International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Finding Maximal Clique In A Graph

Wiwin Apriani[1], Novida Sari Sihombing[2], Zullidar Habsyah[3], Saib Suwilo[4]

[1,2,3]Graduate School. of Mathematics, University of Sumatera Utara, Campus USU, Medan 20155, Indonesia
[4]Departement of Mathematics, University of Sumatera Utara, Campus USU, Medan 20155, Indonesia

Abstract: ***Given a simple graph G with n vertices. Furthermore, given an algorithm to get maximal clique in graph G. Clique Q of G is a set of vertices, where each vertex are connected to each other. Clique is called maximal clique when it does not have vertices that adjoinable again. For finding maximal clique Q_i used the Maximal Clique Algorithm. This algorithm can be used to obtain the maximum clique on certain graph models. Maximal cliques can be very small, when a graph may contain a non-maximal clique with many vertices and a separate clique of size two which is maximal. A maximum clique is necessarily maximal clique, but does not converse. There are some types of graphs, which every maximal clique is maximum clique, which is complete graphs, triangle-free graphs, and complete multipartite graphs. But the other graphs have maximal cliques that are not maximum.***

Keywords: Clique, Maximal clique, Maximum clique, and Adjoinable vertices.

1. Introduction

Given an undirected graph by $G = (V, E)$, where V is the set of vertices and E is the set of edges. Two vertices are said to be adjacent if they are connected by an edge. A clique is a maximal set of pairwise adjacent vertices. A set of pairwise nonadjacent vertices is called an independent set. The maximal clique problem is the problem of finding in a given graph the clique with the largest number of vertices. For finding a maximal clique in a graph is an NP-hard problem, and it is difficult to obtain the exact solution efficiently. It is also difficult to obtain even a satisfactory approximate solution. Nevertheless, many practical problems can be formulated as maximal clique problems. Searching for the maximal clique is often the bottle-neck computational step in these applications. The maximal clique problem is NP-hard [2], and probably no polynomial time algorithm will be possible, but improvements to the existing algorithms can still be effective.

In the maximal clique problem, one desires to find one maximal clique of an arbitrary undirected graph. This problem is computationally equivalent to some other important graph problems, for example, the maximal independent set problem and the minimum vertex cover problem. Since these are NP-hard problems [2], no polynomial time algorithms are expected to be found. Applications for this problem exist in signal processing, computer vision and experimental design for example (see [3]).

The algorithms for finding a maximal clique are frequently used in chemical information, bioinformatics and computational biology applications [4], where their main application is to search for similarity between molecules. These algorithms are used for screening databases of compounds to filter out molecules that are similar to known biologically active molecules and are feasible to be active themselves [5]. Also, these algorithms are used for comparing protein structures, to provide the information about protein function [6] and also the information about possible interactions between proteins [7, 8]. This paper for a given undirected graph G and find a maximal clique of G.

In 1972, Karp [1] introduced a list of twenty-one NP-complete problems, one of which was the problem of finding a maximal clique in a graph. In Section 2, provide precise notations and definitions of all the terminology used. In Section 3, present a formal description of the algorithm for finding a maximal clique in a graph. In Section 4, given a example to find a maximal clique with ten vertices in undirected graph shown below in Figure 1. In Section 7, conclusion. In Section 8, list the references.

2. Notations and Definitions

In this section to begin with precise definitions of all the terminology and notation used in this paper. This paper usually use the notation $\lfloor x \rfloor$ to denote the floor function i.e. the greatest integer not greater than x and $\lceil x \rceil$ to denote the ceiling function i.e. the least integer not less than x.

A simple graph G with n vertices consists of a set of vertices V, with $|V| = n$, and a set of edges E, such that each edge is an unordered pair of distinct vertices. Note that the definition of G explicitly forbids loops (edges joining a vertex to itself) and multiple edges (many edges joining a pair of vertices), whence the set E must also be finite. The complement G^C of a graph G is a simple graph with the same set of vertices as G but $\{u, v\}$ is an edge in G^C if and only if $\{u, v\}$ is not an edge in G. This paper may label the vertices of G with the integers $1, 2, ..., n$. If the unordered pair of vertices $\{u, v\}$ is an edge in G, it say that u is a neighbor of v and write $uv \in E$. Neighborhood is clearly a symmetric relationship $uv \in E$ if and only if $vu \in E$. The degree of a vertex v, denoted by $d(v)$, is the number of neighbors of v. The minimum degree over all vertices of G is denoted by δ.

