

A Comparative Study of Graph Isomorphism Applications

Rachna Somkunwar
Research Scholar of JJTU
Rajasthan, India

Vinod Moreshwar Vaze, PhD
Research guide at JJTU &
Assistant Professor at JJTU
Rajasthan, India

ABSTRACT

Graphs are powerful data structure to represent objects and their concepts. Objects are nothing but nodes and edges describes relation among objects. Number of real world problems is represented by graph. Graph isomorphism is the area of pattern matching and widely used in various applications such as image processing, protein structure, computer and information system, chemical bond structure, Social Networks. This paper surveys both various applications of graph isomorphism and their importance in the society.

Keywords

Graph, Subgraph, Graph Isomorphism, Pattern matching.

1. INTRODUCTION

From more than 30 years, numerous applications of graph and subgraph have been studied. It includes computer and information system, chemistry, social media, images, protein structure and researchers have implemented algorithms for these types of applications which may increase performance of the system which is also useful for the society [1].

Most of the applications are based on unlabelled graph structure [2][3]. Unlabelled graphs are graphs where labels are not necessary that means all vertices are considered as same. Handling such types of graphs is a very time consuming method.

Pattern recognition is the area where two different objects are matched on the basis of some features. Two types of pattern matching systems are in use. First type is where input is given and on the basis of that input it will match different objects. Second is where input is not given means unknowingly it will match the different objects. Graph Isomorphism is a method to check whether two different graphs are similar or not and subgraph isomorphism is nothing but to identify whether an input graph is a part of full graph or not.

Using number of applications this paper will verify how it is useful for the society and how this differentiates the area of graph and subgraph isomorphism.

The objective of this paper is to focus on various problems and they are discussed below:

1. How graph and subgraph isomorphism will be used in image processing and which one will take more time for matching.
2. Isomorphism issues on social networks.
3. Chemical bond structure type of matching in graph and subgraph.
4. Matching for protein structure.
5. Isomorphism for city map.

2. BASIC CONCEPTS OF GRAPH

Graph is a set of nodes and edges. Nodes represent objects and edges represent relations between objects. Graphs are useful in the area of pattern matching and computer vision. Numerous algorithms are presented for pattern matching but still it is having some complexity issues. Two nodes are adjacent if there is a direct link between them. If source node and destination node both are same then it is said to be Loop. A graph is Simple if it has no loop and parallel edges. If there are no nodes and no edges then it is called as Null Graph. Graphs can be Directed or Undirected. If there is a direct link between two nodes then it is Directed Graph and if no directions are given then it is said to be Undirected Graph. Total number of edges connected to a node is Degree of node. For Directed graph it has indegree and outdegree. Indgree means if there is an incoming edge to that node. Oudegree means there is an outgoing edge from that node. Complete Graph is a graph if all nodes are connected by unique edge or simply if node has a degree $N-1$. Connected Graph is a graph if there is path between every pair of nodes. Graphs are remains same if and only if we are not changing their label.

3. GRAPH ISOMORPHISM

A Graph $G(V, E)$ is a set of vertices and Edges. If there is an Edge between the vertices, it is said to be adjacent of neighbor and if there is a relation between vertices, it is called as adjacency relation. Adjacency matrix is used to represent graph in memory, where 1 represents edge and 0 represents no edge [4].

Graph Isomorphism

- If two graphs are looking same, but number of vertices and edges are not same then it is not isomorphic.
- If there is one to one mapping function between the vertices then it is called as isomorphic.
- If all properties of graph isomorphism are satisfied then two graphs are said to be isomorphic else not. Properties of graph isomorphism are as follows:

Properties of Graph Isomorphism

1. Number of vertices should be equal for both the Graphs.
2. Number of Edges should be Equal for both the Graphs.
3. In Degree and Out Degree should be equal.
4. Number of connected components should be same.
5. Number of loops should be same.
6. Number of parallel edges should be same.

These all are properties to verify if two graphs are isomorphic or not and it is denoted by $G1=G2$.

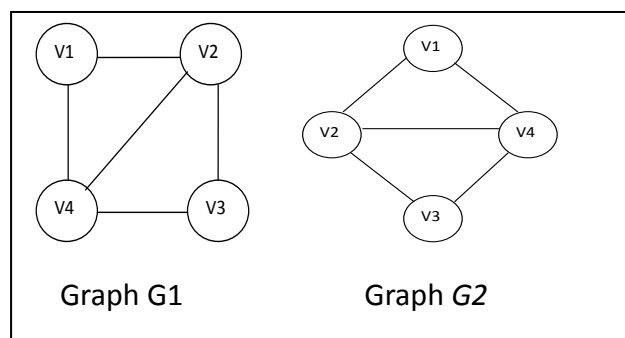


Figure 1. Labeled Graphs

From figure 1, Graph G1 and G2 are Isomorphic graphs.

