BAT Algorithm for Improving Fuzzy C-Means Clustering for Location Allocation of Rural Kiosks in Developing Countries under E-Governance

[1]Rajan Gupta, [1]Sunil K Muttoo and [2]Saibal K Pal
[1]Department of Computer Science, Faculty of Mathematical Sciences,
University of Delhi, Delhi, India
[2]Scientific and Analysis Group Lab, DRDO, Metcalfe House, Delhi, India
guptarajan2000@gmail.com

Abstract

Rural Kiosks are important infrastructural pillar in rural regions for internet and basic technology facility all around the world. They are also known as Tele-centers or Common Service Centers and are majorly used by government to promote Electronic Governance. The major characteristic of setting up of Rural Kiosk is their appropriate location so that people from rural region can avail the services at minimum travel cost and time. There are lot of traditional schemes used by researchers in past for location allocation but this paper proposes the usage of Fuzzy C-Means clustering and BAT algorithm to optimize the location of Rural Kiosk. The meta-heuristic approach has produced better results as compared to normal graph theories in past. The experiment has been conducted on a random data set of 72 village locations from India and their clusters are formed. It is found that using only Fuzzy C-Means clustering to allocate the center and by using it in combination with BAT algorithm produced up to 25% of efficient results. This can drastically help the key stakeholders in allocation of these Rural Kiosks at right places so as to maximize their utility.

Keywords: *BAT Algorithm, Location Allocation, Rural Kiosks, Fuzzy C-Means, Clustering, E-Governance, Tele-centers, Common Service Centers*

1. Introduction

Rural Kiosks or Tele-Centers or Common Service Centers, interchangeable terms used worldwide, are the centers to access internet facilities for public usage [1]. It can be used for accessing E-Government services or for different purposes in various regions. Their main aim is to connect rural region with the urban region and bring all the people on equal platform in terms of infrastructure as everybody can't afford to individual setup in developing nations. In countries like India and Africa, these Rural Kiosks becomes even more important as the population of rural regions is very high and thus more care has to be taken for people in these countries [2].

The Rural Kiosk can be a small center having basic facilities like Computer System, Internet, Telephone, Fax and photocopying machine. These are used to get connected to the world and thus give a basic infrastructure to the people living in rural regions. Although their main purpose is to provide facility to the people from rural regions, but their location becomes

an important criterion while setting them up in various states, districts and regions. In case the travel time and travel cost is not appropriate for the users living in the region of establishment, the citizens may lose interest. This can defeat the whole purpose of setting up of the Rural Kiosks to bring in technology access related equality.

The world-wide Tele-center movement put Tele-centers on the world map only due to the fact that they were able to come together as a global network, and in many cases, much more strongly as national networks of Tele-centers. The Philippine Community e-Centers Network (PhilCeCNet), the Thaitelecentre.org, Tele-center-Europe, Colombian National Tele-center Network, Chilean National Network of Tele-centers (ATACH) and the Sri Lankan Tele-center Family are only a few of the many such efforts around the world. The United Nations Economic Commission has helped to establish regional networks of Tele-centers in various regions of the world, with the Asian one, the Asia-Pacific Tele-center Network (APTN) currently housed in Thailand [3].

Within India too, the Common Service Centers have been planned as the key infrastructural aspect for delivery of effective E-Governance services. These centers are integral part of the National E-Governance plan but due since today the roll out phase has not completed and these centers are inefficient wherever they are opened. Some of them are not working effectively due to inadequate facility and others are not working due to inappropriate locations [4].

Therefore, the aim of this study is to study the optimization of location allocation of these Rural Kiosks or centers so that their implementation benefit can reach out to maximum people. Unless their location is optimized with respect to cost and time, it would be difficult for the centers to run successfully. Some schemes have been studied by researchers in past in the area of facility location allocation but less work exists in respect to Rural Kiosks. This study will first discuss various researches done in past in this area followed by existing schemes to solve facility location allocation. Then the proposed scheme and its experimental results are mentioned followed by discussion and conclusion.

