International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-9 Issue-3, February 2020

Hybridized Gradient Descent Spectral Graph and Local-global Louvain Based Clustering of Temporal Relational Data

L. Jaya Singh Dhas, B. Mukunthan, G. Rakesh

Abstract: ***Temporal data clustering examines the time series data to determine the basic structure and other characteristics of the data. Many methodologies simply process the temporal dimension of data but it still faces the many challenges for extracting useful patterns due to complex data types. In order to analyze the complex temporal data, Hybridized Gradient Descent Spectral Graph and Local-Global Louvain Clustering (HGDSG-LGLC) technique are designed. The number of temporal data is gathered from input dataset. Then the HGDSG-LGLC technique performs graph-based clustering to partitions the vertices i.e. data into different clusters depending on similarity matrix spectrum. The distance similarity is measured between the data and cluster mean. The Gradient Descent function find minimum distance between data and cluster mean. Followed by, the Local-Global Louvain method performs the merging and filtering of temporal data to connect the local and global edges of the graph with similar data. Then for each data, the change in modularity is calculated for filtering the unwanted data from its own cluster and merging it into the neighboring cluster. As a result, optimal 'k' numbers of clusters are obtained with higher accuracy with minimum error rate. Experimental analysis is performed with various parameters like clustering accuracy (C_{ACC}), error rate (Err_{Rate}), computation time ($Time_C$) and space complexity (S_{com}) with respect to number of temporal data. The proposed HGDSG-LGLC technique achieves higher C_{ACC} and lesser $Time_C$, minimum Err_{Rate} as well as S_{com} than conventional methods.***

Keywords: Temporal data analysis, Gradient Descent Spectral graph clustering, Local-Global Louvain method, change in modularity

I. INTRODUCTION

The cluster contains complete knowledge about temporal data and understands dataset structure. The processing of large volume of data is more complicated. Therefore, clustering is a process of dividing large dataset into clusters to minimize complexity of processing data.

A novel approach called Evolutionary Clustering based on Graph regularized Nonnegative Matrix Factorization (ECGNMF) was introduced in [1] to find the dynamic communities. The designed clustering technique failed to use the similarity index for achieving higher accuracy. In [2], bi-weighted ensemble approach divided temporal data into different clusters with the help of a hidden Markov model. But, designed approach has more computation time and space complexity.

A general transformation approach was introduced in [3] for grouping the data with similarity. The designed approach failed to accurately discover the patterns in the dataset. Hidden Markov models (HMMs) was presented in [4] to analyze the sequential time series data. But the complex time series data were not analyzed with lesser complexity. A copula-based distance measure was introduced in [5] for clustering data depending on dependence structure. But the clustering accuracy was not improved.

A k-Medoids algorithm was designed in [6] to cluster data using affinity search. The designed algorithm was not improved the clustering accuracy while processing the complex data. A finite mixture model was designed in [7] to perform a dynamic clustering of Spatio-temporal data. The clustering accuracy was not improved by designed model. A novel method called multivariate reconstructed phase space (MRPS) was developed in [8] for finding the multivariate temporal patterns. A hierarchical aligned cluster analysis (HACA) was developed in [9] to divide multi-dimensional time series into segments. However, the computational complexity was not reduced. Consistent algorithms were designed in [10] for clustering the time series with a minimum error rate. But the algorithm was not effectively minimizing the time complexity in the data prediction.

1.1 Proposal Contributions

The problems identified from the existing survey are addressed by novel HGDSG-LGLC technique. The overall contribution of the proposed work is:

- To improve the temporal data clustering accuracy, the HGDSG-LGLC technique is introduced with the hybrid technique. The hybrid of spectral clustering and Louvain method is applied to form similar data into cluster with minimum time. The similar data are identified by determining distance between data and cluster mean (i.e. centroid). The gradient descent function is used for finding similar data with lesser distance from centroid.

