A Multi-Agent Bio- Inspired System to Map Learners with Learning Resources using Clustering Based Personalization

Nageswara Rao Gottipati, A. Rama Prasath

Abstract*: The work makes an examination of clustering methods in a multi-agent system which is fully decentralized. This has the goal of grouping agents that have similar data or objectives as in the case of traditional clustering. But, this adds to some more additional constraints wherein the agent will have to be in the same place as opposed to being collected within a centralized database. For doing this, it will connect to agents within a random network and will search for them in a peer-to-peer based fashion for the other agents that are similar. The primary aim here was to tackle the basic problem in clustering on the Internet scale thus creating methods where the agents may be grouped thus forming coalitions. For the purpose of investigating the decentralized approaches and their feasibility, the work presents the K-means clustering, the multi-agent Firefly Algorithm, (FA) and the Differential Evolution (DE). This is done for a reasonable number to times and will be surprisingly good. The results of the experiment prove that the multi-agent firefly clustering has better performance compared to that of a multi-agent K-Means clustering or a multi-agent DE clustering.

Keywords*: Clustering, *Data Mining, Differential Evolution (DE), Multi Agent Systems (MAS), K-Means Clustering, Clustering and Firefly Algorithm (FA).

I. INTRODUCTION

Data Mining and the Multi-Agent Systems (MAS) have been technologies that are well-established that have been finding a lot of application. One present challenge in data mining was to cope with the size which is ever increasing. There may also be an answer to a high level which was adopting and applying a higher power of computation. This was further achieved in various types of manners. One such approach was to use either distributed or a parallel level of processing techniques to ensure several processors that are applied to this problem. This further assumes the fact that suitable "farms" of the processors were available. But another disadvantage here was a centralized control and its lack of generality. Both the distributed and the parallel techniques of data mining ideally assumed a new "master" process which will direct the task of data mining which was to control and this was centralized as the process and the system will as a consequence lack robustness. Furthermore, these parallel and distributed approaches are normally directed at certain specific applications of data mining and are also challenging to generalize (owing to the inherent centralized control to all other systems). The MAS offers another alternative to handle data of large quantities by means of harnessing the numbers of processors and their power and this has the additional advantage of control not being centralized. This makes all these systems more versatile and robust [1].

The MAS is a collection of the entities of the software (the agents) intended for cooperating in a manner in which a certain task of processing was undertaken. One more important aspect was that the agents had been behaving in a manner which was autonomous and they tend to negotiate with each other for completing the task as opposed to them being directed so by a master process. The primary idea that adopts the MAS technology for the purpose of data mining which was very attractive. The Multi-Agent Data Mining (MADM) thus permitting distributed data which has to be effectively mined without moving of the data to the data warehouse [2]. The MADM will thus support creating a framework which is allowed in an anarchic manner, where the agents are incorporated within the framework until such time they are able to comply with the protocols that are specified.

The agents were defined as the computer entities which were well-suited to the environment and so they can carry out these tasks autonomously. The actual environment where agents were operating was uncertain, unpredictable and dynamic. So, this is very desirable for the agents to be able to exhibit computer intelligence for interacting with this environment. The abilities that all intelligent agents were to display [3]:

- Reactivity – the intelligent agents need to anticipate based on their environment in order to give a timely response.
- Pro-activeness – All intelligent agents will have to explore some more alternatives for achieving the objectives of design.
- Social ability – the intelligent agents need to interact with the other agents for satisfying the objectives of design.

It is very evident that the capabilities above may be able to increase the autonomy of performance for the agents.

The MAS was a new software system employing various interactive agents for solving problems in an uncertain environment which is decentralized. The central feature of the MAS was it not having any control mechanism which was centralized; the agents also needed to collaborate achieving a design objective for a given MAS. The collective abilitiesare as follows [4]:

Revised Manuscript Received on August 20, 2019.
***** Correspondence Author
Nageswara Rao Gottipati *, Research Scholar, Dept. of CSE, Hindustan Institute of Technology & Science, Chennai, Tamil Nadu, India. Email: rao.nageswara@gmail.com
Dr.A.Rama Prasath, Assistant Professor(Sel.Gr), Dept. of MCA, Hindustan Institute of Technology & Science, Chennai, TamilNadu, India Email: mraprasath@gmail.com

- In the computational resources of the MAS with the capability of being distributed in the network of all interconnected agents in solving problems that are large for the interconnected agents. There is a centralized system that was plagued with some limitations of resources, critical failures or bottlenecks of performance.
- The MAS permits interaction and interconnection of various legacy systems; by means of building a new agent wrapper for these systems, they are incorporated into the agent society.
- In the case of a problem of the MAS, they are modelled as agents of autonomous interacting components working in a self-directed way.
- Using the MAS gives certain solutions where there is a need for spatially and temporally distributing them. In connection with the agents and their social ability, both resources and expertise were shared.
- Using the MAS can enhance the system and its performance in the dimensions of the efficiency of computation, flexibility, reuse, responsiveness, maintainability, extensibility and so on.

