

## *Retraction*

# **Retracted: Library Management System Based on Data Mining and Clustering Algorithm**

### **Wireless Communications and Mobile Computing**

Received 13 September 2023; Accepted 13 September 2023; Published 14 September 2023

Copyright © 2023 Wireless Communications and Mobile Computing. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] L. Pang, "Library Management System Based on Data Mining and Clustering Algorithm," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 1398681, 6 pages, 2022.

## Research Article

# Library Management System Based on Data Mining and Clustering Algorithm

Lu Pang 

Shandong Women's University, Jinan, Shandong 250014, China

Correspondence should be addressed to Lu Pang; 1710211228@hbut.edu.cn

Received 28 June 2022; Revised 10 August 2022; Accepted 17 August 2022; Published 2 September 2022

Academic Editor: Aruna K K

Copyright © 2022 Lu Pang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to solve the problem of building system services between readers and libraries, this paper proposes a library management system based on data mining and clustering algorithm. The library management model is built based on data mining technology and clustering algorithm, and the hybrid clustering algorithm in the data mining platform Weka is used for library data mining. The experimental results show that with the same amount of data, the hybrid clustering algorithm takes 5.5 seconds to process information from 0 to 300, which is at least 1 second faster than the other two algorithms. *Conclusion.* The algorithm is not only a means of library system automation management, but also an effective means to realize library information modernization.

## 1. Introduction

Modern library management systems produce a large amount of information data every day. These data have become valuable resources for data mining and machine learning. The literature in the library is an important way for people to acquire knowledge [1]. However, with the rise of information technology and the popularization of the Internet, libraries not only have traditional paper books, but also more and more e-book resources can provide information resources to the public in the library [2]. The library system also records readers' information resources and changes new data to provide convenience for readers [3]. However, as time goes on, the data will become more and more, the book materials will become larger and larger, and the relationship between readers and libraries will become more complex. Therefore, a better system is needed to process information data to provide data support for library construction [4]. The emergence of data mining technology has solved the problem of huge data. It can not only quickly search the books that readers want, but also analyze readers' usage habits to recommend literature and put forward reasonable procurement suggestions through literature analysis [5]. Therefore, data mining technology combined

with library management system uses association technology to search documents, understand the internal relationship between readers and library, and put forward personalized recommendations.

Because the current library management system cannot find the knowledge hidden in the massive data, and cannot predict the demand information of readers, it is unable to reasonably optimize the collection structure and interlibrary distribution of Libraries in multiple regions. It mainly applies data mining technology to analyze the data in the library management system, find the readers' demand information, and then provide it to the library deployment management system as the basis for decision-making [6]. The main contribution is to reasonably analyze historical data and develop a practical decision support system by using the important algorithms in data mining. The system can provide a more reasonable guidance for each batch of new books on the shelves. This has produced great benefits for optimizing the allocation of book resources in multiple regions [7].

Book classification is the focus of the system. For example, the traditional PAM algorithm technology can effectively solve the classification of different books. Clarans algorithm is also a means of data processing, but both of them have limitations in the amount of data [8].

## 2. Literature Review

Mobile Internet service optimization refers to data collection, data analysis, and efficient data processing for the running network [9]. Data analysis is the focus of Internet optimization. In the information age, there is a huge amount of data, but effective and useful data are hidden by a large amount of data. What we need to solve is to find the data we need, find out the relationship between the data, and make decisions for decision makers, so as to get the desired results [10].

Data mining refers to extracting or “mining” knowledge from a large amount of data. It is an important step in the process of knowledge discovery. Figure 1 shows a typical data mining process, which includes the following: ① pre-processing the source database to get the target data; ② data mining of target data and extracting data patterns; and ③ evaluate the patterns, get really interesting patterns, and use knowledge representation technology to provide users with knowledge [11].

The most important thing of data mining is to clarify the mining objectives and tasks; select different mining algorithms according to different tasks; and determine whether to carry out data classification, clustering, association rules, or time series analysis [12]. Data mining tasks can be descriptive, describing the general nature of the database, or predictive, inferring and predicting the current task. To select an appropriate mining algorithm, we should not only consider the characteristics of the data, but also consider the needs of users, and clarify whether we prefer to acquire descriptive and easy to understand knowledge or predictive knowledge with high accuracy [13]. After selecting the mining algorithm, data mining operations can be carried out to obtain useful patterns.

