2022 IEEE North Karnataka Subsection Flagship International Conference (NKCon)

An Effective Storage Management for University Library using Weighted K-Nearest Neighbor Algorithm

1st K.B.Glory
Department of Engineering English
KoneruLakshmaiah Education Foundation
Vaddeswaram, India
kommalapatiglory@gmail.com

2nd D.Venkatesan
Department of Artificial Intelligence and MachineLearning
St.Martin's Engineering College
Secunderabad, India
dvenkatesanme@gmail.com

3rd G. Naga Rama Devi
Department of Computer Science & Engineering
CMR Institute of Technology
Hyderabad, India
ramadevi.abap@cmritonine.ac.in

4th Chappeli Sai Kiran
Department of Mechanical Engineering
CVR College of Engineering,
Mangalpalli, India
csaikiran001@gmail.com

5th Prerana Nilesh Khairnar
Department of Electronics and Telecommunications Engineering
SirVisveavaraya Institute of Technology
Nashik, India
autadeprerana@gmail.com

6th S. Priya
Department of Computer Science
Government First GradeCollege
Kolar, India
priya12mithul@gmail.com

2022 IEEE North Karnataka Subsection Flagship International Conference (NKCon) | 978-1-6654-5342-4/22/$31.00 ©2022 IEEE | DOI: 10.1109/NKCON56289.2022.10126745

***Abstract*—The most fascinating topic in economic geography is the storage location-allocation problem. The storage serves as a transition point to lower the cost of transmission. To produce an accurate and approximate response, two a model based hybrid of k-means –Particle Swarm Optimization(KPSO) proposed in this work. When compared to the existing model, the proposed model formulation is simpler and easier to understand. The testing findings show that the proposed model makes better use of the computer's Random Access Memory (RAM), allowing us to solve medium-sized tasks. This approach cannot outperform the MIP model in terms of run time.The multi-assignment facility location queries are included in the extension of the CP formulations. Initial PSO solutions are produced using the well-known data clustering technique K-means. The experimental results demonstrate that in terms of time, objective value, and reliability of performance metrics, the KPSO method is superior to the PSO.**

Keywords—Transportation model, SCM, Material management, K-means algorithm.

I. INTRODUCTION

A component of data mining known as data clustering seeks to categorize or group data items within a dataset based on their similarities and differences [1-2]. To make data items inside a cluster more similar to one another than to those in other clusters, a dataset is split into clusters [3]. In other words, data grouping is done to increase inter-cluster distance while decreasing intra-cluster distance between data items [4]. Numerous applications, including biological data, analysis of social networks mathematical programming, customer segmentation, picture segmentation, data summarizing, and consumer research, have benefited greatly from the use of data clustering for categorizing data [5-7]. The process of clustering information may be done in a variety of ways. The region-based segmentation methods and the hierarchical clustering methods are the two main groups into which these techniques fall. The dendrogram produced by the hierarchical clustering approach shows the order in which the dataset's data items were clustered by iteratively hierarchically grouping them. Requiring a particular goal function, the partitional clustering approach creates a single dataset partitioning to recover underlying natural groups within the dataset without using any hierarchical structure [8]. The well-known K-means classification algorithm is one of the several partitional similarity measures.

Without previous domain expertise, it might be challenging to select optimal cluster numbers for datasets including high dimensional data items of different densities and sizes. The K-means technique is ineffective for automated clustering due to the need to pre-define the number of clusters. Because the appropriate number of clusters in a dataset is determined automatically for automatic clustering methods without the need for baseline knowledge about the dataset's data items, As a result, metaheuristics derived from nature have been used to solve automated clustering issues. This standard k-means technique has been enhanced in terms of both performance and autonomous clustered problem handling by combining a few nature-inspired metaheuristics algorithms. In this article, we discuss and evaluate the many nature-inspired optimization methods that have recently been combined with K-means or any of its variations to address automated statistical data analysis issues [9]. Numerous evaluations of the usage of association rules inspired by nature have been published, many of which only focus on autonomous grouping. Current research on all significant existing metaheuristic techniques for automated clustering issues [10].

