Cliques and Sub-groups

Nomula Rohith Yogi (IMT2016072)

October 5, 2019

Abstract

This paper will talk about the concept of cliques and its variants maximal and maximum cliques along with the notions of plex and core. This paper also briefly talks about the maximal clique problem and the Bron-Kerbosch algorithm that is used to find the maximal ciques in an undirected graph.

1 Introduction

Cliques are referred in the context of a graph. Cliques in the graph are defined as strongly connected components of a graph and strongly connected component in a graph is a largest subset of vertices such that there is path between every pair of vertices in the subset. The famous problem of finding the maximal cliques is said to be NP-Complete. But there are various algorithms that can say whether the given clique is maximal clique or not. One such algorithm is the **Bron-Kerbosch** algorithm which this paper is going to cover briefly about. There are two variants of the cliques: maximal and maximum cliques. Groups in general are referred in the context of social network graphs. According to Reicher, S.D. (1982) a group in a social network is defined as a collection of individuals who consider themselves to be a group. There are also several sub-groups associated within a group. Two types of sub-groups called Plex and Core discussed in this paper. Their generalizations k-plex and N-Core will provide some insights to look deeper into the concept of sub-group.

2 What is a Clique

Formal definition of a clique is as follow:-

Consider simple undirected graph G having V nodes and E edges i.e G(V,E). Let U be subset of nodes U $\subseteq$ V then U is called a clique if G(U) is a complete graph, where a complete graph is defined as a simple undirected graph in which every vertex pair is connected by a unique edge.
By this definition we can infer the following properties of a clique :-

- For a clique of size k each vertices in the clique should have minimum of k-1 degree to be completely connected. (And for a perfectly connected graph there should be k-1 vertices and edge connectivity.)
- Clique must have diameter 1 as there is path from every vertex to every other vertex.
- Cliques are closed under exclusion i.e if U is a clique and v is vertex of clique then U-{v} is also a clique. This can be explained as follows:-

  If U is a clique of size K , $\forall u \in U$ degree(u) is at least K-1, now vertex v is removed from the clique, then all its corresponding edges to remaining K-1 vertices are also removed, which implies that now each of the remaining vertex has minimum of K-2 degree. Consider a vertex set V=U-{v} of size K-1 ,then $\forall u \in V$ degree(u) is at least K-2. Now we can say that U-{v} (i.e V) is also a clique of size K-1. Therefore Cliques are closed under exclusion.

A Clique is also referred to as perfectly dense groups. Perfectly dense groups in a graph with k vertices will have at least k-1 degree for each vertex and is said to be perfectly connected with k-1 vertices and edge connectivity. Few examples of cliques are shown below.

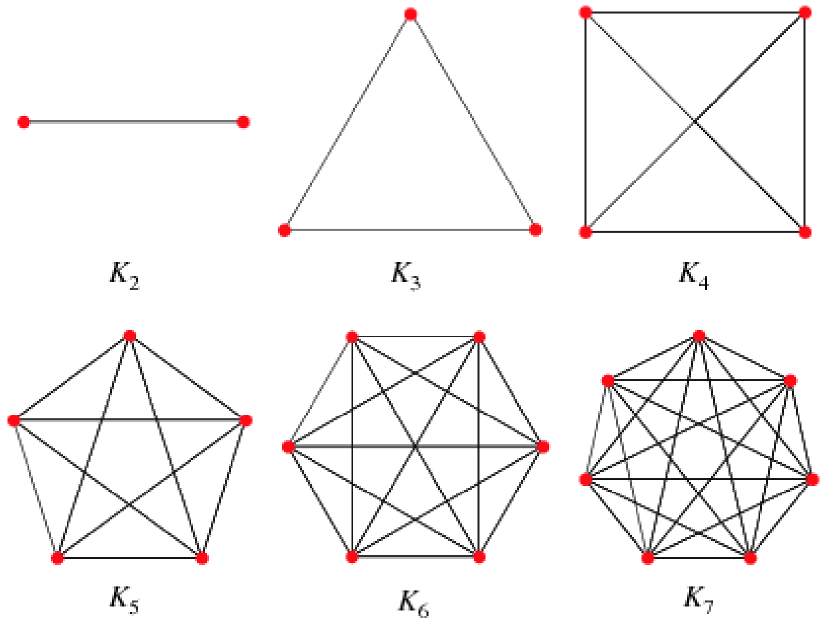

Figure 1: Example of Cliques

3 Maximal and Maximum Cliques

We will talk about two variants of cliques maximal cliques and maximum cliques, Both these variants has their own characteristic definitions.