Personalized systems of learning and learning personalisation had become very popular in the field of scientific literature in recent years. Personalization was observed from two perspectives: personalisation and selection of a single learning object in a learning scenario or choosing the composition. For instance, the personalization for the learning scenario was by finding the learning path. The earlier perspective will learn the problem of object selection and the next one will solve the problem of curriculum sequencing. Some intelligent technologies like semantic web systems, ontologies and semantics were to learn an enhancement by means of personalization [5].

The primary goal to clustering was similar objects as clusters and at the same time, it separated all dissimilar objects. This being a technique of unsupervised classification where clustering is critical to different fields in data analysis, linguistic simulation, investigation of gene expression, community detection, processing of signals, data mining, and image segmentation and so on were considered. In the recent decade, there have been various algorithms that were developed to solve problems in clustering such as the G-N algorithms and the graph partitioning algorithms and this includes the algorithms of spectral clustering [6].

Aside from the types mentioned above, there had been some more algorithms based on the network dynamics analysis. A random walk algorithm was moving in a random trail where every step was next to the neighbouring vertex. The algorithms of the random walk were for detecting the structure of clustering by means of imbuing the network along with flows of a random walk. At the same time, for network synchronization, with a particular model of Kuramoto oscillator, there was an oscillation of synchronization that occurs within the cluster.

The methods of clustering are different based on properties like data type which is clustered but also as its final partitioning that is found in assumptions as regards the shape of their clusters along with the given parameters. Every algorithm in clustering will look out for clusters in accordance with another different criterion. This may also be an issue in several other datasets which may present other shapes and also the sizes of clusters which have one single objective clustering which has not been exposed. Furthermore, the same type of data may be able to have different relevant structures that is connected to a different definitions or level of refinements [7].

The new cluster of agents which agree on a new plan for solving problems or for reacting to a certain event that may be evident when displayed. Furthermore, when there are various events occurring in the control system along with other agents that have other behaviours that have been called upon. Here the relationships among the agents may change where all relationships may be identified and then visualized. The information on the coupling or the agent clustering may be used with an integrator or a system developer for mapping the hardware of agent execution. The placing of agents belonging to every single cluster within the computation unit which may be able to improve the performance of the system, in case the communication cost is within the unit of computation which may be lower than the communication that is found outside. This may be true in the MAS wherein the messages that are sent to that of the same execution unit that does not have to be either serialized or parsed[8].

In recent times, there are many algorithms that are biologically inspired which were introduced for solving the problem in clustering. The algorithms were further characterised by means of an interaction with a large set of simple agents interacting in the MAS. The agents may further perceive and then change their environment and they are also inspired using swarms of bees, flocks of birds and the ant colonies. For the purpose of this work, a multi-agent k-means, the FA-based clustering and the DE to the learning resource and their usage mining. The rest of the investigation has been organized thus. Section 2 explains the related work found in the literature. The different methods used in this work has been explained in Section 3. The experimental results were explained in Section 4 and the conclusion was made in Section 5.

II. RELATED WORKS

Ozaeta and Graña [9] had designed another new multi-agent dynamic system moving within the virtual space in accordance with the attractive and repulsive forces defined in a very complex structure. For the purpose of this approach, every agent was considered for not having any access to information with regard to the network and its general structure as it may also be non-attainable for having a certain network representation which is inside every agent. These links for the conditions of neighbours and the movement of agents by pulling it using certain attractive forces.

Melesko and Kurilovas [10] had made another presentation of a semantic clustering along with the Artificial Neural Network (ANN) that is based on the learning analytics of the software agent used for a personalized adaptive system of multi-agent learning.

In the first stage, there was literature review which was systematic that was on the application of the ANN along with the semantic clustering of a learning material found in a personalised adaptive system of multi-agent learning that was presented. For the purpose of this work, any personalization in a learning system had been based on the model of various learning styles which need using a psychological questionnaire for the determination of the learning styles of the students. The results proved that it may not be correct as certain students may be able to answer this questionnaire in an irresponsible and dishonest manner or even making mistakes in the process of self- diagnosis. This may result in creating a student model which is not correct. A modified clustering approach using k-means was proposed by Selvi and Sivasankar [11] for eliminating these issues. There was also a novel and supervised method of Adaptive Genetic Neural Network (AGNN) that had been proposed for locating the data points that are most favoured in the cluster for delivering some recommendations which are effective. The proposed system of recommender had been measured by means of conducting a new experimental analysis on the benchmark Movie Lens datasets and the Netflix datasets. Experiments then compared with other methods like ANN and the system based on the fuzzy recommender system in order to show the proposed method and its effectiveness. Lin et al., [12] had made a discussion of a new distributed design used for the K-means algorithm which was based on clustering found in a multi-agent network which was switching for cases wherein the data were decentralized and stored and also unavailable to the agents. There was a proposal of an algorithm based on consensus by the authors which was a Double-clock Consensus-based K-means Algorithm (DCKA). Using mild conditions of convergence of the DCKA for guaranteeing any distributed solutions to problems in clustering was made. Furthermore, the authors gave experimental results on the clustering datasets for illustrating the effectiveness of the DCKA. A new Genetic Algorithm (GA) that was based on a Multi-Agent Negotiation Scheme for the Group Recommender System (GA-MANS-GRS) had been proposed by Choudhary and Bharadwaj [13]. This modelled one-to-one schemes of bilateral negotiation. For the negotiation phase, the GA was applied in order to obtain utility for every user and this generated a proper ranking for every individual. For the phase of recommendation generation, the Ga had employed for producing a rating list which could minimize the sum of the distances among various experiments that had been presented to establish the proposed model's superiority over the techniques of the Group Recommender Systems (GRSs). Jin et al., [14] had discussed a clustering analysis by combining all the advantages of the Particle Swarm Optimization (PSO) found in clustering as the PSO had a good global search ability. The centre and the number of the clustering are first found and the results are optimized using the K-means algorithm that is combined with the optimization algorithm. All simulated experiments prove that this combining algorithm was superior to that of the other clustering algorithms since they were more efficient and robust.