If the mining patterns are found to have redundant or irrelevant knowledge after evaluation, they need to be eliminated. If the patterns cannot meet the needs of users, they need to be re mined. The patterns obtained from data mining are often not visual and difficult to understand. They need to be reasonably explained to users. They can be transformed into forms that are easy to understand by users with the help of visual tools or graphical user interfaces.

The main task of the data mining module is to use the corresponding mining algorithm to find unknown knowledge, capture the readers’ demand information hidden in the massive data, and provide support for better deployment of book resources. The module adopts the object-oriented design idea to minimize the control coupling of the system and facilitate the update and maintenance of the algorithm. The task of the core management module is to issue control commands to other sub modules. For example, start the pre-processing module to read the original data, and call the data mining module to find the unknown reader demand information. The book deployment strategy creation module uses the rules provided by data mining and the existing prior knowledge to provide corresponding decision support for the shelving and collection adjustment of books. The whole data mining process is a dynamic and reciprocating process, which needs to be constantly modified and improved. In the

process of mining, the expected results may not be achieved if the data cleaning is not in place, the type conversion is wrong, the attribute selection is improper, or the mining algorithm is improperly selected. The mining steps must be reviewed and corrected [14].

Clustering is the process of clustering data objects. The objects within clusters are very similar, while the objects between clusters are highly different. The degree of dissimilarity is evaluated according to the attribute value of the description object and is usually measured by distance. The difference between clustering and classification is that clustering does not depend on predefined classes and does not need training sets. The commonly used clustering algorithms include the following: partition clustering, hierarchical clustering, density-based clustering, and grid-based clustering algorithm.

The library management system has accumulated a large amount of business data in the long-term use process. Through data mining of readers’ borrowing records and access logs, we can find that the readers’ demand preferences hidden in the data actively provide personalized information services to meet the needs of different readers and improve the reader service quality of the library system [15].

This paper uses the hybrid cluster analysis technology to classify documents and provides the basis for the collection and construction of library documents through the hidden data laws in the process of document borrowing and returning. Data mining technology can find out the potential needs of readers, provide personalized help, and help readers choose to buy e-books so that readers can quickly and accurately use the resources of the library. Through the implementation of the algorithm, the system service between readers and libraries is built, and compared with other algorithms, the superiority of the hybrid clustering algorithm is obtained, which proves the rationality and effectiveness of the algorithm.

## 3. Method

*3.1. System Modeling.* Mobile Internet service optimization refers to data collection, data analysis, and efficient data processing for the running network. Data analysis is the focus of Internet optimization. In the information age, there is a huge amount of data, but effective and useful data are hidden by massive data. What we need to solve is to find the data we need, find out the relationship between the data, and make decisions for decision makers, so as to get the desired results [16].

*3.1.1. Establishment of Loan/Return Model.* Library management system is a computer system built according to the specific business needs of the library. The system mainly provides two models to provide services for the actual business of the library. One is the book borrowing and returning management model, and the other is the reader library management model. The “book borrowing and returning management” is mainly responsible for the general business of the library, which mainly includes querying books, lending and returning books, and booking books [17]. The

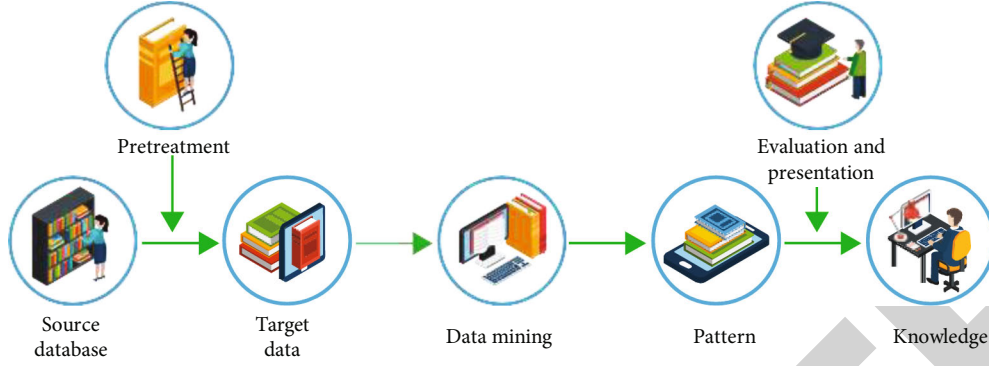


FIGURE 1: Data mining process.

model is shown in Figure 2. Each reader user is set as  $x_i$ , and the book is set as  $c_j$ . The model establishes the relationship between  $x_i$  and  $c_j$ .