1. Number of Vertices= $4(G1\&G2)$.
2. Number of Edges = $5(G1\&G2)$
3. Number of Degree= $2(G1\&G2)$.
4. Connectedness= Fully
5. Pairs of connected vertices= Yes.

4. ALGORITHMS AND APPLICATIONS OF GRAPH ISOMORPHISM

Graphs are used to represent an image which will be very useful in the area of Image processing where graph shows structural description of image. In such representation vertices represent region of image and relation of regions are represented by links called as edges [5][6][7].

In [8], author suggested a new concept for application related problems which may improve the efficiency of the system.

A new algorithm is implemented for labeled and unlabeled graphs which are used to verify how recently pattern of images differ from other situated segments of images. They have also discussed about image representation using Quad-tree technique [9].

Subgraph Isomorphism algorithm for chemical structures is Maximum Common Subgraph which is used for matching 2D and 3D Structures. John W. Raymond and Peter Willett has reviewed this MCS algorithm in [10].

Author introduced a subgraph isomorphism algorithm for biochemical data [11]. This algorithm is not using pruning rules and domain reduction procedure which reduces the search space. In this paper author has compared this algorithm with recent algorithms.

In [12], author proposed a technique for modeling, processing and analysis of an image on their system. They have discussed about numerous applications and segmentation of an image, recognition of an object.

Vertex and Edge Approximate graph Miner (VEAM) is an algorithm proposed by [13], which is used for frequent approximate subgraph mining. They have introduced a framework where frequent subgraph act as features. Using this framework they have done image classification by VEAM algorithm. If labels are occurred in matrices then it matches all the labels and mines those vertices which are occurred frequently. They have also used DFS strategy for finding the candidates if any subgraph is using a new edge.

Author aims to develop path finding algorithm for Route in Geographical System [14]. In this paper, A* and Blind search algorithms are compared. Various algorithms are available for searching routes in city map and they are Blind search, A* algorithm, DFS, Iterative DFS, BFS, Bidirectional BFS, Best First Search, Dijkstra algorithm, but here they have focused on Heuristic with A* and Blind search algorithm. Sometimes Blind search algorithm is faster than A* algorithm but not that much accurate. This paper is also used for 2D and 3D city maps.

For finding shortest path authors presented Fuzzy Routing Algorithm (FRA). Aim of this algorithm is to start searching from source till it reaches to its destination. It is used for path cost update process which is advanced version of Dijkstra algorithm to reduce the path request and improve the performance of the system [15].

Several applications of graphs isomorphism are discussed below:

Image processing: Images are represented as objects and it can be 2D, 3D, 4D objects [16]. Information of images can be stored into database and if it grows we need to retrieve the information from the database. First we have to preprocess the image to convert color image into black and white image. To retrieve the information from image Optical Character Recognition (OCR) will be used. For extracting the features from the image Principal Component Analysis (PCA) algorithm will be used. Image processing can be improved by graph theory.

Image analysis is a method by which we can extract the information from the image and that images may be digital image hence digital images can be processed by digital image techniques. For medical kind of images, segmentation is performed i.e. Divide the image into number of meaningful parts [17].

Protein Structure: Nodes represents protein and Edges represents their interactions between nodes. Proteins are represented by networks [18]. To compare two proteins structures number of algorithms are reviewed in [19][20], which plays an important role in biological networks.

Computer Design System: Computer Aided Design or Computer Aided Mechanical Drawings comes under the subgraph Isomorphism applications [21].

For graph Isomorphism applications are finding two states are symmetric or not [22].

Examples of computer science are database designing, software engineering, computer networks etc. [23].

Chemical Structures: In this structure nodes represents atoms and edges represents covalent bonds. Structural formula shows chemical information and this formula should be uniquely indexed and identified [24]. Molecule structures are studied by graph theory.

Social Networks: In this type of networks nodes represents person and edges represents their relationships. Pattern matching is useful method for finding positions and their locations. Facebook, Twitter are some examples of social networks [25]. Social network is a part of sociology. It is a way to identify structure through the social networks. In this type of networks, it is necessary to evaluate shortest path between two persons on the basis of their relation and from this evaluation they can also find their location or psychology

of that person. For this reason, try to create a cluster which will store similar information of different persons.