İnkaya et al., [15] had presented another new methodology of clustering that was based on the Ant Colony Optimization (ACO-C). This ACO-C was employed based on a multi-objective setting and yielded solutions that were non-dominated. It consisted of two steps of pre-processing which are: neighbourhood construction and the reduction of the dataset. The former tends to extract all local traits of the data points and the latter is employed for its scalability. The results of the experiment proved that the ACO-C was able to outperform all the competing approaches. This mechanism that was multi-objective had been related to the neighbourhood and enhanced extraction of clusters of arbitrary shape with variations in density. Abbasian et al., [16] had also proposed another multi-objective version of the Gravitational Search Algorithm (GSA), called the Clustering based Archive Multi-Objective GSA (CA-MOGSA). This was created with the Pareto principles and it selected the solutions that were non-dominated and were stored in the external archive. For controlling the archive size, all solutions having less distance of crowding had been removed. The strategy further guaranteed diversity and elitism as their features in the multi-objective algorithms. This was clustered and such a cluster was chosen randomly for every agent. The choice of the cluster was based on the actual distance between the representatives of clusters and the population member (agent). So, there was a proper trade-off between the exploration and the exploitation. Fister [17] had further presented a multi-agent system that was based on the self-adaptive DE algorithm to solve problems that were dynamic. These observed the solution and its quality and changes to it over time. Such changes take place once there was a predefined number of generations complete that had online responses desired using evolutionary algorithms. For the purpose of this work, each problem was considered and solved with the help of multi-agent systems. Every such agent found in the system was implemented to be a self-adapted De that had a constant size of the population. As a result of this, the most suitable multi-agent system had been searched and the results of the multi-agent system that was the best were compared to the other state-of-the-art algorithms.

III. METHODOLOGY

The agents who did not want to cooperate with this system had to find one another. There was a solution that was straightforward and this created a new server to the central directory for matching requests. This, however, limited autonomy of the agents in connection to their selection of partners and also their scalability. In ideal conditions, the agents on their own will be able to identify all their potential partners (such as the members in their social circle) and also directly negotiated newer partnerships on the basis of further information compared to a directory server. This could be looked at as a problem in clustering [18]. For the purpose of this model, the agents had been characterized by means of their respective attributes and this was for clustering them as one set of items. Every agent in this had a certain number of links to that of the other agents and the links indicated the channels of communication which defined every agent's neighbourhood.

The primary aim of this system was for rearranging the links and for choosing some to form certain connections or some more connected links among agents thus generating a new graph of connections. The initial links, when connected were taken to be a bootstrapping problem and it was further assumed that they had been derived from the actual placement agents found in the network. They could also be based on a source that was application-dependent and modelled as a random one. For the purpose of this section, there was a general framework used for learning a resource usage, a multi-agent DE clustering, a multi-agent K-means clustering or a multi-agent firefly clustering-based methods.

3.1 Framework for Learning Resource Usage

The principal elements in the learning resource and its personalization are the modelling in the objects of web learning (such as pages) and subjects (like users) and the categorization of either the subjects or the objects. The general architecture for the personalization of a learning resource on the basis of content mining and usage was depicted in Figure 1. The entire process had been divided based on two different components: an offline component consisting of data preparation along with certain specific tasks of web mining and an online recommendation engine. The tasks of data preparation will result in various aggregate structures that consist of a pre-processed usage with content data that was used in the stage of mining [19]. The structures have a user transaction file that captures some meaningful units of semantics for the user activity. Also, only the page views that are relevant are taken to the file. Every page view indicates the collection of a set of Hyper Text Transfer Protocol (HTTP) based requests that contribute to one single display found in a user browser. Pertinent features are extracted from the text/meta-data available on the site and their weights for every page view.