**3.1.2. Establishment of Reader Base Model.** The reader library management model is mainly used to protect, modify and report the loss of information by readers. In addition, it also includes readers' handling of certificates and reissue of certificates in the library [18]. The model is shown in Figure 3. There are two ways for readers: One is to timely handle their certificates to the management personnel and to report the loss of their certificates, and the other is that readers can conduct business processing through the online main page of the library to save time. The last card replacement should be handled by the management personnel.

**3.2. Hybrid Clustering Algorithm Design.** The library management system consists of two modules, including the background system for readers, users, and managers. The two modules are divided into several sub-blocks to realize their respective functions. The function design of the algorithm is as follows.

Reader management is divided into user information registration, user login, and browsing and modification of user personal information. The user registration process on the main page of the system includes filling in useful information such as name, ID number, work unit, and binding amount to realize registration. When readers log in to the system, they can improve, view, and modify their personal information. The backstage of the management personnel manages the massive information of the library books and realizes the functions of adding, deleting, editing, and displaying the book information. In addition, the management and technical personnel must regularly repair the system, install patches in time, and upgrade the system [19].

The hybrid clustering algorithm is used to analyze library books. The first step is to determine the target of hybrid clustering: given a set of  $n$  a-dimensional books or corresponding user data  $X = \{x_1, x_2, \dots, x_i, \dots, x_n\}$ , and  $x_i \in \mathbb{R}^a$ , determine the number  $m$  of subsets of book data to be generated. The hybrid clustering algorithm classifies each reader's books and unsold books and performs  $m$  partitions  $C = \{c_m, i = 1, 2, \dots, m\}$ . The type of information represents a book and user  $c_m$ . For all kinds of  $c_m$ , there is a category

center value  $u_i$ .  $u_i$  is the most representative numerical information of this category, that is, the center value score. The Euclidean distance is used as the basis to judge the similarity. The sum of the squares of the distances from each point in each book category to the  $u_i$  is calculated as the similarity between the point and the central value. Then, the sum of the squares of the Euclidean distance is

$$J(c_m) = \sum_{x_i \in c_m} x_i - u_m^2. \quad (1)$$

The objective function of hybrid clustering is the sum of squares of distances. If  $J(c)$  is the smallest, Formula (2) is

$$J(c) = \sum_{m=1}^M J(c_m) = \sum_{m=1}^M \sum_{x_i \in c_m} x_i - u_m^2 = \sum_{m=1}^M \sum_{i=1}^n a_{mi} x_i - u_m^2. \quad (2)$$

In Formula (2),  $a_{mi} = 1(x_i \in c_i)$  or  $a_{mi} = 0(x_i \notin c_i)$ . It can be seen that the central UI of hybrid clustering should be taken as the average of the data points of each  $c_m$  category and each book category.

The hybrid clustering algorithm starts from the initial  $M$  category [20]. In the hybrid clustering algorithm, the total distance sum of squares increases according to the category of the number  $m$ , but the distance sum tends to decrease. In special cases, when  $M = 0$ ,  $J(c) = 0$ . Therefore, it can be concluded that the minimum value of  $J(c)$  can be obtained only when the sum of squares of the total distance is under the determined number of categories  $M$ .

The hybrid clustering algorithm divides the book data set into  $M$  categories. The flow of the algorithm is as follows:

Step 1: Randomly select  $M$  initial clustering centers from the book data set

Step 2: For each data object in the book data set, calculate the distance between the object and all other clustering centers, and divide it into the nearest category according to the nearest neighbor criterion

Step 3: After the calculation in the previous step, recalculate the cluster center of each new cluster according to the calculation results, and calculate the sum of the squares of the distances of all book data