Revised Manuscript Received on February 05, 2020.
* Correspondence Author
L. Jaya Singh Dhas*, Research Scholar, Department of Computer Science, Jairams Arts and Science College, (Affiliated to Bharathidasan University, Tiruchirapalli) Karur - 639003, Tamilnadu, India.
Orchid –id: https://orcid.org/0000-0002-0136-3941 Email: jayasinghdhas@yahoo.com
B. Mukunthan, Research Supervisor & Assistant Professor, Department of Computer Science, Jairams Arts and Science College, (Affiliated to Bharathidasan University, Tiruchirapalli) Karur - 639003, Tamilnadu, India. Orchid –id: https://orcid.org/0000-0001-8452-3164 Email: dr.mukunthan.bmk@gmail.com
G. Rakesh, Dean of Science, Department of Computer Science, Jairams Arts and Science College, (Affiliated to Bharathidasan University, Tiruchirapalli) Karur - 639003, Tamil Nadu, India.
Orchid –id: https://orcid.org/0000-0001-6978-418X Email: rakesh.gnana@gmail.com

- The Local and global Louvain method is applied to validate the cluster and the data. If data is not a member of particular cluster, then it filtered from own cluster and merging into the nearest cluster centroid. This process helps to correctly group the similar temporal data into the cluster with lesser S_{com} and Err_{Rate}.

1.2 Paper Outline

The manuscript is organized as: Section 2 elaborates the issues and challenges during the bitemporal data clustering in the literature survey. Section 3 describes the HGDSG-LGLC technique for bitemporal data analysis. In section 4, experimental settings are described with time series dataset and the performance results of various parameters are discussed in section 5. Finally, section 6 concludes the work.

II. LITERATURE SURVEY

A variational Expectation-Maximization algorithm was designed in [11] for partitioning the data into different clusters. The designed algorithm failed to improve clustering accuracy. Dynamic random graph clustering of temporal data was presented in [12] using Variational expectation–maximization. The clustering algorithm failed to perform cluster analysis in efficient way. Generalize d k -means-based clustering algorithm was presented in [13] for data clustering with lesser time consumption. But the algorithm was not minimized the error rate. A hybrid clustering algorithm was presented in [14] with data similarity value. Though the algorithm improves the accuracy, clustering time was not reduced.

A sparse subspace clustering (SSC) algorithm was presented in [15] to partition data with minimum error rate. The developed algorithm was not minimized the space complexity. Continuous Clustering of Trajectory Stream Data was presented in [16] for achieving the higher clustering quality. But the clustering technique requires more storage space.

The sliding window method combined with dimension reduction techniques was presented in [17] for time-series analysis and pattern identification. But the technique failed to achieve higher clustering accuracy. Hidden Markov model ensembles were presented in [18] for dividing the temporal data. The designed model improves the accuracy but the performance of the error rate was not reduced.

A dynamic fuzzy cluster (DFC) was introduced in [19] for organizing the data. The performance of clustering accuracy remained unsolved. The two incremental fuzzy clustering techniques were developed in [20] for partitioning the time series data with higher accuracy. The clustering time was not minimized.

The problems identified from above mentioned works are addressed by implementing a new clustering technique termed HGDSG-LGLC. The brief description of HGDSG-LGLC method is given in next section.

III. HYBRIDIZED GRADIENT DESCENT SPECTRAL GRAPH AND LOCAL-GLOBAL LOUVAIN CLUSTERING FOR BITEMPORAL DATA CLUSTERING

Due to large volume of gathered temporal data and increase in complexities, data processing has become a more significant procedure. In order to minimize the complexity involved in the processing and analyzing such a volume of data, a novel technique called HGDSG-LGLC is introduced. The HGDSG-LGLC technique is the graph-based clustering technique. The clustering is the process of grouping the temporal data into different categories or clusters. The detection of clusters in large graphs is becoming a major problem. Therefore, the hybridization technique improves the clustering performance with minimum time.

3.1 Network model

Let us take, temporal dataset D_t and collect the number of time series data $d_1, d_2, d_3, \ldots.. d_n$. After collecting, the hybrid method is applied for grouping the similar data for further processing. The proposed HGDSG-LGLC technique constructs the undirected graph model $G = \langle v, e \rangle$ where 'v' represents a number of nodes (i.e. temporal data $d_1, d_2, d_3, \ldots.. d_n$), '$e$' represents the edges i.e. link between the data. The undirected simple graph is shown in figure 1.

