Notations and definitions

Ankith Mohan

1 Notations

1. *Graph* $\mathcal{G}(\mathcal{V}, \mathcal{E})$ represents a collection of nodes $\mathcal{V}$, interconnected by a set of edges $\mathcal{E}$. The number of nodes $|\mathcal{V}| = n$ and the number of edges $|\mathcal{E}| = m$.
2. *Edgeset* $\mathcal{E} = \{e_1, \ldots, e_m\}$ where the i-th edge, $e_i = (u, v); u, v \in \mathcal{V}$. Each edge can either be directed or undirected.
3. *Adjacency matrix* $\mathbf{A}$ is used to represent a graph. $\mathbf{A}_{u,v} = \mathbb{1}[u \text{ and } v \text{ are connected}]$.
4. *Subgraph* $\mathcal{S}(\mathcal{V}_{\mathcal{S}}, \mathcal{E}_{\mathcal{S}})$ is a subgraph of $\mathcal{G}$, where $\mathcal{V}_{\mathcal{S}} \subseteq \mathcal{V}$ and $\mathcal{E}_{\mathcal{S}} \subseteq \mathcal{E}$.
5. *Node feature* $x_v \in \mathbb{R}^d$ describes the attributes of node v. Node feature matrix $X \in \mathbb{R}^{n \times d}$.
6. *Edge feature* $x^e_{u,v} \in \mathbb{R}^d$ descibes the attributes of the edge (u, v). Edge feature matrix $X^c \in \mathbb{R}^{m \times c}$.
7. *Label* is a target associated with either a node Y_v, an edge $Y_{u,v}$, a subgraph $Y_{\mathcal{S}}$, or a graph $Y_{\mathcal{G}}$.

2 Definitions

Walk A walk of length l from node v_1 to v_l is a sequence of nodes and edges $v_1 \xrightarrow{(v_1,v_2)} v_2 \xrightarrow{(v_2,v_3)} \ldots \xrightarrow{(v_{l-2},v_{l-1})} v_{l-1} \xrightarrow{(v_{l-1},v_l)} v_l$.

Path A walk with all distinct nodes.

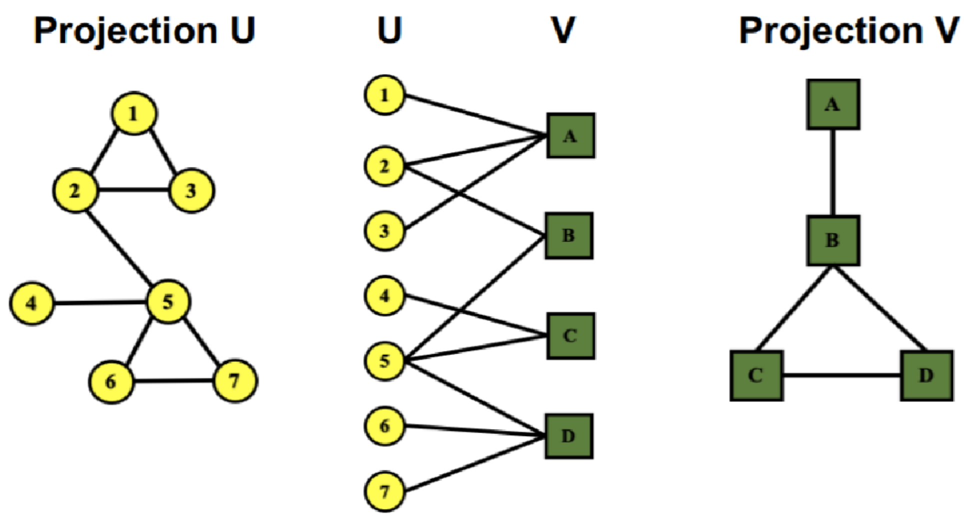

Figure 1: Example of a bipartite graph and the projections of the independent sets U and V (Source: CS224W)

Distance Distance *dist(u,v)* is the length of the shortest path between nodes u and v.

Diameter Longest distance among all pairs of nodes in the graph.

Bipartite graph A graph whose nodes can be divided into two disjoint sets U and V such that every edge can only connect a node in U to one in V, i.e, U and V are independent sets. Figure 1 illustrates an example of a bipartite graph.

Neighborhood For a node v, its neighborhood $\mathcal{N}(v)$ is the collection of nodes that are connected to v.

k-hop neighborhood is the set of all the nodes that are k hops away from node v: $\mathcal{N}^k(v) = \{u | dist(u,v) = k\}$

Clustering coefficient For a node v, it measures how connected the nodes of $\mathcal{N}(v)$ are to one another. Formally, $e_v = \frac{\#(\text{edges among the nodes of } \mathcal{N}(v))}{\binom{deg(v)}{2}}$ Figure 2 gives examples.

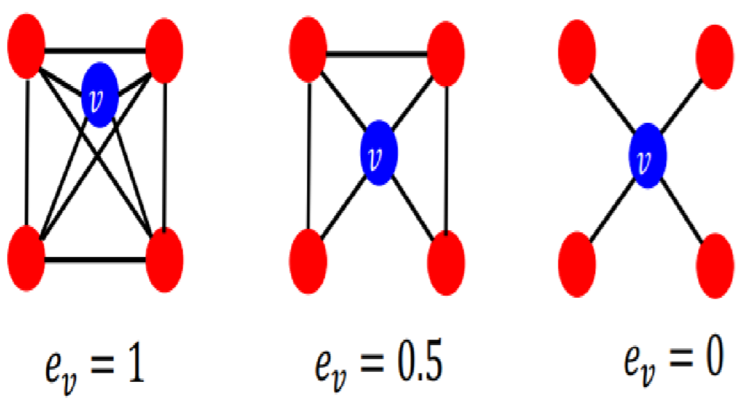

Figure 2: Examples of clustering coefficient of graphs (Source: CS224W)