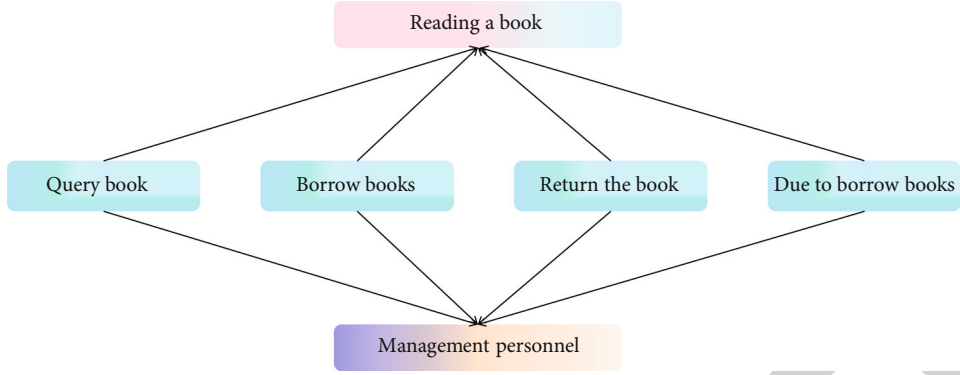


FIGURE 2: Book borrowing and returning management.

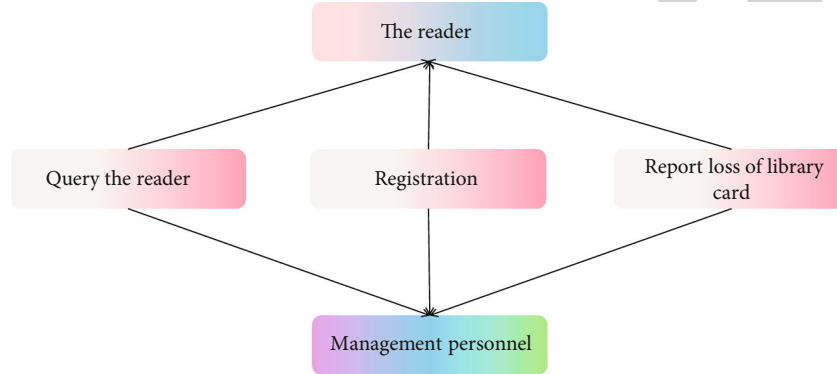


FIGURE 3: Reader library management.

Step 4: Judge whether the  $J(c)$  value of the obtained cluster center has changed. If it has changed, repeat Steps 2 and 3. If the cluster center does not change, the algorithm ends. If there is no change, the algorithm ends directly.

Let  $A_{i,j}$  be the similarity between the book information  $M_i$  and  $M_j$ , then

- 1)  $A_{i,j} = \pm 1 \iff M_i = m M_j$  ( $m$  is a constant and  $m \neq 0$ )
- 2)  $|A_{i,j}| \leq 1$  ( $i, j$  is any number)
- 3)  $A_{i,j} = A_{j,i}$  ( $i, j$  is any number)

#### 4. Results and Discussion

The experimental object of the algorithm is the school library of a school. The test environment includes server and client. The server-side part for the test is Lenovo Windows Server 2003. The desktop computer used is Intel Core i7 with a CPU frequency of 3.2 Hz and a memory of 132gb ddr3a. Finally, the experimental results are analyzed by running the simulation script [21].

The system obtained by the hybrid clustering algorithm is shown in Figure 4. The four modules of book registration form, book registration, inventory books, and registry form are the result categories of the algorithm. Book registration is the core technology of the hybrid clustering algorithm method. It is specific to each class through the algorithm, so the process of design refinement can be completed. The system can effectively complete the

realization and management of the huge data in the library and is conducive to the effective contact between users and the library.

Cluster analysis method is used to mine and evaluate the contents of books and score books. In this way, good data can be presented in the system interface to provide readers' suggestions. Each good book becomes a collection group. The value at the center of the collection and the representative books are the central value, and the central value score is the scoring index of such books [22].

The system has the function of evaluating books, as shown in Table 1, including cover design, book materials, content value, and purchase intention. The final total score can provide the basis for other readers and users to read and purchase and also help the construction of the library. It is the embodiment of personalized services.

In addition to the hybrid clustering algorithm in this paper, there are many traditional algorithms for library information data processing, which can effectively carry out system management. The advantage of the hybrid clustering algorithm lies in its fast processing speed, larger amount of processed data, and more advantages in system maintenance and upgrading. With the gradual growth of time, the hybrid clustering algorithm takes 5.5 seconds to process information from 0 to 300, and the speed is at least 1 second faster than the other two algorithms. Figure 5 shows the processing speed comparison between this algorithm and other algorithms [23–25].