2. Literature Review

There are different studies that took place in the area of location allocation of Rural Kiosks or Tele-centers. The first of such study was carried out by Kolokolov & Zaozerskaya [5] in 2006 on the Greek Islands. These authors used the discrete optimization problem approach to allocate the location to the Tele-centers. They considered the set of finite and countable options for the location of Tele-centers and selected the most optimized location based on cost and distance. Similarly, Naik et al [6] also focused on the financial aspect of the center in the rural regions of India. The total cost for the functioning of Common Service Centers or Tele-centers was found and it was concluded that Government to Citizen Services cannot be the only criterion for the sustainability of the centers. And if these centers are not opened at convenient location for the citizens, even G2C aspect of finances will not be able to work and fetch revenue from citizen's point of view. So from the government's viewpoint, Tele-centers location will play an important role.

Shourijeh & Kermanshah [7] in 2012 worked upon the location allocation of Tele-centers in Tehran, Iran. It was a mathematics oriented model for the optimization of location

within the particular region of Tehran. Linear program (Bi-objective) was developed which primarily considered primitive structured fuzzy goal programming method. The region was divided into various districts for which centroids were calculated for the allocation of the Tele-centers. The telecommuters were given importance and their distance from the Tele-centers was considered as key input for solving the problem. However, the minimum and maximum capacity parameters were kept constant in their solution which was not the real world situation while solving such problems. They proposed to extend the work to a more dynamic solution.

Kolokolov & Zaozerskaya [8] in 2013 presented the extension of their work done in 2006 for bi-criteria based center location allocation optimization. Maximization of efficiency and minimization of the cost were considered as the two criterions for location allocation. Set cover approach was used and linear integer programming was the technique used in deriving the solution along with the L-partition approach. The subsets of the pareto-optimal sets were found using the decomposition, trade-off method and the LCE algorithm.

Asik [9] in 2014 worked on the optimization of the location allocation of Tele-centers in rural regions of Turkey. He discussed the importance of sustainability of the Tele-centers and how it should equally serve majority of the persons in a region efficiently. Set cover problems were used along with p-median and p-center solutions. The author appropriately discussed the advantages and disadvantages of the various techniques and also gave a decent background about different techniques used to solve facility location allocation.

Apart from Tele-center or Rural Kiosks location allocation optimization, there are lots of researchers in the past who worked on other different kinds of facility locations [10, 11, 12, 13, 14]. But, with advent of more dynamic solutions and better algorithms, the allocation of Rural Kiosks can be better optimized and studied for more efficiency. In this regard, Kolokolov & Zaozerskaya [5] proposed the usage of evolutionary algorithms for location allocation optimization problems in future. Other studies also proposed to do the same and hence this paper focuses upon the Rural-kiosk's location allocation using the meta-heuristic algorithm. BAT algorithm is used in conjunction with the Fuzzy C-means clustering which is a prominent clustering technique used in the area of location allocation in past by various researchers [15, 16, 17].

3. Existing Scheme

There are various existing schemes in the area of Facility location allocation but the scope of this paper is limited to the discussion of FCM clustering [15] and BAT algorithm [18] for which the algorithms are described as follows. BAT algorithm has already been proved to be a better meta-heuristic algorithm [19,20,21] and it has performed well as compared to other intermittent search strategies [22]. Moreover its performance with clustering algorithms also looks encouraging [23, 24], so this particular scheme is restricted for the usage in this study from experimental point of view.

Algorithm 1. Fuzzy C Means Clustering

Initialize cluster center, and stopping criteria which is to minimize the following objective function

$$J_m = \sum_{i=1}^{n}\sum_{J=1}^{c} \mu_{ij}^{m} \left\|x_j - v_i\right\|^2$$

Initialize a random fuzzy partition matrix.

Calculate the cluster center value with following Equation

$$v_i = \frac{\sum_{j=1}^{n} {\mu_{ij}}^{m} x_j}{\sum_{j=1}^{n} {\mu_{ij}}^{m}}, \qquad 1 \le i \le c,$$

For each cluster

Calculate the membership function value in matrix using Equation

$$\mu_{ij} = \left[\sum_{k=1}^{c}\left(\frac{\|x_j - v_i\|^2}{\|x_j - v_k\|^2}\right)^{1/(m-1)}\right]^{-1}, \; 1 \le i \le c, 1 \le j \le n$$

End

Update membership matrix values.

