**Project Progress Report**

**Graph Coloring algorithm - Data Clustering**

Algorithms CSCE 5150

Yuede Ji

University of North Texas

A picture containing text, clipart

Description automatically generatedLogo, company name

Description automatically generatedA picture containing clipart

Description automatically generated

BY

KOMAL PAWAR

**Table of Contents:**

* Abstract
* Algorithm Implementation
* **Experimentation:**
* Result
* Conclusion
* Difficulties faced

**Abstract:**

The main aim of this project is to assign different colors to the different components that are present in the graph while having some constrains taken into consideration. There are three coloring forms that are most popular which are edge, vertex, face coloring .I had noticed the main problem with the graph coloring is to know number of colors required to make sure no two adjoining edges or vertices having same hue.

A graph's colorful number G is the smallest number of colors necessary to color it.. A well-known NP Complete Issue is the Graph Coloring problem. NP complete issues are those with an uncertain state.

A correct coloring solution cannot be obtained using a polynomial time technique. Nobody has yet proven that no polynomial-time algorithm exists for any of these. For various limitations, different graph coloring difficulties develop. To some extent, the answer to graph coloring is heavily reliant on graph structure.

**Algorithm Implementation:**

The primary purpose of this project is to assign colors to specified graph regions while keeping certain constraints in mind. Out of the three types of coloring, edge coloring, face coloring, and vertex coloring are the most common graph coloring issues. Graph coloring assures that no two adjacent vertex colors are the same.

Edge coloring may be divided into two types: vertex coloring and list coloring, map coloring, path coloring, and total coloring.

In graph theory, graph coloring is an important notion. It is utilized in a wide range of real-time computer science applications, including  image capturing, clustering, data mining,, and image segmentation.

* The four-color theorem states that any map must be colored with four different colors. A country or state with geographical map colors is one in which no two contiguous cities can be allocated the same colour. A vertex represents each zone, while an edge represents each of its adjoining zones.
* Keeping data in a specific sequence, like in television channels or voice signals in speech processing. Stacks and Queues implementation Adjacency Matrix Graph Hash tables, heaps, and representations are all supported. Color may be applied to components of transitively orientable (coTR0) networks in 0(n4) time, where n is the number of vertices.
* Graphs have the same color problem. Register Allocation is the process of assigning a high number of target program variables to a restricted number of CPU registers in compiler optimization.
* Despite being a well-known NP-complete task, vertex coloring for some types of graphs can be solved in polynomial time [lo].
* They simplify the graph coloring problem to the register allocation problem by showing that a program can be built for any given graph in which the program's register allocation (with registers representing nodes and machine registers representing available colors) colors the original graph.

**Experimentation:**

I acquired datasets.  To find a cluster partition with effective internal cluster coherence and cluster separation, use dissimilarities between each pair of objects. Assume that the data is organized as an unsupervised advantage graph G(V, E), where V is the vertex set and E is equal to V. The edge set is V. Edge-weights, where G vertices stand in for data items and edges for neighborhood associations, explain how two linked vertices differ from one another.

Hence, the purpose of this technique is to use variance between each pair of elements to find a clustering subdivision with efficient governance cluster coherence and clustering separation. It facilitates the development of a data set division with an undetermined number of nodes of clusters. These two problems suggest the nodes between a pair of vertices in one cluster should be small weighted, while those between two clusters should be huge weighted. A cluster must fulfill two basic conditions it must have high internal homogeneity, and have large heterogeneity across objects from clusters.

Consequently, our clustering technique requires the construction of a superior threshold graph, which is a subgraph of the original G. (V, E). Let G (V, E) be the superior threshold graph associated with a D threshold value in a dissimilarity table. G is represented by the vertex set V and the edge set (vi, vj)|D(vi, vj) = di,

An innovative technique to b-coloring that is ideally suited to the cluster analysis problem. The key idea of this research is to apply this method to the superior threshold graph G. After all, the goal is to assign colors to the network G vertices so that no two surrounding pair vertices with dissimilarities bigger than the threshold have the same hue and at least one dominant vertex has the same hue is next to at least one dominant vertex for each color class Every other color's vertices sets.

**Result:**

Text

Description automatically generated

**Difficulties faced:**

It was challenging to gather datasets and give colors to graph nodes. I've begun working on the threshold graph. While a cluster is a collection of related items, pieces from various clusters are not comparable.

**Conclusion:**

This article provides an overview of diagram shading applications in a few areas, however it mostly concentrates on graph coloring applications in software engineering. An idea is used to demonstrate the concept of graph coloring method. Analysts may learn about the graph coloring process and its applications while also gaining knowledge about their specific subjects of study. Some researchers attempted to solve a combinatorial optimization problem using an evolutionary algorithm, which harnesses nature's evolution mechanism to provide a near-optimal solution.

**Reference** –

Ng, R., & Jiawei Han. (2002b). CLARANS: a method for clustering objects for spatial datamining. IEEE Transactions on Knowledge and Data Engineering, 14(5), 1003–1016. <https://doi.org/10.1109/tkde.2002.1033770>

Donato Malerba, Floriana Esposito, Vincenzo Gioviale, & Valentina Tamma. (2000). Comparing Dissimilarity Measures for Symbolic Data Analysis. In: Joint Conferences on New Techniques and Technologies for Statistics and Exchange of Technology and Know-how,

Applications\_of\_Graph\_Coloring\_in\_Modern\_Computer\_Science.(n.d.).https://citeseerx.ist.psu.edu/viewdoc/download? doi=10.1.1.736.8618&rep=rep1&type=pdf