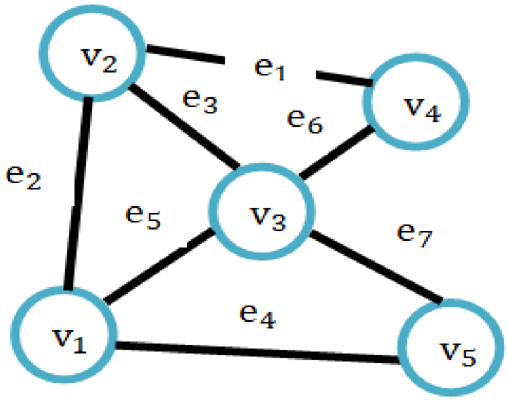

Figure 1 Undirected graph

Figure 1 illustrates the undirected graph with vertices v_1, v_2, v_3, v_4, v_5 and edges are symbolized as $e_1, e_2, e_3, e_4, e_5, e_6, e_7$. The temporal clustering problem consists of finding the partitions of the vertex in the form of $C = \{c_1, c_2, c_3, \ldots. c_k\}$ such that $1 \le j \le k$ exhibits the cluster structure. Then the gradient descent spectral clustering algorithm is applied to a Louvain method for partitioning the vertices into different groups.

3.2 Graph-based Bitemporal data clustering

The Graph-based Bitemporal data clustering technique called HGDSG-LGLC is used for analyzing the complex bitemporal data. The HGDSG-LGLC technique uses the hybridization of the spectral clustering with the Louvain method.

International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-9 Issue-3, February 2020

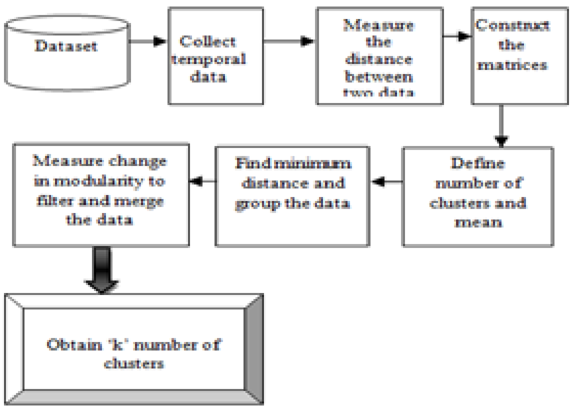

Figure 2 Flow process of HGDSG-LGLC technique

As shown in figure 2, the flow process of HGDSG-LGLC technique to analyze the bitemporal data through the clustering process. Let us dataset D_t and collect the number of time series data $d_1, d_2, d_3, \ldots. d_n$. The gradient descent spectral clustering uses the spectrum (eigenvalues) of the similarity matrix of temporal data for grouping process. Therefore, it initially computes the similarity between the vertices (i.e. data). Here, the distance function is used for determining similarity between temporal data which is expressed as follows,

$$t_{ij} = \sqrt{(d_i - d_j)^2} \qquad (1)$$

Where, t represents the distance, d_i and d_j denotes the two temporal data. Then the similarity matrix is constructed with the distance similarity values,

$$s_{ij} = t_{ij} \qquad (2)$$

In (2), s_{ij}denotes a similarity matrix between the two data. Then the unnormalized Laplacian matrix is constructed with the diagonal matrix and similarity matrix as follows,

$$L_{ij} = d_{ij} - s_{ij} \qquad (3)$$

Where, L_{ij} denotes an unnormalized Laplacian matrix, d_{ij} represents the diagonal matrix, s_{ij} is the similarity matrix. The diagonal matrix with the degrees $g_1, g_2, g_3, \ldots. g_n$ on the diagonal which is expressed as follows,

$$d_{ij} = \begin{bmatrix} g_1 & 0 & 0 & 0 \\ 0 & g_2 & 0 & 0 \\ 0 & 0 & \ldots. & 0 \\ 0 & 0 & 0 & g_n \end{bmatrix} \qquad (4)$$

Where, d_{ij} represent diagonal matrix. The first 'K' largest eigen values of Laplacian matrix and corresponding eigenvectors, $E_1, E_2, E_3, \ldots E_m$ are determined and form matrix by Eigen vectors in column. The matrix is given as follows,

$$W_{ij} = [E_1, E_2, E_3, \ldots E_m] \qquad (5)$$

From (6), W_{ij} symbolizes the matrix. Then, normalized Laplacian matrix is constructed through renormalizing every row of 'W_{ij}'matrix. The new matrix is constructed whose columns represent the eigenvectors corresponding to the 'm' smallest eigen values.

$$L_{ij}{}^{n} = \frac{w_{ij}}{\sum_{j=1}^{n} w_{ij}} \qquad (6)$$

Where, rows of the matrix 'W_{ij}' as collection of 'n' temporal data and group them into 'k' number of clusters by k-means clustering algorithm. Initially, the mean is assigned to each cluster. Then, the proposed technique uses gradient descent function to find minimum distance between data and cluster mean value.

$$f = arg\ min \sum_{i=1}^{n} \sum_{j=1}^{k} \left\| d_i - m_j \right\|^2 \qquad (7)$$

Where, f denotes a gradient descent function, m_j denotes a mean of the clusters, d_i is the temporal data. By applying k means, the mean (i.e. centroid) value for each cluster is defined. The $arg\ min$ abbreviated as argument of the minimum at which the function values are minimized. The temporal data d_i close to the mean is assigned to particular cluster 'j' if row of matrix W_{ij} is allocated to cluster j.