If $||Membership_{Current} - Membership_{New}|| \le \in (0.001)$

Then stop otherwise go to cluster center calculation step

Where, J_m is sum of square error function for all the fuzzy clusters which is represented by membership matrix and related set of cluster center; μ_{ij} is the degree of membership for datapoint x_j in the i^{th} cluster; v_i is the cluster center and m as the constant valuethat controls the influence of membership in performance index which has been taken as 2.

Algorithm 2. BAT Algorithm

Objective function $f(X), X = (x_{1,\ldots,} x_d)\, T$

Initialize the bat population X_i $(i = 1, 2, \ldots, n)$ *and* V_i

Define pulse frequency f_i *at* X_i

Initialize pulse rates r_i *and the loudness* A_i

While (t <Max number of iterations)

Generate new solutions by adjusting frequency, and updating velocities and locations/solutions

$$fi = fmin + (fmax - fmin)\,\beta,$$
$$v_i^t = v_i^{t-1} + (x_i^t - x_*)fi$$
$$x_i^t = x_i^{t-1} + v_i^t,$$
$$X_{new} = X_{0ld} + \epsilon A^t,$$
$$A_i^{t+1} = aA_i^t, \; r_i^{t+1} = r_i^o\,[1 - exp(-yt],$$
$$A_i^t \to 0, \; r_i^t \to r_i^o, \; as\; t \to \infty$$

if $(rand > r_i)$

Select a solution among the best solutions

Generate a local solution around the selected best solution

end if

Generate a new solution by flying randomly

if $(rand < A_i \;\&\; f(x_i) < f(X_*))$

Accept the new solutions

Increase r_i *and reduce* A_i

end if

Rank the bats and find the current best X_*

end while

Postprocess results and visualization

Where, Each bat is associated with a velocity v_i^t and a location x_i^t, at iteration t, in a d-dimensional search or solution space. Among all the bats, there exists a current best solution $x*$. A_i^t is the loudness and r_i^t is the pulse rate of bat at iteration *t*.

4. Proposed Scheme

The proposed scheme combines the clustering and meta-heuristic technique to derive the result.