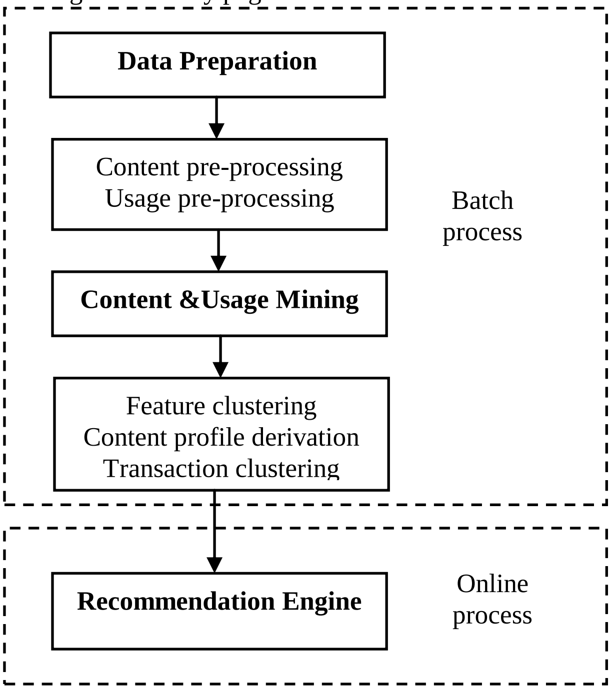

Figure 1 Framework for learning resource personalization based content mining and usage

Considering the pre-processed data given, there are several tasks in data mining that may be performed. For instance, the tasks of usage mining include the discovery of the rules of the association, patterns which are sequential, transaction clusters and page view clusters and sometimes methods of pattern discovery from various user transactions. The tasks and their content mining include feature clustering, a page view clustering and finally the actual discovery of association rules. For the purpose of this work, the focus was on deriving usage profiles from the transaction clusters and deriving content profiles from the feature clusters. The recommendation engine considers active server sessions with the patterns found and the other profiles for providing some personalized content. This may be in the form of products or links, texts, graphics or targeted advertisements.

3.2 Multi Agent K Means Clustering Algorithm

One of the most popular methods of clustering is the K-means clustering in cluster analysis aiming to divide n observations into the k clusters where every observation is part of the cluster to its nearest mean. This will be a new algorithm of partition clustering which is very effective for datasets of a small size. Firstly, k initial centres that are based on the desired cluster number is chosen. For this, a user will be able to specify a k parameter value. Every data point will be assigned to its closest centroid and the actual set of the points that are assigned to a centroid is known as a cluster. Every cluster centroid will then be updated on the basis of the points that have been assigned to this cluster. This process will then be repeated until such time the centroids stay as they are or there are no point clusters [20]. For the purpose of this algorithm, the Euclidean distance (as per equation (1)) is normally used for identifying the actual distance between the data points and the centroids.

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^{N} (x_{ik} - x_{jk})^2} \quad (1)$$

Wherein the $d(x_i, x_j)$ denotes the actual distance between the x_i and x_j. x_i and the x_j which are the attributes for a certain given object wherein the i, j, and k are different from 1 to N where it denotes the total attributes of this object given, indexes i, j, k and N all being integers. Based on these distances, there is a need to be able to generate a partition by means of assigning a sample to the cluster that is the closes. After this, a new cluster centre is computed as the centroids of clusters and the distances are also computed to generate a partition. This has to be repeated until such time the cluster membership reaches stability. This process is called the “K-means”, and gives efficient partitions in terms of variance of class. Further, the K-means process is computationally economical. It is a very feasible process for large samples. The K-means algorithm is probably the where the data analyst makes use of this for investigating another new dataset since it is simple arithmetically, robust and also provides some answers that are ‘good enough’ and answers various datasets.

The steps involved in K-Means clustering are [21]:

Step 1: Choose a value for 'k', as the cluster number.

Step 2: Calculation of the initial centroids in the actual sample of the dataset. Divide the data points into various 'k' clusters.

Step 3: Move all data points into the clusters by making use of the Euclidean's distance equation (1). Now recalculate all the new centroids. The centroids will then be calculated based on the means or average.

Step 4: Repeat Step 3 when there is no data point that is moved.

The algorithm makes use of a criterion of square-error to re-assigning a sample from one cluster to that of another and this results in a decrease in the squared error. The sum of the square-error value 'E' is depicted in equation (2).

$$E = \sum_{i=1}^{L}\sum_{j=1}^{M}\sum_{k=1}^{N}(F-C)^2 \qquad (2)$$

Wherein $(F - C)^2$ denotes the actual distance between data points.

The MAS has four agents of the K-means clustering for a generation of the initial centroids employing various methods. There is a discussion of the initial centroid generation that is based on an actual sample of data points that make use of the range method, the random number method, the outlier method and the inlier method [22].

3.3 Proposed Multi Agent Differential Evolution Clustering Algorithm

A multi-agent algorithm of clustering that is based on the DE chooses randomly the centre from a dataset of the encode clustering and will structure its initial population. It further performs the operation of a DE algorithm like the mutation operations, the selection operations, the crossover operations and so on for obtaining an optimum individual. The individual that is the best is decoded and further clustered to its best initial clustering centre that had been obtained. Given below are the details on the multi-agent clustering which is based on the DE [23].

The Initiation of population: here matrix X (N.d=D) is generated for storing data in the current population. N will express the population scale and this represents a combination style of the N group and its multi-agent clustering centre. D will express an individual dimension in the population. N will be set about 5-10 times to the d. After this, the multi-agent data and their samples were chosen randomly from the centralized data will serve as an individual of the initial population. This is repeated n times where initial population structure is as per equation (3).

$$X_i(t) = (x_{i1}(t), x_{i2}(t), \ldots, x_{ij}(t), \ldots, x_{iK}(t)) \; (i = 1, 2, \ldots, N) \qquad (3)$$

x_{ij} express clustering centre j for individual i.