Then the Louvain method uses two functions namely merging and filtering of bitemporal data in order to obtain the local and global information. The local information represents the knowledge of the neighbors of each vertex to perform graph clustering. The global information represents the knowledge of the whole graph structure. This function helps to retrieve the clusters of highly interconnected data.

In order to connect the local edges of the graph, the temporal data which is closer to their mean is grouped into the cluster. If the temporal data is not closer to their cluster mean value, the data is removed from the current cluster and verifies the distance similarity to the other clusters to insert the data. This helps to connect the vertex into the global edges of the graph. The merging and the filtering process of the Louvain method is expressed as follows,

$$q_m = \left[\frac{\sum w_c + 2e_c}{2b} - \left(\frac{\sum w_t + \alpha_i}{2b}\right)^2\right] - \left[\frac{\sum w_c}{2b} - \left(\frac{\sum w_t}{2b}\right)^2 - \left(\frac{\alpha_i}{2b}\right)^2\right] \qquad (8)$$

Where, q_m denotes a change in graph modularity, $\sum w_c$ denotes sum of all link weights between nodes inside cluster inserted, $\sum w_t$ denotes sum of all link weights to node in cluster moving into, α_i represents the weighted degree of the node 'i', e_c denotes a sum of the link weights between 'i' and other nodes in the cluster that 'i' is moving into, 'b' is the sum of the weights of all links in the network. The modularity is used to measure the structure of graphs and also used to measure the strength of partition of a graph into clusters. The change in graph modularity is used for identifying the links within a cluster and the links between clusters. Hence the proposed technique is called as local and global Louvain method.

This clustering process is repeated until all the data are clustered into a cluster. The merging and filtering process of the Louvain method connects the local and global edges of the graph to improve the clustering performance and minimize the error rate.

The algorithmic process of HGDSG-LGLC is described as,

Input: Bitemporal dataset D_t and number of time series data $d_1, d_2, d_3, \ldots, d_n$
Output: Improve clustering accuracy and minimize the time
Begin
for each data $d_i \in D_t$
Measure the distance t_{ij}
Construct similarity matrix s_{ij}
Construct unnormalized Laplacian matrix L_{ij}
Construct diagonal matrix d_{ij}
Find first 'k' eigenvectors
Construct normalized Laplacian matrix $L_{ij}{}^{n}$
Define 'k' clusters and mean value m_j
for each d_i
for each cluster mean m_j
Find distance $\sum_{i=1}^{n}\sum_{j=1}^{k}\|d_i - m_j\|^2$
If $(arg\,min \sum_{i=1}^{n}\sum_{j=1}^{k}\|d_i - m_j\|^2)$ **then**
Assign data 'i' into the patellar cluster 'j'
else
Calculate the change in modularity q_m
Filter the node 'i' from its own cluster
Merging the node-link 'i' into the neighbor cluster
end if
End for
End for
End for
End

Algorithm 1 Hybridized gradient descent spectral graph and local-global Louvain based bitemporal data clustering

Algorithm1 describes the hybridization of spectral clustering and the Local-Global Louvain method for clustering the bitemporal data. At first, similarity matrix is formed with the help of distance between two vertices (i.e. data) in the graph. Based on the similarity matrix, the unnormalized Laplacian matrix is formed. Then, the diagonal matrix is constructed and subsequently, first 'k' Eigenvectors is computed and constructs the new matrix. With identified Eigenvectors, the normalized Laplacian matrix is formed. Finally, 'k' means algorithm is used and defined the 'k' clusters with mean value. For each data in the cluster, the distance is calculated and identifies the minimum distance between them. If the distance is minimal, then the data is assigned to the particular cluster. Otherwise, the change in modularity is calculated for filtering the data from the current cluster and merging the data into the neighboring clusters which have minimum distance. This process is repeated until all data are combined into the cluster. Finally, the optimal 'k' number of clusters is obtained with higher accuracy.