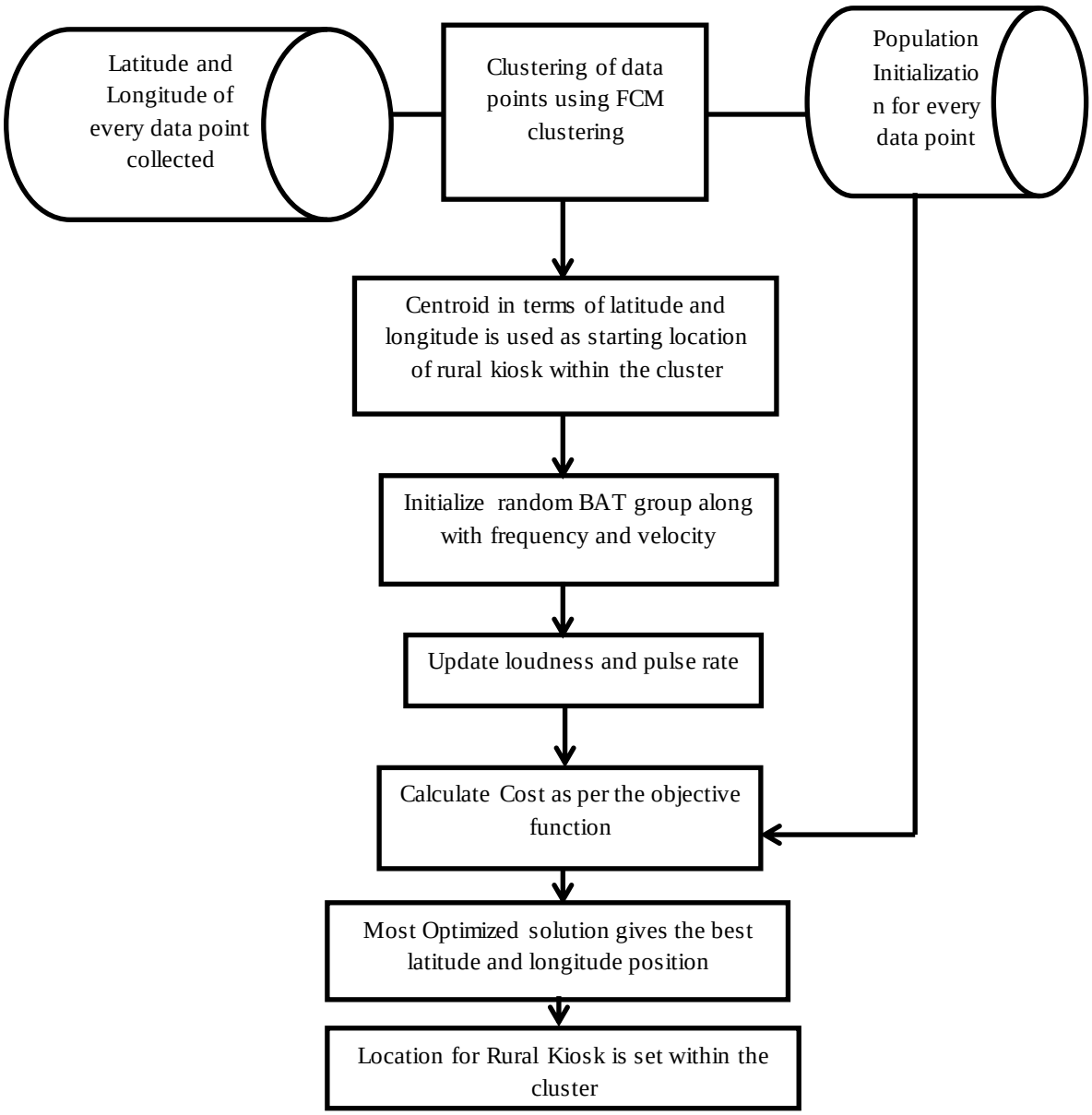

Figure 1. Proposed scheme combining the FCM Clustering and BAT algorithm

In the proposed scheme, as shown in Figure 1, the central idea is to initialize the Fuzzy population with the village positions (latitude and longitude) and they are clustered based on near-ness of their location calculated through membership function. Once the clusters of villages (6 villages per cluster) are formed, their centroid based on latitude and longitude

positions are calculated. This centroid becomes the point of interest for location allocation of Rural Kiosk. Traditionally the centroid should be the rural kiosk allocation point, but it may not be a good option if village population varies in that cluster. So to satisfy the majority of people in the cluster, the population and centroid are initialized in the BAT algorithm. This centroid keeps on changing as per the BAT characteristics and objective function is calculated w.r.t. travelling cost which is as follows. Once the minimum cost function is achieved, that location of Rural Kiosk is finalized for that particular cluster.

*Travel Cost = Population*Travel Cost per unit*Summation of Village distance from Kiosk position*

5. Results & Discussion

The schemes were implemented in MATLAB 2010a version software. A sample location of villages was taken from one of the developing country – India in this experiment. Since, India is aggressively moving towards digitization and electrification of villages [25], and improvement of E-Governance has also led to the set-up of various Rural Kiosks (Common Service Centers in Indian terminology), so optimizing the location allocation of the rural kiosks of such data points would be more beneficial. A total of 72 village points were considered for the experiment from (http://www.mapcoordinates.net/en) whose latitude and longitude on map were considered for the calculations. These 72 villages were clustered in 12 clusters of 6 villages. This is as per the Indian Government's national E-Governance plan according to which a cluster of 6 villages will be considered for set up of one Rural Kiosk in that cluster [26].