II. LITERATURE REVIEW

The choice of providers has drawn a lot of attention in supply chain management, especially when it comes to the purchasing departments of every company [11]. A range of Multiple-Criteria Decision Analysis techniques was used to choose the best bidder based on the specifications outlined by management staff [12]. The procurement teams use arbitrary evaluation standards to evaluate the suppliers. In all, there are two processes involved in choosing the right vendors. One is NN-DEA to address the measuring criterion's lack of data. Data Envelopment Analysis and Neural Network methods are combined in the Analytical Hierarchy Technique [13]. AHP is used to evaluate dimensions, DEA is used to evaluate standardized guidelines, and NN is used to assess the effectiveness of providers [14].NN algorithms were also employed in conjunction with AHP to determine variable weighting and NN to choose appropriate suppliers. To limit alternative amounts and choose the best cluster throughout the selection phase, mix "AHP with NN but employs Fuzzy Set Theory, whereas integrates FST with AHP analyses and clustering analytics [15]. To enhance training search technology, PSO is used to obtain main weights and build a network, and NN chooses the best provider based on previous data. Reverse process evaluation of requirements is used, subjecting potential providers to ANN. The assessment procedure ensures that PSO will be used to determine which supplier is the best. "DEA is planned to be linked with SVM," reads [16]. The optimal supplier may be chosen using SVM once the performance numbers are obtained using DEA. For choosing green suppliers, researchers have proposed the artificial neural network - multi-attribute decisions analysis technique, which combines DEA, Analytical Network Method, and NN models [17-19]. This could handle missing values and gets around the drawbacks of DEA models.

Although this learning has a lot of promise, it also has a lot of restrictions. Concerns about biases in algorithms, security, accountability, and information protection are among the many questions related to ethical concerns. The ideas of explicability and interpretability in the setting of human learning are also highlighted in this special issue [20].To make AI more dependable for users in learning contexts and to avoid misunderstandings, we need much more research and evidence-based dialogue. In addition to already issued patents in the sector, we conducted a thorough study of interdisciplinary computerized bibliographic databases [21]. We have found development tools that can help with different levels of digital merging. Having created a big data-driven, AI-enhanced reference model that guides developers toward a fully functional DT-enabled solution. Furthermore, we revealed problems and presented prospects to demonstrate the passion for research of AI-ML for digital twinning.

III. PROPOSED SYSTEM

In this paper, two new ways to solve the existing problem are proposed. The p-HLAP has been solved using a variety of heuristic and meta-heuristic techniques, although accurate solutions have seldom been created. Additionally, it has been demonstrated to be superior to alternative precise solution techniques for a variety of situations. Since there isn't a precise solution for p-HLAP and CP is effective at solving many other kinds of issues, this study's main innovation is the development of a CP formulation. As a result, the problem is written in a CP-appropriate manner. To address various HLAP kinds, such asmultiple assignment p-HLP and single HLP with restricted capacities, the CP formulation is further expanded. Furthermore, a combination of KPSO is developed to provide high-quality solutions. Various clustering is produced by the K-means method in different runs. This is generally a flaw. This flaw served as a strength for us as we came up with several PSO first remedies. The findings are then examined to determine which strategy is best for the various sizes of the issue after the strengths and drawbacks of the two novel solution strategies are contrasted with MIP and conventional PSO.

A. Problem description

To make travel between them easier and more affordable, p-HLAP is concerned with placing p Stores among n nodes and assigning other networks to one of the storage nodes. The transit is carried out utilizing Storage systems rather than direct transfers among locations, which reduces the cost of transport [25]. In Fig.1, a schematic of this technology is shown.

