Lecture 2: Routing parameters: Diameter and expansion

12th and 13th August 2008

In this lecture, we will study some routing parameters of a network viz *diameter*, *edge expansion* and *node expansion* which can have a major bearing on the number of steps required to route packets through it.

2.1 Diameter of a Graph

The diameter of a graph is an important parameter to study routing in networks. Consider a network represented by a line graph on n nodes.

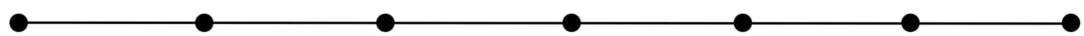

Figure 1: A line graph

To send m packets in the network, it can take $(n-1+m)$ steps. The factor n in this expression is due to the fact that the *diameter* of a graph is a natural lower bound on the number of steps required, since it is the maximum distance a packet may have to travel.

Definition 2.1 *Let $G = (V, E)$ be a graph. For any u,$v \in V$, let $d(u, v)$ denote the shortest distance between u and v. Then the diameter of the graph G, denoted by $Diam(G)$ is defined as:*

$$Diam(G) \;=\; \max_{u,v \in V} d(u, v).$$

Before proceeding further, let us look at diameters of some standard graphs having n nodes.

- When G is a complete graph of n vertices, $\mathrm{Diam}(G) = 1$.
- When G is a Star of n vertices, $\mathrm{Diam}(G) = 2$. (See figure 2)

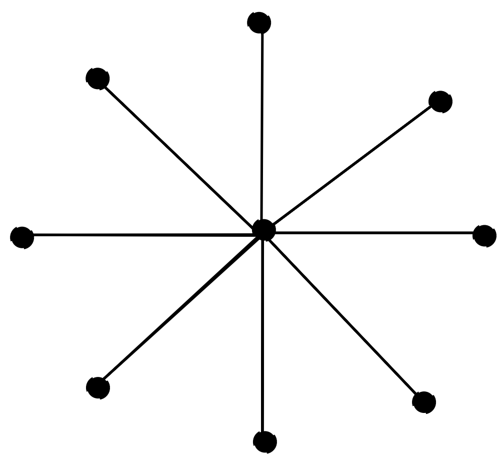

Figure 2: A star graph on n nodes

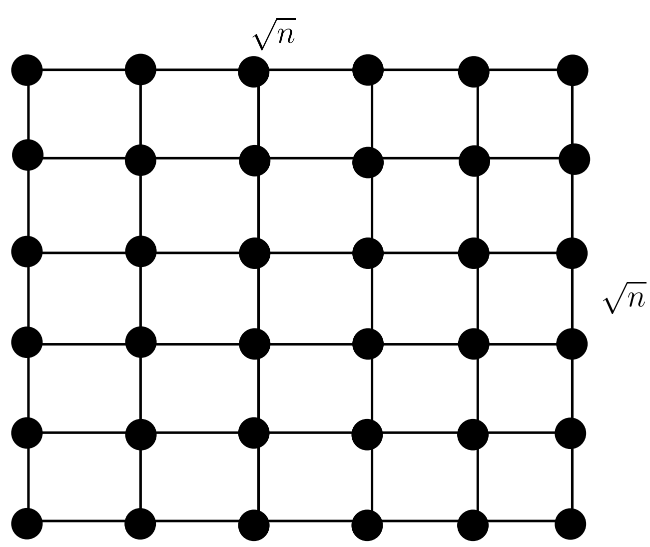

Figure 3: A mesh on n nodes

- When G is a line/path graph of n vertices, $\mathrm{Diam}(G) = n - 1$. (See figure 1).
- When G is a $\sqrt{n} \times \sqrt{n}$ mesh, $\mathrm{Diam}(G) = 2\sqrt{n}$. (See figure 3)

So one can see that there's a lot of variation in the diameters of the graphs. One can also see that as the maximum degree Δ of the graph increases, the diameter decreases.

The next theorem formalises this correlation between maximum degree of a graph and its diameter.

2.1.1 Relation between Diameter and degree of a graph

Theorem 2.2 *Let $G = (V, E)$ be any graph with maximum degree $d > 2$ and size n, then Diam(G) is at least* $\lfloor \log_{d-1} n \rfloor - 1$

Proof. Let v be any arbitary vertex in V. Consider the breadth first traversal of V starting with v. Now the BFS tree formed will have v at level 0. v will have at most d neighbors which occupy the level 1 of the tree. So level 1 will have at most d nodes. Each node at level 1 has at most d neighbors and one of these is in level 0 while the rest occupy level 2. So level 2 will have at most $d(d-1)$ nodes. Let n_i denote the number of the nodes in the graph covered up to level i of the BFS tree rooted at v. From the above discussion we conclude that:

$$\begin{aligned}
n_0 &= 1 \\
n_1 &\leq 1 + d \\
n_2 &\leq 1 + d + d(d-1) \\
n_i &\leq 1 + d \sum_{k=1}^{i-1} (d-1)^k
\end{aligned}$$