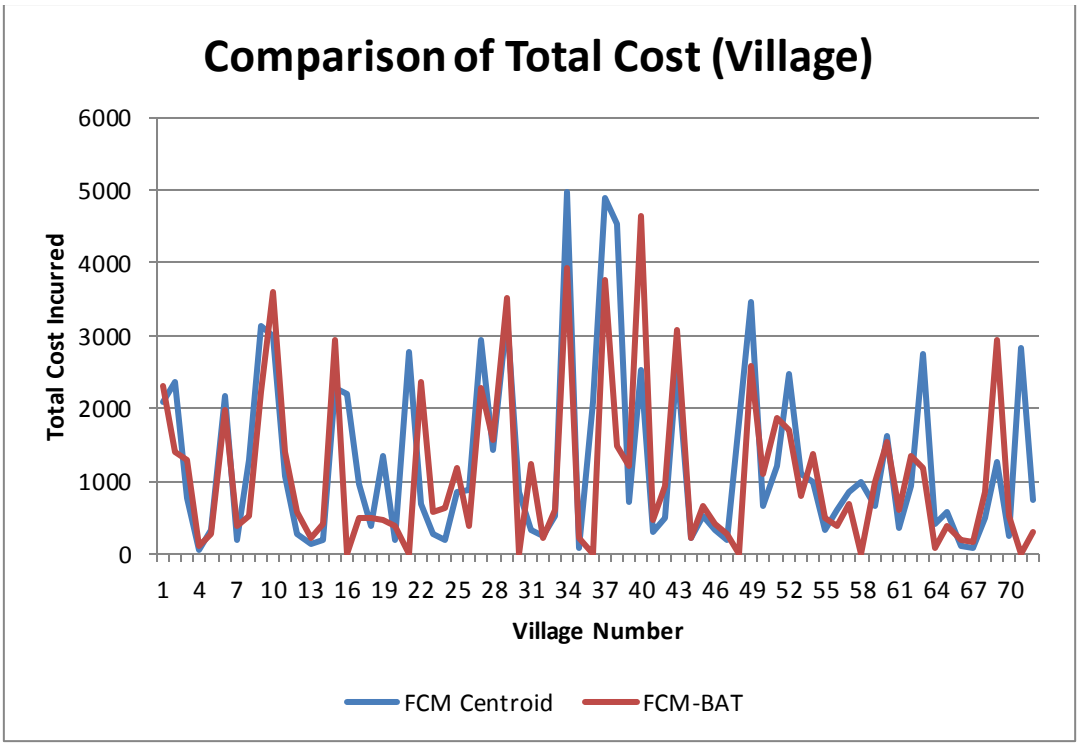

Figure 2. Comparison of Total Cost incurred by 72 villages from two different schemes of Rural Kiosks

As shown in Figure 2, total cost for every village was calculated from both the schemes and compared against each other. For some villages the FCM-BAT scheme had higher cost, but for others it was drastically down as compared to the cost calculated by FCM-Centroid scheme. Although on individual basis the BAT algorithm inclusion could not reduce the results for every data point, but it made a significant impact on the overall cluster's cost as shown in Figure 3.

As shown in Figure 3, the total cost of 12 clusters is compared against the two schemes. It is clearly visible that the cost calculated by FCM-BAT scheme is lower in each case as compared to FCM-Centroid. The cost reduction is just not marginal but it is significant in almost every case. The percentage change in cost for every cluster varies from 4% to 26% which is very encouraging. There were almost 7 clusters found for which the percentage change in cost was more than 15%.