The Mutation operation: this is based on the vector difference of the individual in its current population. Suppose X_i (t) is an individual in its current population,then three different individuals X_l (t), X_m (t), X_n (t) is chosen randomly. The vector difference existing between the individual X_m (t) and X_n (t) is the disturbance factor which is scaled by a factor and thus the mutated individual is obtained (4):

$$U_i(t) = (u_{i1}(t), u_{i2}(t), \ldots, u_{ik}(t)) \; .$$

$$u_{ij}(t+1) = x_{1j}(t) + F(x_{mj}(t) - x_{nj}(t)) \qquad (4)$$

The F was the scaling factor.

The Crossover operation: Crossover occurs between a mutated individual u_{ij}(t + 1) and its current individual X_i (t) found in the population, and the generated cross test individual C_i (t+1) by (5):

$$C_i(t+1) = \bullet c_{i1}(t+1), c_{i2}(t+1), \ldots\ldots, c_{ik}(t+1) \bullet \qquad (5)$$

The j component of the individual had been expressed as in (6):

$$c_{ij}(t+1) = \begin{cases} u_{ij}(t+1) \; if \;\; rand(0,1) \le CR \; or \; j = rand(i) \\ x_{ij}(t) \qquad\qquad\qquad\qquad\qquad\qquad else \end{cases} \qquad (6)$$

Wherein the $rand_j$andthe crossover probability CR is random number between (0, 1);. The Rand (i) was the random integer falling between 0 and D.

The Selecting operation: the evolutionary individual X_i (t) had been compared to the intermediate cross test individual which was C_i (t+1), and for this, the best individual had been chosen in the subsequent generation of the population which was done by employing the greedy algorithm. The fitness value for the individual f (X_i (t)) = 1/E. Equation (7) is:

$$c_{ij}(t+1) = \begin{cases} X_i(t) \; if \; f(X_i(t)) > f(C_i(t+1)) \\ C_i(t+1) \qquad\qquad\qquad else \end{cases} \qquad (7)$$

The Algorithm termination: all individuals in the population of X (t+1) had been tested. The algorithm was terminated in case the iterations could meet this condition or the time there were similar optimal results in the process that exceeded a certain fixed value. Else, the iterations plus 1 after which the algorithm had to return to step 2 to get the best clustering centre. From the above-mentioned analysis, the DE algorithm that had been applied for optimizing its initial cluster centre for the algorithm of multi-agent clustering, there was a remarkable improvement in the quality of clustering and also the design of the structure was simple. The variation operation along with the cross operation ensured evolution diversity and also strengthened its global search ability. The ability of local search however needed strengthening. The speed of convergence in its later stage also required a lot of improvement. The article had put forth a new type of evolution algorithm and this strengthened its local search ability based on the premise of being able to the ability of global optimization [24]. It was also quite important that in the population scale N, the scale factor F and the crossover probability CR for control parameters found in differential algorithms were employed.

The effect of this was that the F had controlled the scaling degree for the information of individual difference. The CR will directly affect the variation function of individuals for testing its individual structure. So, a choice of suitable, as well as dynamical F and the CR ensured the diversity of population in its early stage to enhance global search ability. The process of evolution and the speed of convergence or its accuracy will have to be accelerated for making this algorithm converge to its optimal solution as early as possible.

3.4 Proposed Multi Agent Firefly Clustering Algorithm

The Firefly Algorithm (FA) is a biologically inspired algorithm in meta-heuristic optimization that had been proposed in the year 2008 by Xin-She Yang. This FA was based on the behavior of communication of the tropical fireflies and their flashing patterns. The FA makes use of three different idealized rules in building a mathematical model in the algorithm [25]:

- All the fireflies are presumed to be unisex;
- The attractive is directly proportional to its brightness and will decrease with an increase in the distance;
- A firefly's brightness will be affected by the objective function. So, a problem of maximization will be the brightness that is proportional to the objective function value.

For a standard FA, there were two different points: construction of the intensity of light and change in attractiveness. First, it can be presumed that the firefly's brightness is computed by means of an encoded objective function of the landscape. Next, it will define its variation in light intensity which decreases with the actual distance from the source where the media absorbs its light and therefore in a simulation, it will support the intensity of light I will vary with distance r and light absorption parameter γ which is exponentially. This will be (8):

$$I = I_0 e^{-\gamma r^2} \quad (8)$$

Where in I_0 denotes its original light intensity which is the source and γ which is the coefficient of light absorption. From these idealized rules [26] it is called simulation and so the attractiveness of the firefly which is proportional to the intensity of light i. Thus, it will be able to define the light of the firefly which is proportional to light intensity i. Thus, it can now define the light attractive coefficient β of the firefly in a way in of light intensity coefficient I. This is (9):

$$\beta = \beta_0 e^{-\gamma r^2} \quad (9)$$

Wherein the β_0 denotes its original light attractiveness which is at r = 0.

