Network Science

Kexing Ying

May 15, 2020

Contents

1 Fundamental Graph Theory **1**
1.1 Basic Concepts 1
1.2 Graph as Models 2

2 Working with Networks **2**
2.1 Degree Distribution 3

1 Fundamental Graph Theory

1.1 Basic Concepts

We begin with a few basic definitions.

Definition 1.1 (Graph)**.** A graph G is a tuple $(V(G), E(G))$ equipped with a function $\sim: E(G) \to V(G) \to V(G) \to \text{Prop}$ where $V(G)$ is the *vertex set*, $E(G)$ is the *edge set* and for all $e \in E(G)$ there exists a unique pair $v_1, v_2 \in V(G)$ such that $v_1 \sim_e v_2$. We write $v_1 \sim_e v_2$ as a short hand for $\sim (e, v_1, v_2) = \text{true}$.

Note that this definition works for both directed and undirected graphs as by this definition, a undirected graph is a directed graph with the condition that for all $e \in E(G)$, $\sim_e$ is symmetric.

Definition 1.2 (Subgraph)**.** Let G be a graph, then H is a subgraph of G if and only if H is a graph such that $V(H) \subseteq V(G)$, $E(H) \subseteq E(G)$ and the restriction $\sim^G|_H = \sim^H$. We will write $H \leq G$ for H is a subgraph of G.

Definition 1.3 (Loop)**.** Let $G := (V(G), E(G))$ be a graph and $e \in E(G)$ be an edge. We say e is a loop at some $v \in V(G)$ if and only if $v \sim_e v$.

Definition 1.4 (Multiple Edges)**.** Let $G := (V(G), E(G))$ be a graph and $e, f \in E(G)$ be edges. We call e, f be multiple edges if and only if there exists $v_1, v_2 \in V(G)$ such that $v_1 \sim_e v_2$ and $v_1 \sim_f v_2$.

Definition 1.5 (Simple)**.** We call a graph simple if it contains no loops nor multiple edges.

If a graph is simple we can then model the edge set of the graph $E(G)$ by a set of unordered tuples where each edge e with end points v_1, v_2 can be uniquely represented by $e = v_1v_2$

(commutative if and only if G is undirected).

Definition 1.6 (Adjacent)**.** Let $G := (V(G), E(G))$ be a graph and $v_1, v_2 \in V(G)$, then v_1 and v_2 are adjacent (or are neighbours) if and only if there exists some edge $e \in E(G)$ such that $v_1 \sim_e v_2$.

Definition 1.7 (Path)**.** Let G be a graph, then a path in G is a simple subgraph P of G such that $V(P)$ can be ordered in a list such that consecutive vertices are adjacent. On top of this, if this ordering resulted in the first element to be adjacent to the last, then we say P is a *cycle*.

1.2 Graph as Models

Definition 1.8 (Complement)**.** Let G be a simple graph, then the complement of G, $\bar{G}$ is the simple graph $(V(G), E(\bar{G}))$ where for all $u, v \in V(G)$, $uv \in E(\bar{G})$ if and only if $uv \notin E(G)$.

Note that this complement graph is unique only if we restrict it to be simple. Suppose G is simple and let $v \in V(G)$, then $vv \notin E(G)$ by the no loop condition. Thus, if we do not restrict $\bar{G}$ to be simple, then we can add how many loops as we want at v, making the complement not unique.

Proposition 0.1. *Let G be a simple graph, then the complement of G is unique.*

Proof. Let G_1, G_2 be complements of G. By definition $V(G_1) = V(G) = V(G_2)$ so $G_1 = G_2$ if and only if $E(G_1) = E(G_2)$. Wlog. it suffices to show that $E(G_1) \subseteq E(G_2)$. Let $uv \in E(G_1)$, then $uv \notin E(G)$ and thus $uv \in E(G_2)$. $\square$

Let us consider a real world problem. Suppose we have n job openings and k applicants but not all applicants are qualified for all jobs. We can easily model this problem by connecting each applicants to their respective qualified jobs and ask whether we can find a subgraph that consist of n pairwise disjoint edges.

Upon examining this question, we find that this particular model has an interesting graph structure in which none of the jobs are adjacent to each other (similarly for the applicants). This type of graphs are called *bipartite* and the set vertices representing people and jobs respectively are called independent.

Definition 1.9 (Independent)**.** Let G be a graph and $S \subseteq V(G)$. S is called an independent set in G if and only if for all $u, v \in S$, $uv \notin E(G)$.

Definition 1.10 (Bipartite)**.** A graph G is called bipartite if and only if $V(G)$ is the disjoint union of two independent sets in G. We call these two independent sets the *partite* sets of G.

2 Working with Networks

In this section we will be less rigorous and focus more on the methods used to analysis graphs (networks) especially really large ones.

2.1 Degree Distribution

Network systems vary in size but they are normally very large (that is they are large enough such that we can't draw them by hand), so, in order to analyse large networks, it is often useful to take a probabilistic approach.

Definition 2.1 (Degree of a Vertex of a Undirected Graph)**.** Let $(V(G), E(G))$ be a undirected graph and $v \in V(G)$, if v is the end point of some $e \in E(G)$, then we say v and e are *incident*. Then the degree of v is the number of incident edges.

Suppose we denote the degree of some vertex v by $d(v)$, then we find the total number of edges is simply

$$|E(G)| = \frac{1}{2} \sum_{v \in V(G)} d(v).$$

Note that the 1/2 factor is because each edge is incident to two vertices.

Definition 2.2 (Average Degree of a Undirected Graph)**.** Let $G = (V(G), E(G))$ be a undirected graph, then the average degree of G is $\frac{1}{|V(G)|} \sum_{v \in V(G)} d(v) = 2\,|E(G)| \,/\, |V(G)|$.

The above, however, does not simply transfer to directed graphs since we would loss the information of "directedness" of the graph. Therefore, the degree is defined slightly differently for directed graphs.

Definition 2.3 (Degree of a Vertex of a Directed Graph)**.** Let $(V(G), E(G))$ be a directed graph and $v \in V(G)$, then the degree of v is simply the difference between number incoming edges and the number of out going edges.

With the definition above, we see straight away the sum of the degrees of all vertices in a directed graph is zero so the definition for average degree does not apply for digraphs[1] either. Therefore, instead defining the average degree by averaging the sum of degrees, we use the average of the sum of either the incoming or outgoing degrees (both of which are equal).

Let us now consider the *degree distribution*, p_k, a characterisation of a graph that provides the probability that a randomly selected vertex has k degree.

Straight away, given some graph $(V(G), E(G))$, let $S := \{v \in V(G) \mid d(v) = k\}$, then $p_k = |S| \,/\, |V(G)|$.

[1] Digraph is an alternative word to directed graph.