3.1 Maximal Clique

A maximal clique is a clique that cannot be further extended by adding more adjacent vertices to it. Consider simple undirected graph G having V nodes and E edges i.e G(V,E). Let U be subset of nodes U $\subseteq$ V then U is called a maximal clique if G(U) is a complete graph and if there is no clique W in the graph such that W $\supset$ U .

3.2 Maximum Clique

Maximum Clique is defined as the largest maximal clique in the graph. i.e Among all maximum cliques in the graph ,the clique which has maximum number of vertices is called maximum clique. Consider simple undirected graph G having V nodes and E edges i.e G(V,E). Let {U1,U2,U3.....UK} be set of all maximal cliques in the graph the U^* is called maximum clique if

$$U^* = \arg\max_U N(U)$$

where N(U) is number of vertices of a clique U.
From this we can infer that for every maximum clique is a maximal clique but not vice versa. This can be further explained using an example graph shown below. In this graph we can see that there are both maximal cliques and maximum clique which is largest among them. First we list down all the cliques in the below graph : {6,4},{4,3},{4,5},{5,2},{5,1},{3,2},{2,1},{5,1,2}. Now the maximal cliques are {6,4},{4,3},{4,5},{3,2},{5,1,2}. We have removed {5,2},{5,1} and {2,1} because they can be extended by adding 1,2 and 5 respectively. Here we can see the maximum clique to be {5,1,2} as it is the largest of all the cliques with size 3.

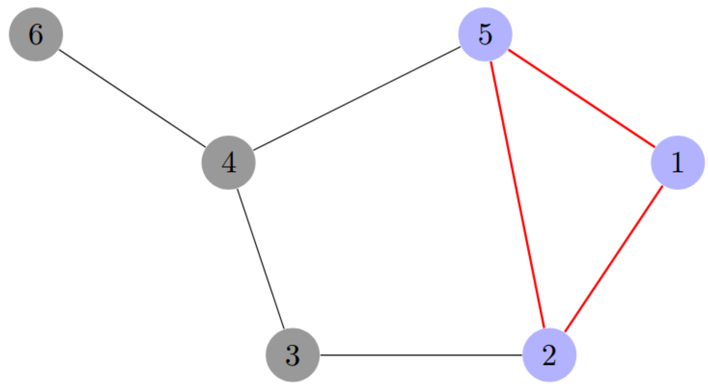

Figure 2: Example of Maximum Clique

4 Plex and Core

4.1 Plex

Plex is a generalized concept of a Clique. Also defined as a generalised notion of clique where the nodes are allowed to "miss" having edges with each other. For example a 1-plex is when a node is allowed to miss one edge i.e for k vertices there has to be at least k-1 edges for each vertex, so the minimum degree should be k-1. This is exactly what a clique is by the previous definitions of clique. Similarly 2-plex is allowed to miss 2 i.e for k vertices there has to be at least k-2 edges for each vertex, 3-plex is allowed to miss 3 and so on and so forth. The formal definition of a N-plex is defined below.

Consider simple undirected graph G having V nodes and E edges i.e G(V,E) with $| V |$= n and $1 \leqslant$ N $\leqslant$ n-1 be a natural number. A subset U $\subseteq$ V is said to be N-plex iff δ(G[U]) $\geqslant$ $| U |$ - N.
where δ(G[U]) is the minimum degree of any node in the graph U.

Similarly to that of Clique, there are a few properties of a Plex that we need to make note of. Plex is closed under exclusion as any sub graph of N-plex is also an N-plex. Consider the 1-plex case where we know 1-plex is a clique and a when we exclude 1 vertex the form a 3 clique the graph formed would be a 2 clique which is also a clique i.e a 1-plex this property is true for every such value of N. Another property is the relationship between N-plex and diameter and connectivity.

An N-plex on a graph with n nodes with n $\geqslant$ 4 and N $< \dfrac{n+2}{2}$ has an edge connectivity at least 2. And N-plex with N $\geqslant \dfrac{n+2}{2}$ has a diameter more than 2N - n +2.

4.2 Core

Core is based on tractable notion of dense subgroup. It is based on the minimum degree in a sub graph. Consider simple undirected graph G having V nodes and E edges i.e G(V,E). A subset U $\subseteq$ V is said to be N-core iff minimum degree in U is at least N. Similar to the notion of maximal clique here we have maximal N-core. So U is said to be maximal N-core if it is not contained in larger N-core within the graph G. A few properties of N-Core are: (N+1)-Core has a minimum degree of N+1 so all we need for N-Core is minimum degree of N so (N+1)-Core is also an N-Core. Another property of core is that any N-core is a (n-N)-Plex.