There was also a Cartesian distance that was used to calculate any distance existing between any of the two fireflies i and j at positions x_i and $x_j$as in (10):

$$r_{ij} = \| x_i - x_j \|_2 = \sqrt{\sum_{k=1}^{d} (x_{i,k} - x_{j,k})^2} \quad (10)$$

Where the d denotes the actual number of dimensions. The movement to firefly i to another attractive (brighter) firefly j has been shown as in (11):

$$x_i = x_i + \beta_0 e^{-\gamma r^2} (x_j - x_i) + \alpha \varepsilon_i \quad (11)$$

Wherein its x_i denotes the present location of the firefly i and 2nd term its attraction and final term is for the randomization having the vector of the random variables ε_i and the Lévy flight. In the final term, α denotes the scaling parameter controlling a step size linked to the interests of problems. Parameter α (t) [27] has been defined as (12):

$$\alpha(t) = \alpha(0) \left(\frac{10^{-4}}{0.9} \right)^{t/t_{\max}} \quad (12)$$

Where the α (t) denotes the parameter of randomization and iteration number.

For the clustering, the firefly's position vector will be i x_i is (c_1, c_2... c_M), and this is where every firefly duly has its place in the cluster centre. Every firefly's attractiveness will be defined by its objective function. Pseudo code for the multi-agent clustering algorithm has a combined FA as below.

Initialize a population of fireflies $x_i (i = 1, 2, ..., N_f)$
Initialize a multi agent cluster centers $c_m (m = 1, 2, ..., N_c)$
Initialize a objects $o_n (n = 1, 2, ..., N_o)$
while t < MaxGeneration do
Each object is assigned the closest
multi agent cluster center
calculate $J_i (i = 1, 2, ..., N_f)$
Light intensity I_i *is determined by* J_i
if $I_i > I_j$ *then*
Move firefly i towards j
end if
Find the current best
end while
Postprocess results and visualization

IV. 4 RESULTS AND DISCUSSION

In this section, the multi agent K Means clustering, multi agent DE clustering and multi agent firefly clustering methods are used. Experiments are carried out using 5 to 50 number of agents. The precision, recall and f measureas shown in tables 1 to 3 and figures 2 to 4.

Table 1Precision for Multi Agent Firefly Clustering

Number of Agents	Multi Agent K Means Clustering	Multi Agent Differential Evolution Clustering	Multi Agent Firefly Clustering
5	0.5202	0.5237	0.5331
10	0.5647	0.5868	0.566
15	0.6463	0.6864	0.6968
20	0.7073	0.7282	0.7413
25	0.7517	0.7823	0.7912
30	0.7574	0.7982	0.8481
35	0.7918	0.824	0.8851
40	0.8265	0.8308	0.925
45	0.8294	0.8581	0.91
50	0.8745	0.8706	0.9304

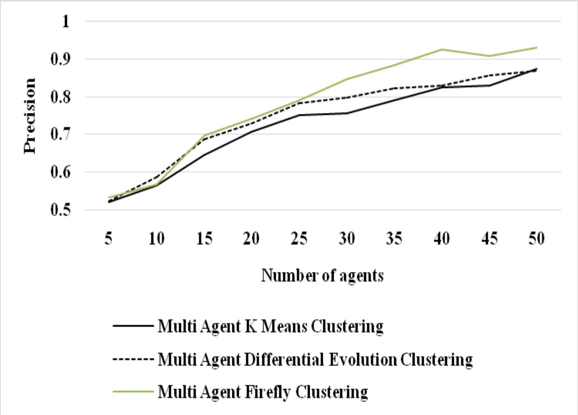

Figure 2Precision for Multi Agent Firefly Clustering

From the figure2, it can be observed that the multi agent firefly clustering has higher precision by 2.44%, 4.69%, 11.29%, 11.24% & 6.19% for multi agent k means clustering and by 1.77%, 1.78%, 6.06%, 10.73% & 6.64% for multi agent DE clustering when compared with 5, 20, 30, 40 and 50 number of agents respectively.

Table 2Recall for Multi Agent Firefly Clustering

Number of Agents	Multi Agent K Means Clustering	Multi Agent Differential Evolution Clustering	Multi Agent Firefly Clustering
5	0.5235	0.5232	0.5314
10	0.5636	0.5825	0.5635
15	0.6428	0.6893	0.6992
20	0.7104	0.7293	0.7407
25	0.7458	0.7755	0.7863
30	0.7607	0.8009	0.8545
35	0.7885	0.8234	0.8838
40	0.8339	0.827	0.9245
45	0.8302	0.8651	0.9117
50	0.8679	0.8701	0.9337

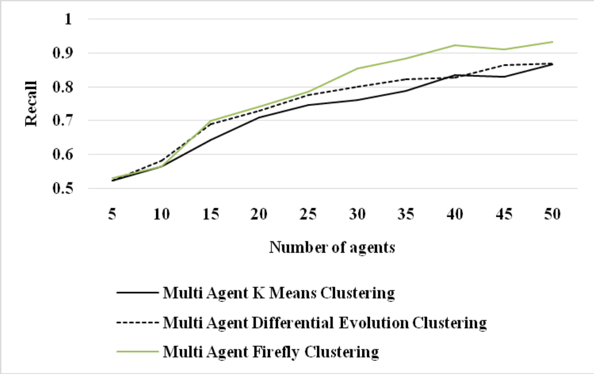

Figure 3Recall for Multi Agent Firefly Clustering

From the figure3, it can be observed that the multi agent firefly clustering has higher recall by 1.49%, 4.17%, 11.61%, 10.3% & 7.3% for multi agent k means clustering and by 1.55%, 1.55%, 6.47%, 11.13% & 7.05% for multi agent DE clustering when compared with 5, 20, 30, 40 and 50 number of agents respectively.