IV. EXPERIMENTAL SCENARIO

In this section, experimental evaluation of HGDSG-LGLC and existing techniques ECGNMF [1] and Bi-weighted ensemble approach [2] is carried out using Java language with Activity Recognition from Single Chest-Mounted Accelerometer Dataset is taken from UCI machine learning repository [21]. This dataset includes the temporal data points from wearable accelerometer fixed on chest. There are 15 participates performing seven different activities are found through clustering process. The attributes characteristics are real. The association tasks are clustering and classification.

Performance analysis of the proposed technique is compared with existing results with certain parameters such as C_{ACC}, $Time_C$, Err_{Rate} and S_{com}. For the experimental consideration, ten different runs are performed with temporal data ranging from 1000 to 10000.

V. RESULTS ANALYSIS

The experimental results of parameters using HGDSG-LGLC and existing techniques [1] and [2] are discussed. The various parameters are C_{ACC}, $Time_C$, Err_{Rate} and S_{com}. The results are discussed using either table or graphical representation and the statistical calculation is provided for each subsection to show the results of proposed technique against conventional clustering techniques.

5.1 Impact on clustering accuracy

C_{ACC} is defined as the ration of number of temporal data that are correctly grouped to form clusters to the total number of data. The clustering accuracy is measured using the following equation,

$$C_{ACC} = \left[\frac{Number\ of\ data\ correclty\ grouped}{Total\ number\ of\ data}\right] * 100 \qquad (9)$$

Where, C_{ACC} denotes clustering accuracy measured in terms of percentage (%).

Table 1 Clustering Accuracy

Number of data	Clustering Accuracy (%)		
	HGDSG-LGLC	ECGNMF	Bi-weighted ensemble approach
1000	91	85	78
2000	90	82	76
3000	92	84	77
4000	91	86	83
5000	93	88	86
6000	92	87	82
7000	95	89	87
8000	96	87	86
9000	95	88	84
10000	94	89	87

Table 1 describes the performance results of temporal data clustering accuracy with three techniques namely HGDSG-LGLC, [1] and [2]. In order to calculate the clustering accuracy, data is taken in range from 1000 to 10000 for conducting the experimental scenario. The reported clustering accuracy of HGDSG-LGLC technique is higher than the existing techniques. The clustering accuracy is improved by applying the hybrid technique. The spectral clustering constructs the similarity matrix to group the similar temporal data by applying the k means algorithm. The local and global Louvain method calculates the modularity for filtering the unwanted data within the cluster and merging into the neighboring cluster. This process of HGDSG-LGLC technique achieves higher clustering accuracy.

The improvement of the accuracy using HGDSG-LGLC technique is proved by the mathematical calculation. Let us consider the 1000 data for calculating the clustering accuracy. The HGDSG-LGLC technique correctly clustered 910 data and their accuracy is 91% whereas [1] and [2] grouped 850 and 780 data and their accuracy percentages are 85% and 78% respectively. Finally, the ten different results of clustering accuracy of HGDSG-LGLC technique are compared to the conventional methods. The average of ten results confirms that the clustering accuracy is found to be higher using HGDSG-LGLC technique. The comparison results prove that the accuracy is improved by 7% and 13% than [1] and [2]

5.2 Impact on error rate

Err_{Rate} is defined as the ratio of number of data incorrectly grouped to the total number of data. It is measured in terms of percentage (%) and formulated as,

$$Err_{Rate} = \left[\frac{\text{Number of data incorrectly grouped}}{\text{Total number of data}}\right] * 100 \qquad (10)$$