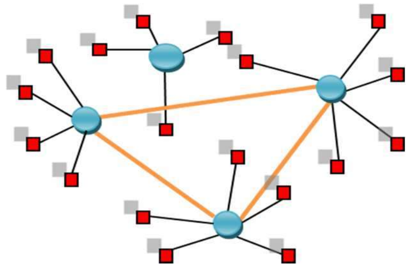

Fig. 1. Proposed architecture of solving p-HLP issues

Equation (1) assurances the creation of p Memory. Each node is given a Memory according to (2). In (3) states that a node must serve itself if it is formed as a repository. This solution has 2I+1 restrictions and a 2I variable.

$$min \sum_{x=1}^{X} \left[C_c \sum_{y=1}^{X} f_{xy} d_{xA_X} + C_d \sum_{x=1}^{X} \sum_{y=1}^{X} f_{xy} d_{A_x A_y} \right]$$

$$\sum_{x}^{X} [R_x \geq 1] = P \qquad (1)$$

$$A_x = \sum_{y}^{X} y[A_x - y] R_y \;\; 1 \leq x \leq X \qquad (2)$$

$$A_x R_x = x[R_x \geq 1] 1 \leq x \leq X \qquad (3)$$

A community of population of solutions is typically the starting point for PSO, which then aims to enhance these solutions. Crossover and mutation are two different sorts of operators that are employed for improvement. Some academics have created novel operators or employed inventive algorithms to construct early answers to speed up problem-solving based on the characteristics of the challenge. Here, we proposed accelerating the PSO for p-HLAP using the k-means method and a novel crossover operator. Fig.2. shows the planned KPSO flowchart.

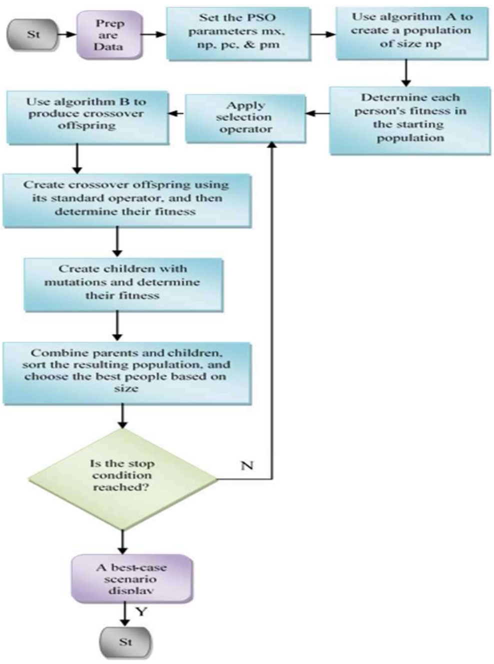

Fig. 2. Proposed KPSO to approximate p-HLAP solution

The original inhabitants are what PSO uses to try to enhance them. Mutation and mutation regulators are used for this. A novel transformation function is also proposed, and Fig.2 illustrates it. To do this, a criterion is established using (4) and (5).

$$rh_x = d_{xk}\left(\sum_y^X f_{xy} + f_{xy}\right) x \in A_k \quad (4)$$

Where A_k are all nodes that are allocated to hub k. A node with the max rh is chosen for changing its hub.

$$rn_x = d_{xk}\left(\sum_y^X f_{xy} + f_{xy}\right) k \in N_k \quad (5)$$

IV. Results and Discussions

The effectiveness of the CP composition over MIP and the KPSO composition over pure PSO are contrasted. The CP and KPSO are then examined along with their benefits and drawbacks. IBM ILOG v12.8.0 is used to code and solve the MIP and CP formulations to compare their performance to that of the MIP formulation. The main dataset with 200 nodes and 8 Storage, 11 preconfigured smaller examples, and a C programme file for producing smaller instances are all included in this dataset. We also utilized the programme file to produce new incidences.

A. Comparing study

Comparisons are made between the superiority of the CP composition over MIP and the KPSO composition over pure PSO . The advantages and disadvantages of the CP and KPSO are then discussed. To compare the performance of the MIP formulation to that of the CP formulation, the MIP and CP formulations are coded and solved using IBM ILOG v12.8.0. The outcomes are compared using a variety of factors, such as runtime, performance depends, and convergence rate. The Comparison of proposed system with existing one is represented in Table I.

TABLE. I. Comparison of proposed system with existing one.

No.	Optimal Objective	MP			CP		
		PSOP (%)	Runtime (s)	Status	PSOP (%)	Runtime (s)	Status
1	13658023	0.00	8	Optimal	0.00	15	Optimal
2	15321232	0.00	88	Optimal	0.00	2193	Optimal
3	15155568	0.00	600	Optimal	0.00	3700	Timeout
4	15236480	0.00	1930	Optimal	0.00	3700	Timeout

TABLE. II. Comparison of performance measures of proposed system with existing one.