The adjacency matrix of G is an $n \times n$ matrix with the entry in row u and column v equal to 1 if $uv \in E$ and equal to 0 otherwise. A vertex cover C of G is a set of vertices such that for every edge $\{u,v\}$ of G at least one of u or v is in C. An independent set S of G is a set of vertices such that no unordered pair of vertices in S is an edge. A clique Q of G is a set of vertices such that every unordered pair of vertices in Q is an edge. Given a clique Q of G and a vertex v outside Q, we say that v is adjoinable if the set $Q \bigcup \{v\}$ is also a clique of G. Denote by $\rho(Q)$ the number of adjoinable vertices of a clique Q of G. A maximal clique has no adjoinable vertices. A maximum clique is a clique with the largest number of vertices.

An undirected graph $G = (V,E)$ consists of a set of vertices V = {1, 2, ..., n} and a set of edges $E \subseteq V \times V$. Two vertices v and w are adjacent, if there exists an edge $(v, w) \in E$. For a vertex $v \in V$, a set $\Gamma(v)$ is the set of all vertices $w \in V$ that are adjacent to the vertex v. $|\Gamma(v)|$ is the degree of vertex v. Let $G(R) = (R, E \bigcap R \times R)$ be the subgraph induced by vertices in R, where R is a subset of V. The number of vertices in a maximum clique is denoted by $\omega(G)$ [9].

A graph $G = (V, E)$ is complete if all its vertices are pairwise adjacent, i.e. $\forall i, j \in V$ with $i \neq j$, we have $(i, j) \in E$. A clique Q is a subset of V such that $G(Q)$ is complete. The clique number of G denoted by $\omega(G)$ is the size of the maximum clique. The maximum clique problem asks for clique of maximum cardinality (the cardinality of a set Q, i.e., the number of its elements which will be denoted by $|Q|$).

$$\omega(G) = \max\{|Q| : Q \text{ is a clique in } G\}$$

3. The Maximal Clique Algorithm

Definitions: Given a clique Q in graph G, and there is a vertex v outside Q, so it can be said that v adjoinable with Q if $Q \bigcup \{v\}$ is clique in G.

Propositions: Q_k is maximal clique in G, if $v \in G \setminus Q_k$ connected to at most k - 1 vertex in Q_k.

Proof: If there are a vertex $v \in G$ connected to at most k+1 vertex in Q_k, than Q_k not a maximal clique in G.

For finding maximal clique in a simple graph can use basic algorithm is:

Procedure MaxClique:

Set $S := V$ and $Q := \phi$;
while $S \neq \phi$ do
choose a vertex $v \in S$ with maximum degree in G;
and choose an adjoinable vertice v_{adj} in G;
if $Q_{max} = Q + v_{adj}$ then
$Q_{\max} := Q \bigcup \{v_{adj}\}$;
if $v_{adj} = \phi$ then
obtain a maximal clique (Q_{max});
else if $Q_{max} < Q$ then $Q_{max} := Q$;
$Q := Q \setminus \{v\}$;
end while

This algorithm starts from a small clique and continue to find the larges clique until one of the clique is verified that the clique is the one that has the largest size. This algorithm have two sets is Q and Q_{max}, where Q consist of the vertices of the clique before and Q_{max} contains consists of vertices of the largest clique currently found. The algorithm starts with an empty set Q, and then will adds vertices and deletes vertices from Q, until it can verify that no clique with more vertices can be found. The next vertex to be added to Q is selected from the set of candidate vertices $S := V$ with maximum degree and an adjoinable vertices in G. At each step, the algorithm selects an adjoinable vertex $v_{adj} \in S$ with the maximum degree in G. if obtained $Q_{max} = Q + v_{adj}$, then Q of the clique before will be combined with the adjoinable vertices $(Q \bigcup \{v_{adj}\})$. if not there anymore adjoinable vertex $(v_{adj} = \phi)$, then the obtain maximal clique Q_{max}. Else if the maximal clique obtained smaller than previous clique, then the maximal clique is a clique Q itself.