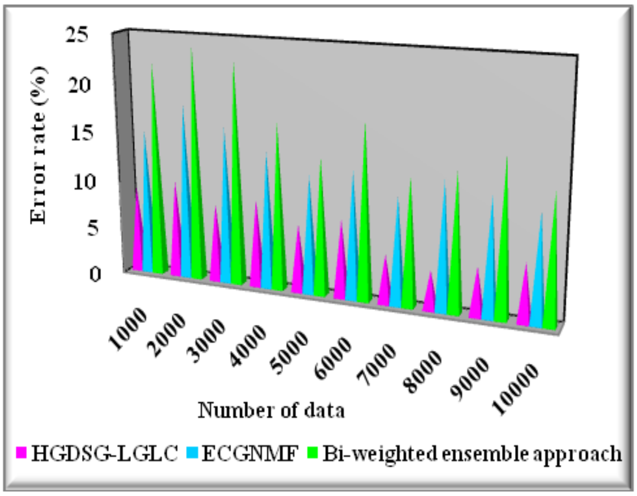

Figure 3 Performance results of error rate

Figure 3 depicts the performance results of error rate for different number of data. The temporal data are gathered from dataset. As shown in graph, the number of data is given in the 'x' axis and output of the error rate is obtained at the 'y' axis. The results of three clustering techniques are represented in three different colors of the cone. The rose color cone indicates error rate of proposed HGDSG-LGLC technique whereas the blue color and green color cone indicate the error rate of two existing methods [1] and [2] respectively. The above figure clearly shows the error rate is found to be minimized using HGDSG-LGLC technique. This is due to the application of the local and global Louvain method. The method initially finds the local relationship i.e. data and cluster mean. When distance between mean and data is lesser, the data is a member of that specific cluster. If the data is not a member of specific cluster, it removed from the own cluster and inserted into the neighboring cluster. The inserting and removing are performed through the modularity. This helps to minimize the error rate of temporal data clustering. The ten various results of an error rate of HGDSG-LGLC technique is found to be minimized by 48% and 59% as compared to [1] and [2].

5.3 Impact on computation time

$Time_c$ is defined as the amount of time taken to group the similar temporal data to form the cluster. The formula for measuring the computation time is given below,

$$Time_c = Number\ of\ data * T\ (clustering\ one\ d) \qquad (11)$$

From (11), T represent time for clustering one temporal data 'd'. It is measured in milliseconds (ms).

Table 2 Computation Time

Number of data	Computation Time (ms)		
	HGDSG-LGLC	ECGNMF	Bi-weighted ensemble approach
1000	20	24	28
2000	22	26	32
3000	27	33	39
4000	32	36	44
5000	38	40	45
6000	40	42	48
7000	43	46	50
8000	46	48	54
9000	50	52	59
10000	53	56	62

Table 2 shows the computation time with respect to a varying number of temporal data ranging from 1000 to 10000 using three different methods. By increasing the number of data packets, the computation time differs due to increasing the number of temporal data for each iteration. From the table, it is inferred that the computation time is found to be minimized using HGDSG-LGLC technique when compared to the existing technique. HGDSG-LGLC technique uses gradient descent function to identify minimum distance between the data and cluster mean. This helps to effectively identify the cluster members with lesser time.

This is evident from the mathematical calculation when the number of temporal data was found to be 1000, the computation time of HGDSG-LGLC technique is $20m$. The computation time was '$24ms$' using [1] and '$28ms$' using [2] respectively. The obtained results clearly evident the computation time using HGDSG-LGLC technique is minimized by 9% compared to [1] and 24% when compared to [2].

5.4 Impact on space complexity

S_{com} is defined as the amount of memory space consumed to store clustered temporal data and it is computed by,

$$S_{com} = n * Space\ (storing\ one\ d) \qquad (12)$$

Where S_{com} is the space complexity, n denotes number of data,' 'd' denotes a temporal data. The unit to measure the space complexity is mega bytes (MB).

Retrieval Number: C5989029320/2020©BEIESP
DOI: 10.35940/ijeat.C5989.029320
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

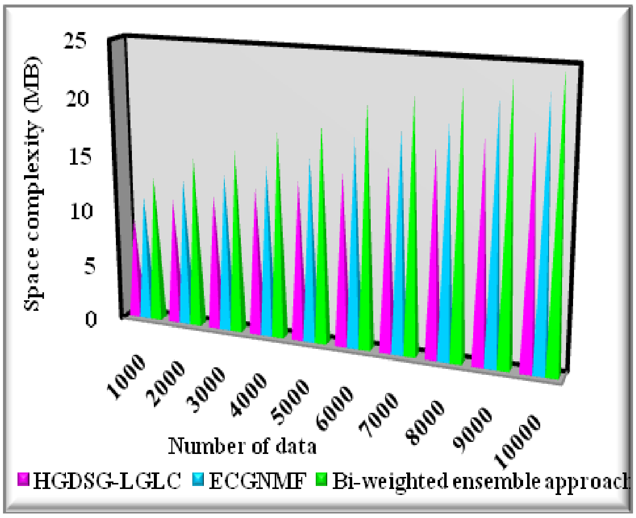

Figure 4 Performance results of space complexity

Figure 4 given above shows the effect of space complexity for different temporal data ranging from 1000 to 10000 using three different methods. The horizontal axis represents the number of temporal data and the vertical axis represents space complexity. By increasing the number of temporal data, the space complexity is found to be reduced. However, from the figure, it is inferred that the space complexity is minimized using HGDSG-LGLC technique. This is achieved by measuring the change in modularity. This helps to remove the unwanted temporal data within the cluster and insert to the other neighboring cluster. From the results, only similar data are clustered inside cluster. In this manner, the space complexity is found to be minimized using HGDSG-LGLC. The average of comparison results evidently proves that space complexity of HGDSG-LGLC technique is reduced by 14% when compared to [1] and 25% when compared to [2].