No	Optimal Solution	PSO				KPSO			
		Initial Solution	Final Solution	PSOP (%)	Runtime (S)	Initial Solution	Final Solution	PSOP (%)	Runtime (S)
1	1359515	1831547	1356120	0.57	40	1359515	1360257	0.00	32
2	1515486	2038166	1588264	2.89	98	1654823	1603572	5.77	88
3	1513570	2465847	1595712	2.24	251	1513570	1521571	0.00	200
4	1532519	2412063	1735585	11.13	580	1564712	1532693	1.38	416

The degree to which a solution can be applied in the actual world and the extent to which expert requirements are satisfied is how we characterized the quality of solutions. For instance, a professional would anticipate that all 3 Containers will be utilized when 3 are taken into account. An additional illustration is that it is preferable to commit neighborhood networks to the same repository when the costs of doing so are equivalent for the 2 additional storage. Table II illustrates the findings.

Additionally, it becomes clear that KPSO is far superior to the PSO after meeting the stop condition. Further evidence that KPSO is trustworthy even in big instances of the problem comes from the fact that the PSOP percent of PSO grows quicker as the problem's magnitude increases than KPSO. For instance, in problem test 15, the PSOP percent of KG is fourth of KPSO, demonstrating the dependability of KPSO. The length of each algorithm's runtime is shown in Fig.3. As can be observed, KPSO takes less time to operate than PSO, which is attributable to the newly incorporated adaptive crossover operator. Additionally, the time required to produce continuous integration using an algorithm or at random takes about the same amount of time. A solution that is simpler to implement in practice is chosen when comparing the KPSO and PSO, taking into account the solution's overall quality and the runtime and final answer. An eye study of the Storage network that a solution has proposed serves as the basis for this qualitative analysis

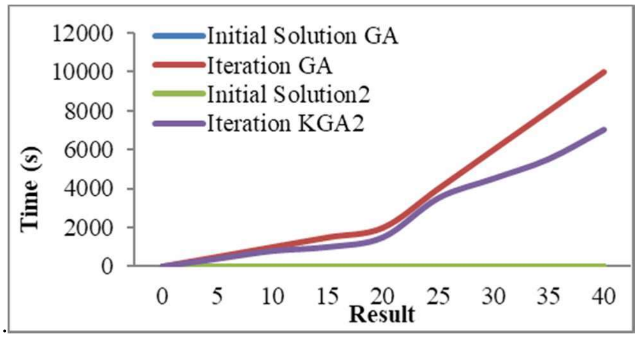

Fig. 3. Execution time.

Those networks that are assigned to storage are shown in Fig.4 in coordinates pages like 7 and 15, where each icon represents a node that is assigned to the same store. A Memory has no assigned nodes. To another Memory are given four connections. A third Storage is designated for all other nodes. The connections between the four blue circle nodes are allotted to separate Storage, which is another issue of quality relevance. The KPSO system is structured so that nearby node is allotted to the same Storage, however. We now have a Storage network that is more structured as a result.