4. Example

In this sections given simple example to demonstrate the steps of the algorithm. Suppose there are 8 villages in the district. These villages is a road between the villages. The villages there are directly connected and not directly connected. For finding maximal clique, this problem can be described in a simple graph form as follow:

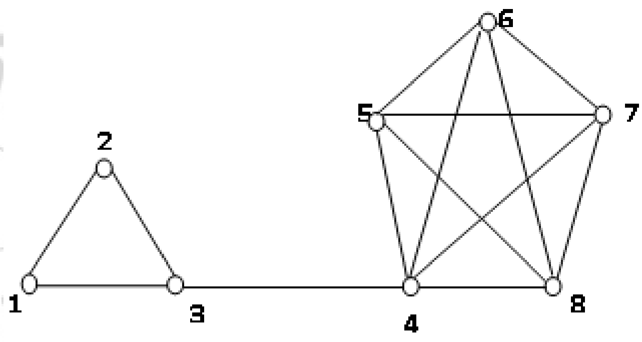

Figure 1: A simple graph to demonstrate the steps of the algorithm

In figure 1. can be searched adjoinable vertex with maximum degree and a clique. Based on the maximal clique algorithm, the process of finding the maximal clique is as follows:

1. In figure 1. there are 8 vertices $V = \{1,2,3,4,5,6,7,8\}$
2. Take the initial vertex with maximum degree i = 4, the result clique $Q_4 = \{i\} = \{4\}$ size 1.
3. Search adjoinable vertex with Q_4, by algorithm, and the results in tabular form:

Table 1: Adjoinable vertex with Q_4

v adjoinable with Q_4	$Q_4 \cup \{v\}$	$\rho(Q_4 \cup \{v\})$
5	6, 7, 8	3
6	5, 7, 8	3
7	5, 6, 8	3
8	5, 6, 7	3

In table 1. is obtained maximal $\rho(Q_4 \cup \{v\}) = 5,6,7,8$ for $v = 5,6,7,8$ the result clique $Q_{5,6,7,8} = Q_4 \cup \{v\} = \{5,6,7,8\}$ size k = 5.

1. The process is completed because there are not vertex adjoinable again. The result is Q = (4, 5, 6, 7, 8) and size k = 5.

The figure is follows:

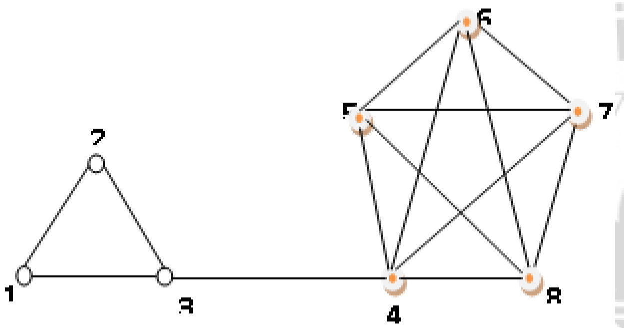

Figure 2: A graph with maximal clique k = 5

The picture above shows that the maximal clique size k = 5. The results is showed that the a maximal clique is a maximum clique. so that the maximal clique algorithm can be used to obtain the maximum clique on certain graph models. A maximal clique sometimes called inclusion-maximal, is a clique that is not included in a larger clique. Therefore, that every clique is contained in a maximal clique. Maximal cliques can be very small, when a graph may contain a non-maximal clique with many vertices and a separate clique of size 2 which is maximal. A maximum clique is necessarily maximal clique, but does not converse. There are some types of graphs, which every maximal clique is maximum clique, that is complete graphs, triangle-free graphs, and complete multipartite graphs. But the other graphs have maximal cliques that are not maximum.

References

[1] R.M. Karp, Reducibility among combinatorial problems, Complexity of Computer Computations, Plenum Press, 1972.
[2] M.R. Garey, D.S. Johnson, Computers and Intractability: A Guide to the Theory of NP-completeness, Freeman, New York, 1979.
[3] E. Balas, C.S. Yu, Finding a maximum clique in an arbitrary graph, SIAM J. Comput. 15(4) (1986) 1054–1068.
[4] S. Butenko, W.E. Wilhelm, Clique-detection models in computational biochemistry and genomics, European Journal of Operational Research 173 (2006) 1-17.
[5] C. Hofbauer, H. Lohninger, A. Asz´odi, SURFCOMP: A novel graphbased approach to molecular surface comparison, Journal of Chemical Information and Computer Science 44 (2004) 837-847.
[6] S. Schmitt, D. Kuhn, G. Klebe, A new method to detect related function among proteins independent of sequence and fold homology, Journal of Molecular Biology 323 (2002) 387-406.
[7] J. Konc, D. Janezic, A branch and bound algorithm for matching protein structures, Lecture Notes in Computer Science 4432 (2007) 399-406.
[8] J. Konc, D. Janezic, Protein-protein binding-sites prediction by protein surface structure conservation, Journal of Chemical Information and Modeling 47 (2007) 940-944.
[9] J. Konc, D. Janezic, An improved branch and bound algorithm for the maximum clique problem, Journal of Mathematical and in Computer Chemistry (2007) 0340-6253