5. COMPARATIVE DISCUSSION OF GRAPH ISOMORPHISM APPLICATIONS

In image processing, graph isomorphism is used to match two different images. In this type of application, first convert the input image into graph. For converting input image to graph an algorithm is needed so try to create an algorithm for conversion using some features. But it is time consuming process. After getting the graph apply graph isomorphism algorithm to verify whether input graph is same as one of the graph of the dataset. In case of subgraph isomorphism follow the same procedure only the difference is it is very lengthy procedure because instead of searching the complete image, only part of the image to be searched. Figure 2 shows matching an input image from complete image.

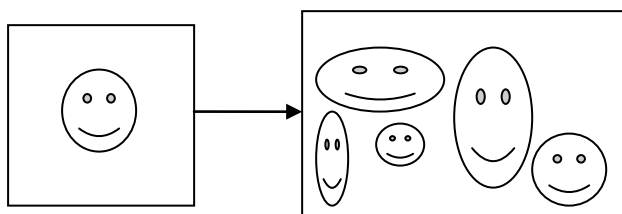


Figure 2. Image based subgraph isomorphism.

In Chemical bond structure, first dataset of chemical structure is to be collected then study each and every structure of bond. Structure of chemical bond is very different than normal graph hence it becomes very tedious job to apply graph isomorphism algorithms for such kinds of structure. Here also first convert bond structure in relevant graph format then go for graph isomorphism. In Chemical bond structure, each node has a unique identification number. If in case it is changed in that case it is assumed that chemical structure is changed. Figure 3 shows chemical structure.

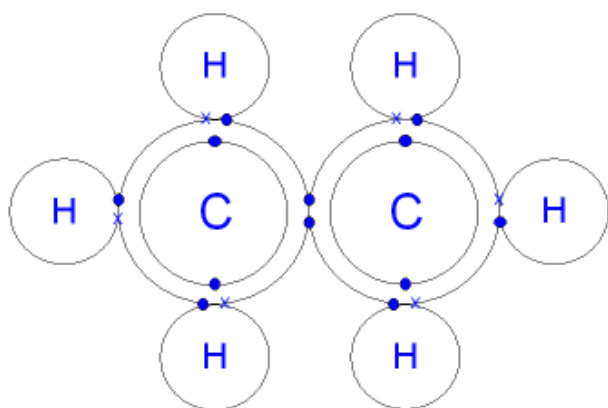


Figure 3. Chemical structure

In social network, to check how many persons are interacting? Example of this type of network is Facebook where numbers of friends are connected to each other.

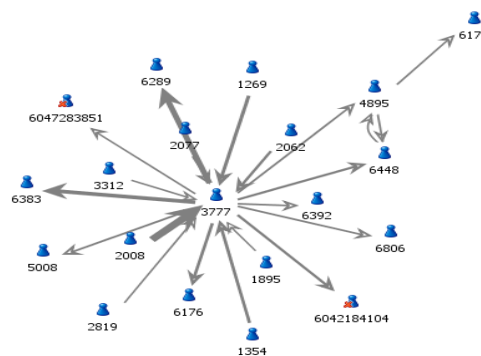


Figure 4. Social networks

Direct or indirect interaction may be done among the friends. So instead of giving name to each friend, try to give unique id to each friend. So it is easy to find how many friends are interacting directly or indirectly. Indirect interaction in terms of friends of friends. If in case their names are changed that means their relation is also changed. Figure 4 shows social networks of people.

In protein structure, subgraph isomorphism will be used. For example collection of food is available and every food is having their own structure that is graph structure. For providing the food and that food contains some proteins, then find protein graph structure first and check where it is available in the food structure or not. Figure 5 shows protein structure and its levels.

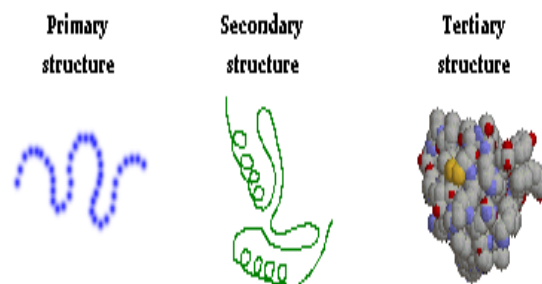


Figure 5. Protein structure

In city map, every city has some unique name in string form but some local cities having same name, here also assign some unique value to each city. So in such case it is easy to find any particular city. Hence subgraph isomorphism algorithm will be used in such kinds of applications.

6. CONCLUSION

Various applications of graph isomorphism and their algorithm are studied thoroughly. These applications play an important role in the society. In this paper graph isomorphism applications and their problems are also discussed. It can be concluded that these applications are useful in pattern matching.