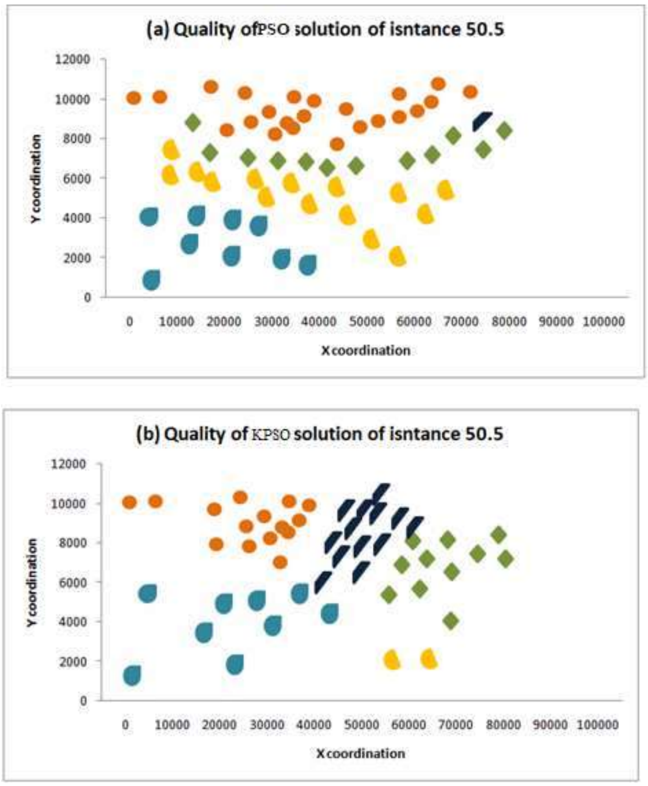

Fig. 4. Quality of (a) PSO and (b) KPSO solution

V. CONCLUSION

When compared to the MIP paradigm, the proposed CP formulation is simpler and easier to comprehend. One key difference from the MINLP paradigm is that the parameters and requirements are scaled down, causing them to expand linearly rather than dramatically as the number of nodes rises. Improved memory use demonstrates its effects. The testing findings showed that this model allows us to answer medium-sized issues, whereas MIP could only handle problems with up to 30 nodes. However, this technique cannot be faster than MIP in terms of runtime. Additionally, we expanded the CP formulation to include single, multi-allocation, and restricted capacity p-HLPs. K-means is a well-known machine learning approach for data clustering, and it is utilized in this case to produce preliminary PSO solutions. We built an algorithm to pick Storage and utilized K-means to cluster data into k groups based on the criterion of x and y coordination of nodes. Additionally, a novel crossover operator is created using this method as inspiration. According to the experimental findings, KPSO outperforms PSO in terms of solution quality, objective, and response time.

REFERENCES

[1] Ikotun, A. M., Almutari, M. S., &Ezugwu, A. E. (2021). K-Means-Based Nature-Inspired Metaheuristic Algorithms for Automatic Data Clustering Problems: Recent Advances and Future Directions. Applied Sciences, 11(23), 11246.

[2] Kaswan, K. S., Dhatterwal, J. S., &Balyan, A. (2022, April). Intelligent Agents-based Integration of Machine Learning and Case Base Reasoning System. In 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) (pp. 1477-1481). IEEE.

[3] Yunita, A., Santoso, H. B., &Hasibuan, Z. A. (2022). 'Everything is data': towards one big data ecosystem using multiple sources of data on higher education in Indonesia. Journal of Big Data, 9(1), 1-22.

[4] MuruPSOppan, E., Subramanian, N., Rahman, S., Goh, M., & Chan, H. K. (2021). Performance analysis of clustering methods for balanced multi-robot task allocations. International Journal of Production Research, 1-16.

[5] YarlaPSOddaa, J., &Malkapuram, R. (2020). Influence of MWCNTs on the Mechanical Properties of Continuous Carbon Epoxy Composites. Revue des Composites et des MatériauxAvancés, 30(1).

[6] Ren, Z., Zhai, Q., & Sun, L. (2022). A Novel Method for Hyperspectral Mineral Mapping Based on Clustering-Matching and NonnePSOtive Matrix Factorization. Remote Sensing, 14(4), 1042.

[7] Toorajipour, R., Sohrabpour, V., Nazarpour, A., Oghazi, P., &Fischl, M. (2021). Artificial intelligence in supply chain management: A systematic literature review. Journal of Business Research, 122, 502-517.

[8] BalamuruPSOn, K., Kuppusamy, A., Latchoumi, T. P., Banerjee, A., Sinha, A., Biswas, A., & Subramanian, A. K. (2022). Multi-response Optimization of Turning Parameters for Cryogenically Treated and Tempered WC–Co Inserts. Journal of The Institution of Engineers (India): Series D, 1-12. https://doi.org/10.1007/s40033-021-00321-x

[9] Sariyer, G., Mangla, S. K., Kazancoglu, Y., Xu, L., &Tasar, C. O. (2022). Predicting Cost of Defects for Segmented Products and Customers Using Ensemble Learning. Computers & Industrial Engineering, 108502.