The above-discussed result of different parameters clearly proves that the proposed HGDSG-LGLC technique efficiently performs the temporal data analysis through the clustering with minimum time.

VI. CONCLUSION

A novel technique called hybrid clustering with the graphical structure for bitemporal data analysis is performed aiming to improve the accuracy and minimize time. The gradient descent spectral clustering is used to measure distance similarity between data and cluster mean for data grouping. With help of distance measure, similar data is combined into form the cluster. The change in modularity is calculated to all the clusters for reducing irrelevant data inside the cluster. The irrelevant data inside the cluster is filtered and inserting into the other neighboring clusters. This, in turn, improves the temporal data clustering accuracy and minimizes the error rate. Experimental analysis of proposed HGDSG-LGLC technique and conventional methods are carried out using human activity recognition dataset. The performance results are evaluated with conventional methods. The compared result shows that clustering accuracy is enhanced with minimum time, space complexity as well as error rate than existing methods.

REFERENCES

1. Wei Yu, Wenjun Wanga, Pengfei Jiao, Xuewei Li," Evolutionary clustering via graph regularized nonnegative matrix factorization for exploring temporal networks", Knowledge-Based Systems, Elsevier, Volume 167, 2019, Pages 1-10.
2. YunYang and Jianmin Jiang, "Bi-weighted ensemble via HMM-based approaches for temporal data clustering", Pattern Recognition, Elsevier, Volume 76, April 2018, Pages 391-403.
3. Leon Bornemann, Tobias Bleifuß, Dmitri Kalashnikov, Felix Naumann, Divesh Srivastava, "Data Change Exploration Using Time Series Clustering", Datenbank-Spektrum, Springer, Volume 18, Issue 2, 2018, Pages 79–87.
4. Nazanin Asadi , Abdolreza Mirzaei , Ehsan Haghshenas, "Creating Discriminative Models for Time Series Classification and Clustering by HMM Ensembles", IEEE Transactions on Cybernetics, Volume 46 , Issue 12 , 2016, Pages 2899 – 2910.
5. Beibei Zhang and Baiguo An, "Clustering time series based on dependence structure", PLoS ONE, Volume 13, Issue 11, 2018, Pages 1-22.
6. Saeed Aghabozorgi, Teh Ying Wah, Tutut Herawan, Hamid A. Jalab, Mohammad Amin Shaygan, and Alireza Jalali, "A Hybrid Algorithm for Clustering of Time Series Data Based on Affinity Search Technique", The Scientific World Journal, Hindawi Publishing Corporation, Volume 2014, March 2014, Pages 1-12.
7. Lucia Paci and Francesco Finazzi, "Dynamic model-based clustering for spatio-temporal data", Statistics and Computing, Springer, Volume 28, Issue 2, 2018, Pages 359–374.
8. Wenjing Zhang and Xin Feng, "Event Characterization and Prediction Based on Temporal Patterns in Dynamic Data System", IEEE Transactions on Knowledge and Data Engineering, Volume 26, Issue 1, 2014, Pages 144 – 156.
9. Feng Zhou , Fernando De la Torre , Jessica K. Hodgins, "Hierarchical Aligned Cluster Analysis for Temporal Clustering of Human Motion", IEEE Transactions on Pattern Analysis and Machine Intelligence , Volume 35 , Issue 3 , 2013, Pages 582 – 596.
10. 10. Azadeh Khaleghi, Daniil Ryabko, Jérémie Mary, Philippe Preux, "Consistent Algorithms for Clustering Time Series", Journal of Machine Learning Research, Volume 17, Issue 3, 2016, Pages 1−32.
11. Hani El Assaad, Allou Saméa, Gérard Govaert, Patrice Aknina, "A variational Expectation–Maximization algorithm for temporal data clustering", Computational Statistics & Data Analysis, Elsevier, Volume 103, 2016, Pages 206-228.
12. Catherine Matias and Vincent Miele, "Statistical clustering of temporal networks through a dynamic stochastic block model", Journal of the Royal Statistical Society Series B, Volume 79, Issue 4, 2017, Pages 1119-1141.
13. Saeid Soheily-Khah, Ahlame Douzal-Chouakria, Eric Gaussier, "Generalized k-means based clustering for temporal data under weighted and kernel time warp", Pattern Recognition Letters, Elsevier, Volume 75, 2016, Pages 63-69.
14. Saeed Aghabozorgi, Teh Ying Wah, Tutut Herawan, Hamid A. Jalab, Mohammad Amin Shaygan, and Alireza Jalali, "A Hybrid Algorithm for Clustering of Time Series Data Based on Affinity Search Technique", The Scientific World Journal, Hindawi Publishing Corporation, Volume 2014, March 2014, Pages 1-12.
15. Yan Wang, Yunian Ru, Jianping Chai, "Time series clustering based on sparse subspace clustering algorithm and its application to daily box-office data analysis", Neural Computing and Applications, Springer, 2018, Pages 1–10.
16. Musaab Riyadh, Norwati Mustapha, Md. Nasir Sulaiman, and Nurfadhlina Binti Mohd Sharef, "CC_TRS: Continuous Clustering of Trajectory Stream Data Based on Micro Cluster Life", Mathematical Problems in Engineering, Hindawi Volume 2017, July 2017, Pages 1-9.
17. Mohammed Ali, Mark W. Jones, Xianghua Xie, Mark Williams, "TimeCluster: dimension reduction applied to temporal data for visual analytics", The Visual Computer, Springer, Volume 35, Issue 6–8, 2019, Pages 1013–1026.
18. Nazanin Asadi, Abdolreza Mirzaei, and Ehsan Haghshenas, "Creating Discriminative Models for Time Series Classification and Clustering by HMM Ensembles", IEEE Transactions on Cybernetics, Volume 46, Issue 12, 2016, Pages 2899 – 2910.
19. Min Ji, Fuding Xie, and Yu Ping, "A Dynamic Fuzzy Cluster Algorithm for Time Series", Abstract and Applied Analysis, Hindawi Publishing Corporation, Volume 2013, March 2013, Pages 1-7.