Suppose the Breadth First Traversal of V finishes at level k. Hence, $\mathrm{Diam}(G) \geq k$ because for every vertex u in the level k of the BFS tree, $d(u, v) = k$. Also,

$n_k = n$. Hence it follows:

$$\begin{aligned} n &\leq 1 + d\sum_{j=1}^{k-1}(d-1)^j \\ &\leq 1 + \frac{d[(d-1)^k - 1]}{d-2} \\ &\leq 1 + \frac{d(d-1)^k}{d-2} \end{aligned}$$

Since $d > 2$,

$$\frac{d(d-1)^k}{d-2} > 1$$

Hence it follows that

$$n \leq \frac{2d(d-1)^k}{d-2}$$

Taking $\log_{d-1}$ on both sides,

$$\log_{d-1} n \leq k + \log_{d-1} 2 + \log_{d-1} \frac{d}{d-2}$$

$\log_{d-1} \frac{2d}{d-2}$ is smaller than 1 for $d > 4$ while for $d = 3$ and $d = 4$, its only slightly bigger than 1. Hence

$$\lfloor \log_{d-1} n \rfloor - 1 \leq k \tag{1}$$

Since the vertex v was arbitarily chosen, for any $u,v \in V$, $d(u,v) \leq k$. So, $\text{Diam}(G) = k$ and from equation (1) we get,

$$\text{Diam}(G) \geq \lfloor \log_{d-1} n \rfloor - 1$$

This concludes the proof. ■

The above theroem gives the minimum diameter of graphs with maximum degree greater than 2. What about graphs with maximum degree equal to 2? Such graphs are either line/path graphs or they are cycles. So the minimum diameter for such graphs is $\lfloor \frac{n}{2} \rfloor$ when the graph is a cycle.

For a complete binary tree which has maximum degree $d = 3$, the diameter is $2\lfloor \log_2 n \rfloor$ which is asymptotically the same as the lower bound stated by theorem 2.2.

The natural question that arises now is that does any graph match the lower bound given by theroem 2.2 exactly i.e. is the lower bound tight? The following result throws some light on this.

Theorem 2.3 *For every even $d > 2$, there is an infinite family G_n of graphs with maximum degree d and diameter at most $\lfloor \log_{\frac{d}{2}} n \rfloor$.*

The proof of the above result is beyond the scope of this course but it shows that the lower bound given by theorem 2.2 is indeed tight.

2.2 Edge Expansion and Node Expansion of a Graph

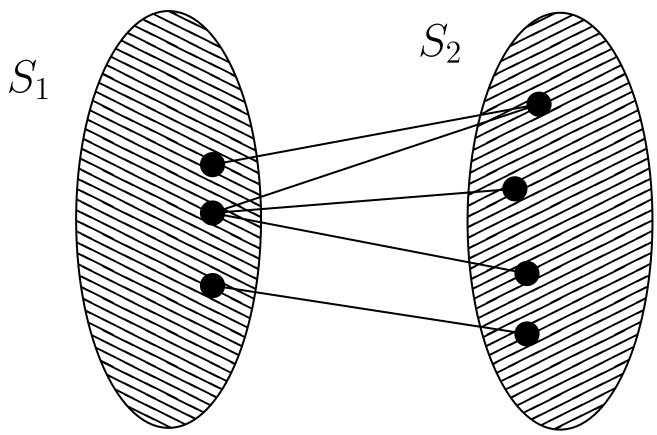

Figure 4: Interface of two sets

To understand the next routing parameter, consider two disjoint subsets S_1 and S_2 of V in a graph $G = (V, E)$, and we have to send packets from S_1 to S_2. Then we would be interested in the ratio of edges going from S_1 to S_2 to the size of S_1. This basically gives us the idea of how many packets can be simultaneously sent from the set S_1 to set S_2,that is, the size of *interface* between S_1 and S_2 (see figure 4), ignoring the internal travel of packets inside the sets. Clearly, this affects how many steps it will take to send m packets in the above situation.
Let us formalize this concept:

Definition 2.4 *Let $G = (V, E)$ be a graph and let $S \subseteq V$. Let $\bar{S} = V - S$ and let $C(S, \bar{S})$ denote the cut between the partition S and $\bar{S}$. Sparsity of the subset S is denoted as $\alpha(S)$ and is defined as:*

$$\alpha(S) = \frac{|C(S, \bar{S})|}{\min\{S, \bar{S}\}}$$

Definition 2.5 *Let $G = (V, E)$ be a graph. The Edge Exapnsion of G denoted as α is defined as:*

$$\alpha = \min_{S \subseteq V} \alpha(S)$$

The *edge expansion* of a graph is a measure of the maximum load any bipartition in the vertex set of the graph will put on the interface separating the 2 vertex sets. As we did in section 2.1, let us look at the edge expansion of some standard graphs having n nodes.

- When G is a line/path graph of n vertices, edge expansion of G is $\frac{1}{\lfloor \frac{n}{2} \rfloor}$. The cut which achieves this minimum is the edge which divides the line graph into 2 equal halves.

- $G = (V, E)$ is a complete graph of n vertices.

