and machine learning algorithms have great potential in solving complex human centered problems. Previous research has certainly involved algorithmic roommate matching, but the specificity of optimizing the chameleon algorithm in this situation provides a new application that adds value to existing methods. The integration of multi-objective optimization in hierarchical matching strategies goes beyond traditional one-dimensional methods and is consistent with emerging academic standards advocating the application of artificial intelligence in the human social environment. The success of this model indicates the transformation of the life paradigm of school students. In addition to the direct impact on student satisfaction and dormitory culture, these findings also indicate the potential application in a broader social matching system. In addition, the methodological advancements presented here may inspire further integration of optimized machine learning algorithms in various administrative and organizational tasks in education and other institutions, potentially leading to more data-driven, efficient, and satisfactory result oriented operations. Although this study has great potential, its application is limited by several factors, especially the scope of the dataset used. The effectiveness of algorithms mainly depends on academic centered datasets, and without significant adjustments or reconfiguration, they may not be able to be directly transformed into other environments. In addition, the dependence on algorithm efficiency overlooks the potential qualitative aspects of roommate compatibility, which may not be fully captured through quantifiable data. Future research can explore the application of optimized chameleon algorithms in different social matching scenarios, potentially improving and adjusting models to adapt to different datasets and matching standards.

5. Conclusion. In seeking an effective solution for intelligent matching of college dormitory roommates, a chameleon algorithm based on optimized partitioning was introduced. This algorithm optimized the traditional Chameleon algorithm to improve its application effectiveness in the intelligent matching problem of college dormitory roommates. The focus of optimization was to improve matching accuracy and reduce running time to meet the needs of practical applications. The experimental results confirmed the significant advantages of the chameleon algorithm based on optimized partitioning. In testing on the Classic4 dataset, the AUC value of this algorithm reached 93.13%, which was 2.08% and 5.04% higher than the other two models. The accuracy was 93.06%, which was also higher than the other two models by 1.94% and 2.87%. These results demonstrated the excellent performance of the algorithm in matching accuracy and runtime. Although this algorithm has demonstrated significant advantages on BBC and Classic4 datasets, its performance evaluation is not yet comprehensive. Because the current testing is limited to these two datasets, it cannot cover more application scenarios. In the next research direction, more in-depth research and improvement are still necessary, including testing on more datasets to comprehensively evaluate the performance of the algorithm.

REFERENCES

[1] Li, W. & Assaad, M. Matrix exponential learning schemes with low informational exchange. *IEEE Transactions On Signal Processing.* **67**, 3140-3153 (2019)

[2] Karakoç, N. & Duman, T. Rate selection for wireless random-access networks over block fading channels. *IEEE Transactions On Communications.* **68**, 1604-1616 (2019)

[3] Dove, N., Klingeman, D., Carrell, A. & Others Fire alters plant microbiome assembly patterns: integrating the plant and soil microbial response to disturbance. *New Phytologist.* **230**, 2433-2446 (2021)

[4] Carlstr"om, C., Field, C. & Bortfeld-Miller, M. Synthetic microbiota reveal priority effects and keystone strains in the Arabidopsis phyllosphere. *Nature Ecology & Evolution.* **3**, 1445-1454 (2019)

[5] Xiong, C., Zhu, Y., Wang, J., Sin, B., Han, L., Shen, J. & He, J. Host selection shapes crop microbiome assembly and network complexity. *New Phytologist.* **229**, 1091-1104 (2021)

[6] Meyer, K., Porch, R., Muscettola, I., Vasconcelos, A., Sherman, J., Metcalf, C., Lindow, S. & Koskella, B. Plant neighborhood shapes diversity and reduces interspecific variation of the phyllosphere microbiome. *The ISME Journal.* **16**, 1376-1387 (2022)

[7] Gaveau, J. & Leturc, X. Martret C J L, Assaad M. Trial and Error Learning for Dynamic Distributed Channel Allocation in Random Medium. *EEE Transactions On Wireless Communications.* **20**, 8177-8190 (2021)