[10] Venkatesh, A. P., Latchoumi, T. P., ChezhianBabu, S., BalamuruPSOn, K., PSOnesan, S., Ruban, M., &Mulugeta, L. (2022). Multiparametric Optimization on Influence of Ethanol and Biodiesel Blends on Nanocoated Engine by Full Factorial Design. Journal of Nanomaterials, 2022. https://doi.org/10.1155/2022/5350122

[11] Lee, I., &ManPSOlaraj, G. (2022). Big data analytics in supply chain management: A systematic literature review and research directions. Big Data and Cognitive Computing, 6(1), 17.

[12] Tutak, M., &Brodny, J. (2022). Business Digital Maturity in Europe and Its Implication for Open Innovation. Journal of Open Innovation: Technology, Market, and Complexity, 8(1), 27.

[13] YarlaPSOddaa, J., & Ramakrishna, M. (2019). Fabrication and characterization of S glass hybrid composites for Tie rods of aircraft. Materials Today: Proceedings, 19, 2622-2626.

[14] Bharathiraja, N., Padmaja, P., Rajeshwari, S. B., Kallimani, J. S., Buttar, A. M., & Lingaiah, T. B. (2022). Elite Oppositional Farmland Fertility Optimization Based Node Localization Technique for Wireless Networks. Wireless Communications and Mobile Computing, 2022. https://doi.org/10.1155/2022/5290028

[15] Bharathiraja, N., Shobana, M., Manokar, S., Kathiravan, M., Irumporai, A., & Kavitha, S. (2023). The Smart Automotive Webshop Using High End Programming Technologies. In Intelligent Communication Technologies and Virtual Mobile Networks (pp. 811-822). Springer, Singapore. https://doi.org/10.1007/978-981-19-1844-5_64

[16] Latchoumi, T. P., Swathi, R., Vidyasri, P., &BalamuruPSOn, K. (2022, March). Develop New Algorithm To Improve Safety On WMSN In Health Disease Monitoring. In 2022 International Mobile and Embedded Technology Conference (MECON) (pp. 357-362). IEEE. doi: 10.1109/MECON53876.2022.9752178.

[17] PSOrikapati, P. R., BalamuruPSOn, K., Latchoumi, T. P., & Shankar, G. (2022). A Quantitative Study of Small Dataset Machining by Agglomerative Hierarchical Cluster and K-Medoid. In Emergent Converging Technologies and Biomedical Systems (pp. 717-727). Springer, SinPSOpore. https://doi.org/10.1007/978-981-16-8774-7_59

[18] He, C., & HQ Ding, C. (2021). Predicting Partner's Digital Transformation Based on Artificial Intelligence. Applied Sciences, 12(1), 91.

[19] Latchoumi, T. P., Kothandaraman, R., &BalamuruPSOn, K.. (2022). Implementation of Visual Clustering Strategy in Self-OrPSOnizing Map for Wear Studies Samples Printed Using FDM. Traitement du

[20] Rajagopal Sudarmani, Kanagaraj Venusamy, Sathish Sivaraman, Poongodi Jayaraman, Kannadhasan Suriyan, Manjunathan Alagarsamy, "Machine to machine communication enabled internet of things: a review", International Journal of Reconfigurable and Embedded Systems, 2022, 11(2), pp. 126-134

[21] Roselin Suganthi Jesudoss, Rajeswari Kaleeswaran, Manjunathan Alagarsamy, Dineshkumar Thangaraju, Dinesh Paramathi Mani, Kannadhasan Suriyan, "Comparative study of BER with NOMA system in different fading channels", Bulletin of Electrical Engineering and Informatics, 2022, 11(2), pp. 854–861.

[22] AbithaKumariDuraisamy, Raja Guru Ramaraj, MathankumarManoharan, ManjunathanAlagarsamy, "Certificateless

linkable ring signature-based blockchains for securing cognitive radio networks", Concurrency and Computation: Practice and Experience, 2022, pp.e7235, John Wiley & Sons, Inc.