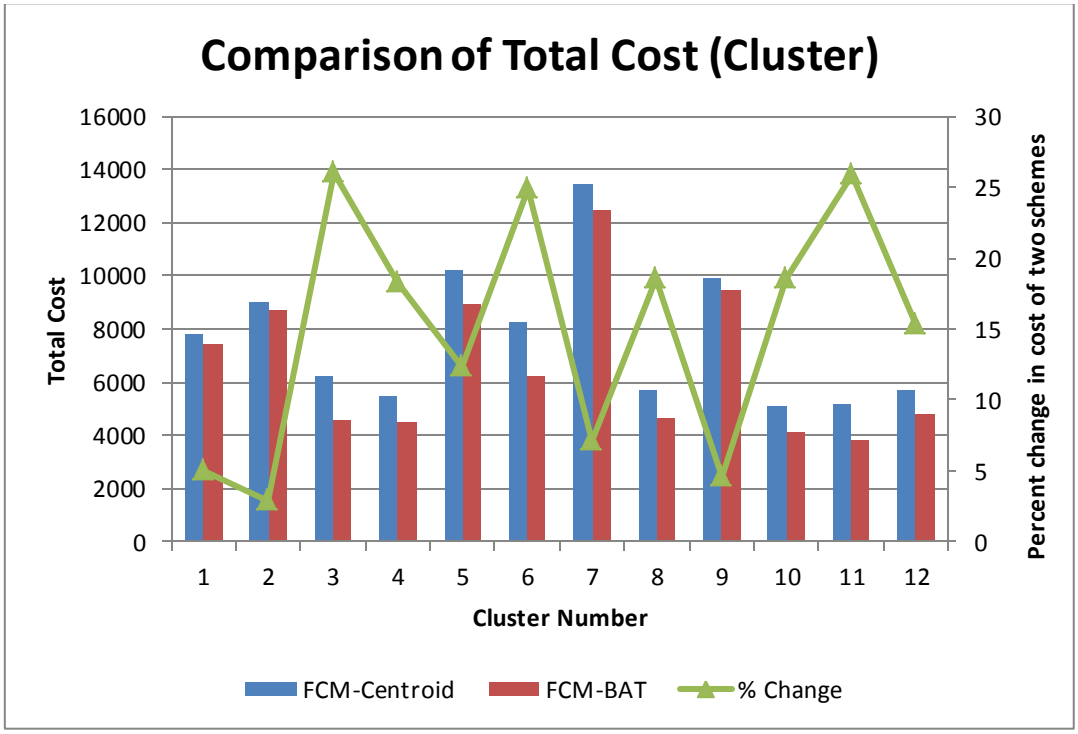

Figure 3. Comparison of Total Cost incurred by 12 village clusters from two different schemes of Rural Kiosks

BAT scheme has the ability to overcome the weakness of FCM clustering i.e. getting trapped with local maxima. Various meta-heuristic schemes have been designed to overcome the local maxima problem and BAT is one of those successful schemes. It has the ability to produce globally optimized results which can overcome the local maxima problem. Moreover, BAT algorithm is superior to other meta-heuristic algorithms because of its ability to zoom into the area of optimized solution and thus producing the best results. It also has the ability to find optimized solution using swarm-intelligence but with variable values for control parameters unlike other such category of algorithms [18].

BAT algorithm has proven to be an optimization option for clustering by different authors in past [19, 20, 21, 22]. It has been found useful for optimizing fuzzy clustering schemes and even K-means clustering schemes [23, 24]. But within the location allocation research, usage of BAT algorithm has been found very rare and previous researchers have only focused on the usage of meta-heuristic schemes in their proposed work. This study has used BAT algorithm with FCM clustering and found efficient results as compared to the various other location allocation scheme solutions from the past.

6. Conclusion

Rural Kiosks are an important infrastructural component for developing regions around the world. Every developing country needs to take efficient measures in order to give access of internet to the rural areas. If government wants to transform their process digitally, there is definitely a need to set up the infrastructure in such a way that people can actually access it. Clustering is general method when resource allocation is considered in a particular region but it may not produce the most optimum results as shown in this study. It can get trapped in the problem of local maxima and thus overall optimized solution may not be achieved. This study proposed the combination of BAT algorithm and FCM clustering scheme to produce most optimized results in terms of travelling cost for people to visit the Rural Kiosk and access the service. It should be beneficial to majority of population of the cluster. The results shows an improvement of up to 25% for the cost saving and thus more people would be encouraged to use the rural kiosk services.

Future scope of this work includes testing of the clustering technique with other advanced meta-heuristic schemes and analyzing the optimization within the clusters. Also, since the scope of study was limited to few data points from Indian region, more number of data points across different countries can be studied w.r.t. to their latitude and longitude position on the map.

References

[1] Toyama, K., &Kuriyan, R. 2007. "Review of research on rural PC kiosks." Microsoft Research India.

[2] Kendall, J., & Singh, N. 2012. "Performance of Internet Kiosks in Rural India Gender, Caste and Location." Review of Market Integration, 4(1): 1-43.

[3] Shadrach, B., & Sharma, S. 2013. "Impact assessment of Indian common services centres." Ministry of Communication and Information Technology, Government of India, New Delhi.

[4] Dass, R., &Bhattacherjee, A. 2011. "Status of Common Service Center Program in India: Issues, Challenges and Emerging Practices for Rollout(No. WP2011-02-03)."Indian Institute of Management Ahmedabad, Research and Publication Department.

[5] Kolokolov, A. A., &Zaozerskaya, L. A. 2006, May. "A bicriteria problem of optimal service centers location." In Information Control Problems in Manufacturing 12(1): 425-429.

[6] Naik, G., Joshi, S., &Basavarajappa, K. P. 2010. "Making E-Governance centers financially sustainable in rural India: A conceptual design for action research." IIM Bangalore Research Paper, (317).