20. Yongli Liu, Jingli Chen, Shuai Wu, Zhizhong Liu, Hao Chao, "Incremental fuzzy C medoids clustering of time series data using dynamic time warping distance", PLoS ONE, Volume 13, Issue 5, 2018, Pages 1-25.
21. Dataset: https://archive.ics.uci.edu/ml/datasets/Activity+Recognition+from+Single+Chest-Mounted+ Accelerometer.

AUTHORS PROFILE

L. Jaya Singh Dhas received his Bachelor of Science in Computer Science from Madurai Kamaraj University - Madurai, India in 1991 and Master of Computer Applications from Bharathidasan University - Tiruchirapalli in the year 1996 and M.Phil from Alagappa University - Karikudi in the year1998. He is currently working as a Assistant Professor, Department of Computer Science, Scott Christian College (Autonomous), Nagercoil since 1998 and his research work focuses on Algorithms, Big Data Analytics, Data Mining..

B. Mukunthan pursed Bachelor of Science in Computer Science from Bharathiar University, India in 2004 and Master of Computer Applications from Bharathiar University in year 2007 and Ph.D from Anna University - Chennai in 2013. He is currently working as Research Advisor in Department of Computer Science, Jairams Arts & Science College, Affiliated to Bharathidasan University, Tiruchirapalli since 2016. He is a member of IEEE & IEEE computer society since 2009, a life member of the MISTE since 2010. He has published more than 25 research papers in reputed international journals. He is also Microsoft Certified Solution Developer. His main research work focuses on Algorithms, Bioinformatics, Big Data Analytics, Data Mining, IOT and Neural Networks. He also invented a Novel and Efficient online Bioinformatics Tool and filed for patent. He has 12 years of teaching experience and 10 years of Research Experience.

G. Rakesh pursed Bachelor of Computer Application from Bharathidasan University, India in 2004 and Master of Computer Applications from Bharathiar University in year 2007 and Ph.D from Bharathiar University - Coimbatore on 2019. He is currently working as Dean of science in Jairams Arts & Science College, Karur. He is a member of IEEE & IEEE computer society since 2009, a life member of the MISTE since 2010. He has published more than 10 research papers in reputed international journals including Thomson Reuters (SCI & Web of Science). His main research work focuses on Algorithms, Bioinformatics, Big Data Analytics, Data Mining, IOT and Neural Networks. He has 12 years of teaching experience and 10 years of Research Experience.