7. REFERENCES

- [1] D. Conte, P. Foggia, C. Sansone, M. Vento, "Thirty Years of Graph Matching In Pattern Recognition", *International Journal of Pattern Recognition and Artificial Intelligence*, vol.18, 2004.
- [2] Avrim Blum, Shuchi Chawla, "Learning from Labeled and Unlabeled Data using Graph Mincuts", pp. 19-26, 2001.
- [3] Pan, Shirui, Xingquan Zhu, Chengqi Zhang, Philip S. Yu "Graph stream classification using labeled and unlabeled graphs." *Data Engineering (ICDE)*, 29th International Conference on. IEEE, 2013.
- [4] Dragos Cvetkovi, "Application of Graph Spectra: An introduction to literature", *Mathematics Subject Classification*, 2000.
- [5] R.A.Finkel, J.L.Brntley, "Quad Trees: A Data Structure for Retrieval on Composite Keys", *Acta Informatics*, pp.1-9, 1974.
- [6] Marfil, Rebeca, and F. Sandoval. "Energy-based perceptual segmentation using an irregular pyramid", *International Work-Conference on Artificial Neural Networks*. Springer Berlin Heidelberg, 2009.
- [7] Morales-González, Annette, and Edel García-Reyes, "Assessing the role of spatial relations for the object recognition task", *Iberoamerican Congress on Pattern Recognition*. Springer Berlin Heidelberg, 2010.
- [8] Peter Czimmermann, "The graph isomorphism problem and its applications", 2003.
- [9] Acosta-Mendoza, Niusvel, Andrés Gago-Alonso, and José E. Medina-Pagola. "Frequent approximate subgraphs as features for graph-based image classification", *Knowledge-Based Systems* pp. 381-392, 2012.
- [10] John W. Raymond, Peter Willett, "Maximum Common Subgraph Isomorphism Algorithms for the Matching of Chemical Structures", *Journal of Computer-Aided Molecular Design*, vol.16, pp.521-533,2002.
- [11] Vincenzo Bonnici, Rosalba Giugno, Alfredo Pulvirenti, Dennis Shasha, Alfredo Ferro, "A subgraph isomorphism algorithm and its application to biochemical data", Bonnici et al. *BMC Bioinformatics*, 2013.
- [12] A. Sanfeliua, R. Alquézarb, J. Andradea, J. Climentc, F. Serratosad and J. Vergésa, "Graph-based representations and techniques for image processing and image analysis", *Pattern Recognition*, pp. 639-650, 2002.
- [13] Niusvel Acosta-Mendozaa,, Andres Gago-Alonsoa, Jose E. Medina-Pagola, "Frequent approximate subgraphs as features for graph-based image classification", Elsevier, 2011.
- [14] Sobota, B., Cs Szabó, and J. Perhac, "Using path-finding algorithms of graph theory for route-searching in geographical information systems", *Intelligent Systems and Informatics*, 6th International Symposium on. IEEE, 2008.
- [15] Boominathan, P., and Kanchan Arora, "Routing planning as an application of graph theory", *International Journal of Scientific & Technology Research*, PP. 61-66, 2014.
- [16] Conte, Donatello, "Graph matching applications in pattern recognition and image processing", *Image Processing Proceeding*, International Conference on IEEE, vol. 2, 2003.
- [17] Barua, Sonali, "A Generic Framework for the Application of Graph Theory to Image Processing", Diss. The University of North Carolina at Charlotte, 2007.
- [18] Ahed Elmsallati, Cannon Clark, Jugal Kalita, "Global Alignment of Protein-Protein Interaction Networks: A Survey", *IEEE/ACM Transaction on Computational Biology and Bioinformatics*, vol. 6, 2007.
- [19] R. Singh, J. Xu, B. Berger, "Global Alignment of Multiple Protein Interaction Networks with Application to Functional Orthology Detection", *Proceedings of the National Academy of Sciences*, vol.105, pp.12763-12768,2008.
- [20] C. Clark, J. Kalita, "A Comparison of Algorithms for the Pairwise Alignment of Biological Networks", *Bioinformatics*, 2014.
- [21] Cordella L.P., Foggia P., Sansone C., Vento M., "Performance Evaluation of the VF Graph Matching Algorithm", *Proceeding of the 10th ICIAP*, IEEE, Computer Society Press, pp. 1172-1177,1999.
- [22] Clarke E.M., Enclers R., Filkorn T., Jha S., "Exploiting Symmetries in Temporal Modal Logic Model Checking", *Formal Methods in System Design*, 1996.
- [23] Singh, Rishi Pal, "Application of Graph Theory in Computer Science and Engineering", *International Journal of Computer Applications*, 2014.
- [24] Alexandru T. Balaban, "Applications of Graph Theory in Chemistry", *J. Chem. Inf Comput. Science*, 1985.
- [25] Wenfei Fan, "Graph Pattern Matching Revised for Social Network Analysis", *ICDT*, 2012.