[7] Shourijeh, M. T., Kermanshah, M., Mamdoohi, A. R., Faghri, A., &Hamad, K. 2012. "A Mathematical Optimization Model for Locating Telecenters."Applied Mathematics, 3: 251-263

[8] Kolokolov, A. A., &Zaozerskaya, L. A. 2013. "Solving a bicriteria problem of optimal service centers location." Journal of Mathematical Modelling and Algorithms in Operations Research, 12(2): 105-116.

[9] Asik, O. 2014. "Location Optimization To Determine Telecenter Network In Rural Turkey." Doctoral dissertation, Cornell University.

[10] Current, J., Min, H., & Schilling, D. 1990. "Multiobjective analysis of facility location decisions." European Journal of Operational Research, 49(3): 295-307.

[11] Wen, M., &Iwamura, K. 2008. "Fuzzy facility location-allocation problem under the Hurwicz criterion." European journal of operational research, 184(2): 627-635.

[12] Hajipour, V., Fattahi, P., Tavana, M., & Di Caprio, D. 2015. "Multi-Objective Multi-Layer Congested Facility Location-Allocation Problem Optimization with Pareto-based Meta-heuristics." Applied Mathematical Modelling.

[13] Mestre, A. M., Oliveira, M. D., & Barbosa-Póvoa, A. P. 2015. "Location–allocation approaches for hospital network planning under uncertainty." European Journal of Operational Research, 240(3): 791-806.

[14] Pereira, M. A., Coelho, L. C., Lorena, L. A., & De Souza, L. C. 2015. "A hybrid method for the Probabilistic Maximal Covering Location–Allocation Problem." Computers & Operations Research, 57: 51-59.

[15] Küçükdeniz, T., Baray, A., Ecerkale, K., &Esnaf, Ş. 2012. "Integrated use of fuzzy c-means and convex programming for capacitated multi-facility location problem." Expert Systems with Applications, 39(4): 4306-4314.

[16] Esnaf, Ş., Küçükdeniz, T., &Tunçbilek, N. 2014. "Fuzzy C-Means Algorithm with Fixed Cluster Centers for Uncapacitated Facility Location Problems: Turkish Case Study." In Supply Chain Management Under Fuzziness (pp. 489-516). Springer Berlin Heidelberg.

[17] Shi, Q. S., &Zheng, X. Z. 2014, May. "Electric Vehicle Charging Stations Optimal Location Based on Fuzzy C-means Clustering." In Applied Mechanics and Materials, 556: 3972-3975.

[18] Yang, X. S., &He, X. 2013. "Bat algorithm: literature review and applications." International Journal of Bio-Inspired Computation, 5(3): 141-149.

[19] Shieh, C. S., Horng, M. F., Ngo, T. G., & Dao, T. K. 2015. "Unequal Clustering Formation Based on Bat Algorithm forWireless Sensor Networks." In Knowledge and Systems Engineering (pp. 667-678). Springer International Publishing.

[20] Cao, Y., Cui, Z., Li, F., Dai, C., & Chen, W. 2014. "Improved Low Energy Adaptive Clustering Hierarchy Protocol Based on Local Centroid Bat Algorithm." Sensor Letters, 12(9): 1372-1377.

[21] Nanda, S. J., & Panda, G. 2014. "A survey on nature inspired metaheuristic algorithms for partitional clustering." Swarm and Evolutionary computation,16: 1-18.

[22] Yang, X. S., Deb, S., & Fong, S. (2014). Bat algorithm is better than intermittent search strategy. arXiv preprint arXiv:1408.5348.

[23] Komarasamy, G., &Wahi, A. 2012. "An optimized K-means clustering technique using bat algorithm." European Journal of Scientific Research,84(2): 26-273.

[24] Sood, M., &Bansal, S. 2013. "K-Medoids Clustering Technique using Bat Algorithm." International Journal of Applied Information Systems (IJAIS) ISSN, 2249-0868.
[25] Asher, M. G. 2016. "The 2016-17 Budget: A Positive Step in India's Transformation Process." Lee Kuan Yew School of Public Policy Research Paper, (16-09).
[26] Mishra, G. 2014. "Telecentres as a Medium for Good Governance in Rural India." Handbook of Research on Democratic Strategies and Citizen-Centered E-Government Services, 56.