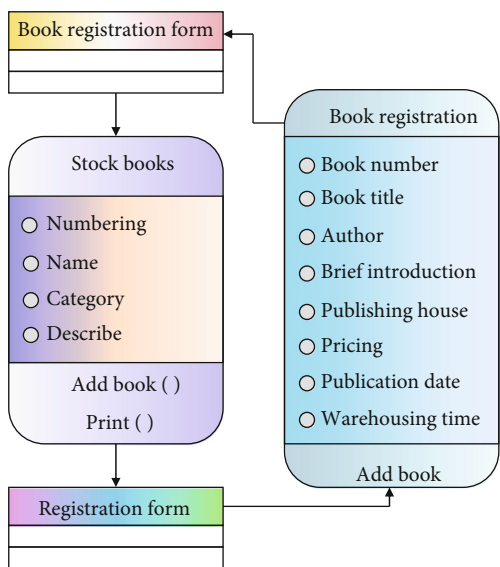


FIGURE 4: Library management.

TABLE 1: Evaluation function of books in the system.

Serial number	Scoring rubric	Total score	Rating
1	Cover design	10	
2	Book materials	10	
3	Content value	10	
4	Purchase intention	10	
Total score 40, grade A B C			

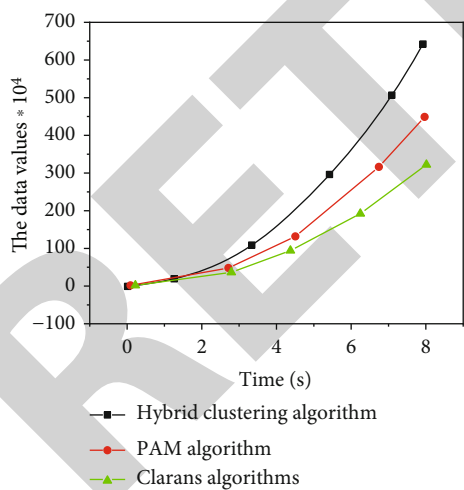


FIGURE 5: Comparison of data processing capacity and processing speed of different algorithms.

### 5. Conclusion

This paper presents the research of library management system based on data mining and clustering algorithm. By building a connection between a large amount of book data accumulated in the library and user information, it is used to help the library to carry out system management. As a huge

database, the introduction of data mining technology makes the management of the library more convenient. After data mining, the book information can be reasonably arranged based on the hybrid clustering algorithm to improve the convenience of the system. Through the algorithm implementation and algorithm comparison, it can be seen that the system combined with the algorithm in this paper can form a good system management order, realize functional visualization, and provide services for the users of book cases and management technicians, so the algorithm is reasonable.

### Data Availability

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

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### References

- [1] I. V. Timoshenko, "The principles of unique identification of library documents in automatic proximity identification systems," *Scientific and Technical Libraries*, vol. 1, no. 2, pp. 65–80, 2021.
- [2] M. E. Luka and J. Hutchinson, "Trust in the system: an introduction to the #aoir2019 special issue," *Information, Communication & Society*, vol. 23, no. 6, pp. 794–801, 2020.
- [3] L. Yi, "RETRACTED: Erratum to "research on china academic library and information system" [Open Access Library Journal, 2020, Volume 7: e6689]," *Open Access Library Journal*, vol. 8, no. 10, pp. 1–10, 2021.
- [4] C. Vathanak, "China's official development assistance: an implication of the transport infrastructure development in Cambodia," *Open Access Library Journal*, vol. 8, no. 8, pp. 1–11, 2021.
- [5] M. Acharya, K. P. Acharya, K. Gyawali, P. Acharya, and Devkota, "Discussing professor yin Kejing's drug use law for mammary hyperplasia based on data mining technology," *International Journal of Clinical and Experimental Medicine*, vol. 5, no. 3, pp. 403–407, 2021.
- [6] Q. Wang and B. Zhang, "Research and implementation of the customer-oriented modern hotel management system using fuzzy analytic hiererchical process (fahp)," *Journal of Intelligent Fuzzy Systems*, vol. 40, no. 4, pp. 8277–8285, 2021.
- [7] X. Zhou, X. Zhang, Z. Dai, R. L. Hermaputi, and Y. Li, "Spatial layout and coupling of urban cultural relics: analyzing historical sites and commercial facilities in district iii of Shaoxing," *Sustainability*, vol. 13, no. 12, p. 6877, 2021.
- [8] F. Wang, L. Zhang, and X. Xu, "A literature review and classification of book recommendation research," *Journal of Information Systems and Technology Management*, vol. 5, no. 16, pp. 15–34, 2020.
- [9] B. Bahmani-Firouzi, "A new hybrid algorithm based on pso, sa, and k-means for cluster analysis," *International Journal of Innovative Computing Information & Control*, vol. 6, no. 7, pp. 3177–3192, 2010.
- [10] C. L. Hsu, H. P. Lu, and H. H. Hsu, "Adoption of the mobile internet: an empirical study of multimedia message service (mms)," *Omega*, vol. 35, no. 6, pp. 715–726, 2007.