Table 3F Measure for Multi Agent Firefly Clustering

Number of Agents	Multi Agent K Means Clustering	Multi Agent Differential Evolution Clustering	Multi Agent Firefly Clustering
5	0.5218	0.5234	0.5322
10	0.5641	0.5846	0.5647
15	0.6445	0.6878	0.698
20	0.7088	0.7287	0.741
25	0.7487	0.7789	0.7887
30	0.759	0.7995	0.8513
35	0.7901	0.8237	0.8844
40	0.8302	0.8289	0.9247
45	0.8298	0.8616	0.9108
50	0.8712	0.8703	0.932

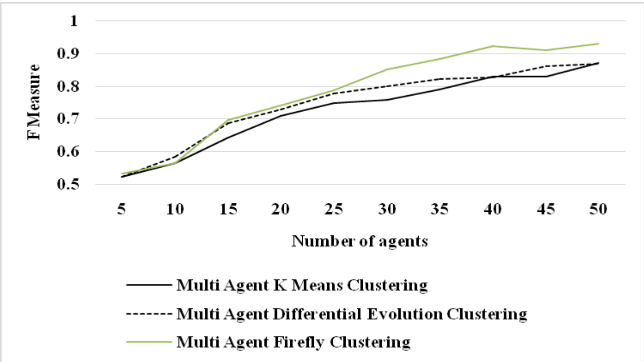

Figure 4F Measure for Multi Agent Firefly Clustering

From the figure4, it can be observed that the multi agent firefly clustering has higher f measure by 1.97%, 4.44%, 11.46%, 10.76% & 6.74% for multi agent k means clustering and by 1.66%, 1.67%, 6.27%, 10.92% & 6.84% for multi agent DE clustering when compared with 5, 20, 30, 40 and 50 number of agents respectively.

V. CONCLUSION

For efficient personalization, the usage and the content attributes for the site have to be integrated within usage of learning resource and its mining framework which is used by its recommendation engine. Clustering is defined as a technique found in data mining for identifying interesting patterns in the dataset. Large dataset which is grouped within clusters which are of smaller sets having similar data with the K-Means algorithm. The initial centroids are needed as the parameters of input while making use of the K-means clustering algorithm. For the purpose of this work, the multi-agent K-means clustering approach optimized with FA was based on the model of client-server for the distributed computing system that has four different intelligent agents that have been deployed on various clients. The DE algorithm is an algorithm of heuristic global optimization which is based on the population and is put forward as a DE algorithm which is combined with the multi-agent clustering algorithm. This improved algorithm had a convergence speed that was faster and a stronger ability of global search. In the multi-agent FA clustering algorithm, there is a parameter which is changed at the time there is no change to the assignment. In this algorithm, all fireflies will move in a random manner. This only means the fireflies can easily escape from its local minimum. With the increasing algorithm iteration, the firefly tries to converge. The results have proved that the multi-agent firefly clustering possesses a precision that was higher by about 2.44%, 4.69%, 11.29%, 11.24% and 6.19% for the multi-agent K-means clustering and further by 1.77%, 1.78%, 6.06%, 10.73% and 6.64% for the multi-agent DE clustering on being compared with the 5, 20, 30, 40 and the 50 number of agents.