[8] Zucker, S. & And, B. and sub-optimal rate and power allocation for random networks. *Ransactions On Emerging Telecommunications Technologies.* **34**, 4730-4754 (2023)

[9] Kiesewetter, K. & Afkhami, M. Microbiome-mediated effects of habitat fragmentation on native plant performance. *Ew Phytologist.* **232**, 1823-1838 (2021)

[10] Bistritz, I., Ward, A., Zhou, Z. & Bambos, N. Smart Greedy Distributed Energy Allocation: a Random Games Approach. *EEE Transactions On Automatic Control.* **67**, 2208-2220 (2021)

[11] Lai, P., He, Q., Xia, X., Chen, F. & Yang, Y. Dynamic User Allocation in Stochastic Mobile Edge Computing Systems. *EEE Transactions On Services Computing.* **15**, 2699-2712 (2021)

[12] Henning Smith, C., Hernandez, A., Lahr, M. & Kozhimannil, K. Rural-Urban Differences in Access to, Attitudes Toward, and Satisfaction with Care Among Medicare Beneficiaries. *Health Services Research.* **55**, 20-21 (2020)

[13] Liang, Z. & Chen, P. An automatic clustering algorithm based on the density-peak framework and Chameleon method. *Attern Recognition Letters.* **150**, 40-48 (2021)

[14] Li, Y., Ma, J., Miao, Y. & Li, H. Choo K K R. *DVREI: Dynamic Verifiable Retrieval Over Encrypted Images.* **71**, 1755-1769 (2021)

[15] Alexiadis, A., Veliskaki, A., Nizamis, A., Bintoudi, A., Zyglakis, L. & Triantafyllidis, A. smarthome conversational agent performing implicit demand-response application planning. *Ntegrated Computer-Aided Engineering.* **29**, 43-61 (2022)

[16] Nie, Q., Zhang, L. & Li, S. How can personal carbon trading be applied in electric vehicle subsidies?. *A Stackelberg Game Method In Private Vehicles.* **313**, 1-17 (2022)

[17] Feng, J., Chen, S., Ye, Z., Miralinaghi, M., Labi, S. & Chai, J. Repositioning Shared Urban Personal Transport Units: Considerations of Travel Cost and Demand Uncertainty. *Ournal Of Infrastructure Systems.* **27**, 1-12 (2021)

[18] Zhu, X., Xiangfei, L. & Feng, X. Simulation Study on Emergency Evacuation Behavior of College Students in Considering the Influence of Dormitory Interpersonal Relationship. *Journal Of Systems Science And Information.* **11**, 124-128 (2023)

[19] Zhao, X., Peng, B., Zheng, C. & Others Closed-loop supply chain pricing strategy for electric vehicle batteries recycling in China. *Environment, Development And Sustainability.* **24**, 7725-7752 (2022)

[20] Yang, D., Yang, L., Chen, X. & Others Research on credit pricing mechanism in dual-credit policy: is the government in charge or is the market in charge?. *Environment, Development And Sustainability.* **25**, 1561-1581 (2023)
[21] Fang, Y., Luo, B., Zhao, T., He, D., Jiang, B. & ST-SIGMA:Spatio-and, L. and interaction graph aggregation for multi-agent perception and trajectory forecasting. *CAAI Transactions On Intelligence Technology.* **7**, 744-757 (2022)
[22] Guo, Y., Mustafaoglu, Z. & Koundal, D. Spam Detection Using Bidirectional Transformers and Machine Learning Classifier Algorithms. *Journal Of Computational And Cognitive Engineering.* **2**, 5-9 (2022)

Edited by: Zhengyi Chai
Special issue on: Data-Driven Optimization Algorithms for Sustainable and Smart City
Received: Nov 16, 2023
Accepted: Mar 5, 2023