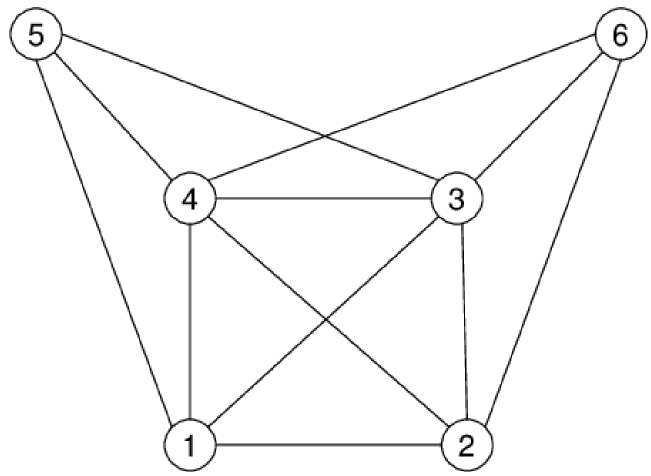

Figure 3: Example of N-plex with N = 1,2,3

5 The Bron-Kerbosch algorithm

In general the detection of maximal clique or N-plex are NP-complete but there are several algorithms that approximately finds the maximal cliques of the graph, one such algorithm is the Bron-Kerbosch algorithm.

Detection of maximal clique is an important problem in graph theory for several reasons. This can help divide large social network graphs into smaller graphs and process it efficiently. This algorithm is extensively used in the field of computational chemistry. We know that till date there is no algorithm that can tell whether a given clique is maximum clique or not in polynomial time. This problem is NP-hard.

The Bron-Kerbosh algorithm is a recursive backtracking algorithm and it basically operates on three sets R, P, X where R is the set of vertices of the maximal clique, P is the set of possible vertices that can be selected as part of the maximal clique and X is the set of vertices that not part of the maximal clique or the vertices that are excluded. The algorithm performs a set operations such as Union, Intersection and Relative compliment. Pseudo Code for this algorithm is given below.

Algorithm 1: The Bron-Kerbosh algorithm

```
Function BronKerbosch1(R,P,X) is
    if P and X are both empty then
        report R as a maximal clique;
    end
    for each vertex v in P do
        BronKerbosch1(R ⋃ v, P ⋂ N(v), X ⋂ N(v));
        P := P \ v;
        X := X ⋃ v;
    end
end
```

This algorithm performs search for all possible maximal cliques in a given graph G(V,E) where V is number of vertices and E is number of edges. We can say that given three disjoint sets of vertices R, P and X which are defined above, it finds the maximal cliques which include all vertices in R, some vertices in P and does not include vertices in X (which were excluded as they are not a part of maximal clique). This algorithm is initiated by setting R to to empty set, P to be the vertex set of the graph and X to be the empty set. The algorithm with each recursive call, considers the vertices in P one by one and checks to see if P and X are both empty if so it reports R as maximal clique else it backtracks.

When these sets are non empty we can see that there are vertices to added to R or there are vertices to be excluded to X. So it recursively checks each and every vertex in P, whether the vertex v from P when added to R forms a maximal clique. So in the next recursive call we include v to be part of maximal clique so the new R would contain v, pass only the neighbours of the vertex v which were included in set of possible vertices that can be selected as part of maximal clique as new P argument (because these are the vertices that can be part of a clique when extended) and pass those vertices that were in X and neighbours of v as new X argument (because they cannot be part of the clique as they were excluded before). After this call we remove the vertex v from P to exclude it from next recursive built cliques and add it to X as these are the new set of vertices that are now excluded. Then the algorithm continues with the next vertex in P.

This algorithm is inefficient in the case of graph having many non maximal cliques as the recursion runs deep. This is the basic form of the Bron-Kerbosh algorithm. The worst case complexity of this algorithm is $O(3^{\frac{n}{3}})$.

6 Conclusions

Cliques and its variants maximal and maximum cliques are in general very hard to detect, as said earlier detecting or finding a maximal clique is NP-complete. But there are approximated algorithms that strive to be very efficient in practice. The Bron-Kerbosh algorithm is one such algorithm and it serves as base for many of today's state of the art algorithms for maximal clique detection.

References

[1] www.dcs.gla.ac.uk/ pat/jchoco/clique/enumeration/papers/MScprojectofDavidSteiner.pdf

[2] https://math.stackexchange.com/questions/61735/maximal-and-maximum-cliques

[3] https://towardsdatascience.com/graphs-paths-bron-kerbosch-maximal-cliques-e6cab843bc2c

[4] https://faculty.ucr.edu/ hanneman/nettext/C11Cliques.html

[5] https://www.slideshare.net/bron-kerbosch

[6] https://en.wikipedia.org/wiki/BronKerboschalgorithm

[7] https://www.researchgate.net/figure/ILLUSTRATION-OF-k-PLEXES-FOR-k-1-2-3