REFERENCES

1. Chaimontree, S., Atkinson, K., &Coenen, F. (2011, May). A multi-agent based approach to clustering: harnessing the power of agents. In International Workshop on Agents and Data Mining Interaction (pp. 16-29). Springer, Berlin, Heidelberg.
2. Chaimontree, S., Atkinson, K., &Coenen, F. (2012). A framework for multi-agent based clustering. Autonomous Agents and Multi-Agent Systems, 25(3), 425-446.
3. Wooldridge, M., & Jennings, N. R. (1995). Intelligent agents: Theory and practice. The knowledge engineering review, 10(2), 115-152.
4. Sycara, K. P. (1998). Multiagent systems. AI magazine, 19(2), 79.
5. Nandigam, D., Tirumala, S. S., &Baghaei, N. (2014, December). Personalized learning: Current status and potential. In e-Learning, e-Management and e-Services (IC3e), 2014 IEEE Conference on (pp. 111-116). IEEE.
6. Cai, N., Diao, C., & Khan, M. J. (2015). A Novel Clustering Approach Based on Group Quasi-Consensus of Unstable Dynamic Linear High-Order Multi-Agent Systems. arXiv preprint arXiv:1501.02588.
7. Santos, D. S., De Oliveira, D., &Bazzan, A. L. (2009). A multiagent, multiobjective clustering algorithm. In Data mining and multi-agent integration (pp. 239-249). Springer, Boston, MA.
8. Kubalik, J., Tichy, P., Sindelar, R., &Staron, R. J. (2010). Clustering methods for agent distribution optimization. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 40(1), 78-86.
9. Ozaeta, L., & Graña, M. (2017, June). Finding Communities in Recommendation Systems by Multi-agent Spatial Dynamics. In International Conference on Hybrid Artificial Intelligence Systems (pp. 577-587). Springer, Cham.
10. Melesko, J., & Kurilovas, E. (2018, June). Semantic Technologies in e-Learning: Learning Analytics and Artificial Neural Networks in Personalised Learning Systems. In Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics (p. 34). ACM.
11. Selvi, C., &Sivasankar, E. (2018). A novel Adaptive Genetic Neural Network (AGNN) model for recommender systems using modified k-means clustering approach. Multimedia Tools and Applications, 1-28.
12. Lin, P., Wang, Y., Qi, H., & Hong, Y. (2018). Distributed Consensus-Based K-Means Algorithm in Switching Multi-Agent Networks. Journal of Systems Science and Complexity, 31(5), 1128-1145.
13. Choudhary, N., & Bharadwaj, K. K. (2018). Evolutionary learning approach to multi-agent negotiation for group recommender systems. Multimedia Tools and Applications, 1-23.
14. Jin, J., Zhong, M., Lin, X., &Changhui, T. (2016). A New Cluster Analysis Based on Combinatorial Particle Swarm Optimization Algorithm. International Conference on Education, Management, Computer and Society, 476-479.
15. Inkaya, T., Kayaligil, S., &Özdemirel, N. E. (2015). Ant colony optimization based clustering methodology. Applied Soft Computing, 28, 301-311.
16. Abbasian, M. A., Nezamabadi-pour, H., &Amoozegar, M. (2015). A clustering based archive multi objective gravitational search algorithm. FundamentaInformaticae, 138(4), 387-409.
17. Fister, I. (2017). Multi-agent system based on self-adaptive differential evolution for solving dynamic optimization problems. Proceedings of the 2017 4th Student Computer Science Research Conference, 35-41.
18. Ogston, E., Overeinder, B., Van Steen, M., & Brazier, F. (2003, July). A method for decentralized clustering in large multi-agent systems. In Proceedings of the second international joint conference on Autonomous agents and multiagent systems (pp. 789-796). ACM.
19. Mobascher, B., Dai, H., Luo, T., Sun, Y., & Zhu, J. (2000). Combining Web usage and content mining for more effective personalization. In Proceedings of the International Conference on E-Commerce and Web Technologies.
20. Gomathi, D., &Tajunisha, N. (2012). Improving the accuracy and efficiency of k-means algorithm by using less similarity based clustering technique for better initial centroids. International Journal of Computer Science and Management Research, 1 (4), 698-706.
21. Altayeva, A., &Im Cho, Y. (2016). Multi-agent based on K-means Algorithm for Intelligent Decision Making Model. Korean Institute of Intelligent Systems, 26(1), 87-88.
22. Khan, D. M., &Mohamudally, N. (2010, June). A multiagent system (MAS) for the generation of initial centroids for k-means clustering data mining algorithm based on actual sample datapoints. In Software Engineering and Data Mining (SEDM), 2010 2nd International Conference on (pp. 495-500). IEEE.
23. Wang, X. M., & Wang, J. B. (2013). Research and Improvement on K-Means Clustering Algorithm. In Advanced Materials Research (Vol. 756, pp. 3231-3235). Trans Tech Publications.
24. Hosseini, M., Sadeghzade, M., &Nourmandi-Pour, R. (2014). An efficient approach based on differential evolution algorithm for data clustering. Decision Science Letters, 3(3), 319-324.
25. Yang, X. S., & He, X. (2013). Firefly algorithm: recent advances and applications. arXiv preprint arXiv:1308.3898.
26. Zhang, L., Liu, L., Yang, X. S., & Dai, Y. (2016). A novel hybrid firefly algorithm for global optimization. PloS one, 11(9), e0163230.
27. Hassanzadeh, T., &Meybodi, M. R. (2012, May). A new hybrid approach for data clustering using firefly algorithm and K-means. In Artificial Intelligence and Signal Processing (AISP), 2012 16th CSI International Symposium on (pp. 007-011). IEEE.

AUTHORS PROFILE

Nageswara Rao Gottipati Completed M. Tech from Acharya Nagarjuna University, in the year 2010. Currently working as Associate Professor in the Department of Information Technology at Sasi Institute of Technology & Engineering, Tadepalligudem, Andhra Pradesh. Now he is research scholar, Department of CSE, Hindustan University, Chennai His research interests are Personalized Education, Data Mining, Multi Agents etc., He is Life member of ISTE, AMIE. He is having13 years of teaching experience.

Dr. A. Rama Prasath Completed MCA form Anna University, in the year 2011 and got doctorate from Hindustan University, Chennai, in the year 2016 in the research area of Image Processing. His research interests are Image Processing, Machine Learning etc., and in his credit, he has eight papers published in National/ International repute Conferences /Journals. Currently working as Assistant Professor (Selection Grade) in Department of Computer Applications at Hindustan Institute of Technology & Science, Chennai.