- [11] K. Kim, J. W. Lee, B. G. Park et al., "Investigation of correlative parameters to evaluate EUV lithographic performance of PMMA," *RSC Advances*, vol. 12, no. 5, pp. 2589–2594, 2022.
- [12] X. Gong, F. Wu, R. Xing, J. Du, and C. Liu, "Lcbrg: a lane-level road cluster mining algorithm with bidirectional region growing," *Open Geosciences*, vol. 13, no. 1, pp. 835–850, 2021.
- [13] C. Tunca, G. Salur, and C. Ersoy, "Deep learning for fall risk assessment with inertial sensors: utilizing domain knowledge in spatio-temporal gait parameters," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 7, pp. 1994–2005, 2020.
- [14] E. Helm, A. M. Lin, D. Baumgartner, A. C. Lin, and J. K ung, "Towards the use of standardized terms in clinical case studies for process mining in healthcare," *International Journal of Environmental Research and Public Health*, vol. 17, no. 4, p. 1348, 2020.
- [15] R. Rehman, M. Osto, N. Parry et al., "Ewing sarcoma of the craniofacial bones: a qualitative systematic review," *Otolaryngology-Head and Neck Surgery*, vol. 166, no. 4, pp. 608–614, 2022.
- [16] C. Hou, Q. Zhao, and T. Basar, "Optimization of web service-based data-collection system with smart sensor nodes for balance between network traffic and sensing accuracy," *IEEE Transactions on Automation Science and Engineering*, vol. 9, pp. 1–13, 2020.
- [17] T. Silwattananusarn and P. Kulkanjanapiban, "Mining and analyzing patron's book-loan data and university data to understand library use patterns," *International Journal of Information Science and Management*, vol. 18, no. 2, pp. 151–172, 2020.
- [18] E. B. Gyau, L. Jing, and S. Akowuah, "International students library usage frequency patterns in academic libraries: a user survey at Jiangsu university library," *Open Access Library Journal*, vol. 8, no. 7, pp. 1–20, 2021.
- [19] G. Enos, "Detroit health system seeks upgrade of inpatient care via partnership," *Mental Health Weekly*, vol. 31, no. 3, pp. 1–7, 2021.
- [20] H. Gao, Y. Li, P. Kabalyants, H. Xu, and R. Martinez-Bejar, "A novel hybrid PSO-k-means clustering algorithm using Gaussian estimation of distribution method and L vy flight," *IEEE Access*, vol. 8, pp. 122848–122863, 2020.
- [21] A. Sharma and R. Kumar, "Performance comparison and detailed study of AODV, DSDV, DSR, TORA and OLSR routing protocols in ad hoc networks," in *2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, pp. 732–736, Wagnaghat, India, 2016.
- [22] M. Raj, P. Manimegalai, P. Ajay, and J. Amose, "Lipid data acquisition for devices treatment of coronary diseases health stuff on the Internet of medical things," *Journal of Physics Conference Series*, vol. 1937, p. 012038, 2021.
- [23] J. Liu, X. Liu, J. Chen, X. Li, and F. Zhong, "Plasma-catalytic oxidation of toluene on Fe<sub>2</sub>O<sub>3</sub>/sepiolite catalyst in DDBD reactor," *Journal of Physics D: Applied Physics*, vol. 54, no. 47, p. 475201, 2021.
- [24] P. Ajay, B. Nagaraj, R. Arun Kumar, R. Huang, and P. Ananthi, "Unsupervised hyperspectral microscopic image segmentation using deep embedded clustering algorithm," *Scanning*, vol. 2022, Article ID 1200860, 9 pages, 2022.
- [25] G. Veselov, A. Tselykh, A. Sharma, and R. Huang, "Special issue on applications of artificial intelligence in evolution of smart cities and societies," *Informatica*, vol. 45, no. 5, p. 603, 2021.