  Claim 2.6 *The edge expansion of a complete graph of n vertices is $\frac{n}{2}$.*

  Proof. Let S and $\bar{S}$ be any cut in G. Let $|S| = x$. Without loss of generality, let $x \leq \frac{n}{2}$. Then:

  $$\begin{aligned} \alpha(S) &= \frac{|C(S, \bar{S})|}{\min\{S, \bar{S}\}} \\ &= \frac{x(n-x)}{x} \\ &= n - x \end{aligned}$$

  Since $x \leq \frac{n}{2}$, from definition of edge expansion we have $\alpha = \frac{n}{2}$. This concludes the proof of the claim. ■

- G is a Star of n vertices.

  Claim 2.7 *The edge expansion of a star of n vertices is* 1.

Proof. Let S and $\bar{S}$ be any cut in G. Let $|S| = x$. Without loss of generality, let $x \leq \frac{n}{2}$. Now either the $(n-1)$ degree vertex lies in S or in $\bar{S}$.

In the former case,

$$\begin{aligned}
|C(S,\bar{S})| &= n - x \\
\alpha(S) &= \frac{|C(S,\bar{S})|}{\min\{S,\bar{S}\}} \\
&= \frac{(n-x)}{x} \\
&\geq 1
\end{aligned}$$

When the $(n-1)$ degree vertex lies in $\bar{S}$,

$$\begin{aligned}
|C(S,\bar{S})| &= x \\
\alpha(S) &= \frac{|C(S,\bar{S})|}{\min\{S,\bar{S}\}} \\
&= \frac{x}{x} \\
&= 1
\end{aligned}$$

Hence the minimum value of $\alpha(S)$ for any subset S of V is 1. This concludes the proof of the claim. ■

Note that the star and a complete graph with n vertices have diameters of the same order but their edge expansions have a significant difference. So how do these 2 characteristics of a graph relate to each other? Let us try to relate the two parameters we have studied till now, Diamater and Edge Expansion. It is intuitively clear that more the degree of the vertices, greater will be the edge expansion. We also proved a relation where maximum degree of a graph was negatively correlated to the diameter of the graph. Hence, on combining these two intuitions, we should get that edge expansion should also be negatively correlated to diameter. This is indeed the case as we'll see below.

2.2.1 Relation between Diameter and Edge Expansion

Theorem 2.8 *Every graph $G = (V, E)$ of size n and maximum degree d has diameter at most $2\log_{(1+\frac{\alpha}{d})} n$.*

Proof. For any k, $\text{Diam}(G) \leq k$ iff for every $u, v \in V$, $d(u,v) \leq k$.

Choose any two vertices u and v in G. Consider the Breadth First Traversal of V starting with u.

Let $S_{(u,i)}$ denote the nodes which have been covered uptil the level i of the BFS tree rooted at u. Let k_u be the largest number such that $|S_{(u,k_u)}| < \frac{n}{2}$. Hence, $S_{(u,k_u+1)} \geq \frac{n}{2}$. Also for any $i \leq k_u$, $|S_{(u,i)}| < \frac{n}{2}$. Since, $|S_{(u,i)}| < \frac{n}{2}$, there are at least $\alpha|S_{(u,i)}|$ edges in the cut $(S_{(u,i)}, V - S_{(u,i)})$. Out of these, at most d edges will go the same vertex. Hence, it follows that for every $i \leq k_u$,

$$\begin{aligned} |S_{(u,i+1)}| &\geq |S_{(u,i)}| + \frac{\alpha}{d}|S_{(u,i)}| \\ &\geq |S_{(u,i)}|(1 + \frac{\alpha}{d}) \end{aligned}$$

Hence, it follows that:

$$\frac{n}{2} > |S_{(u,k_u)}| \geq (1 + (\frac{\alpha}{d}))^{k_u}$$

Taking $\log_{1+\frac{\alpha}{d}}$ on both sides,

$$\begin{aligned} k_u &< \log_{1+\frac{\alpha}{d}} \frac{n}{2} \\ &< \log_{1+\frac{\alpha}{d}} n - \log_{1+\frac{\alpha}{d}} 2 \end{aligned} \quad (2)$$

Since u was arbitarily chosen, the inequality (2) holds for the BFS Tree rooted at v also that is:

$$k_v < \log_{1+\frac{\alpha}{d}} n - \log_{1+\frac{\alpha}{d}} 2 \quad (3)$$

Since $|S_{(u,k_u+1)}| \geq \frac{n}{2}$, $|S_{(v,k_v+1)}| \geq \frac{n}{2}$ and $|S_{(u,k_u+1)} \cup S_{(v,k_v+1)}| \leq n$, the intersection of $S_{(u,k_u+1)}$ and $S_{(v,k_v+1)}$ is non-empty. Let $w \in S_{(u,k_u+1)} \cap S_{(v,k_v+1)}$. So there's a path from u and v to w in at most $k_u + 1$ and $k_v + 1$ steps respectively. Hence, there's a path from u to v in at most $k_u + k_v + 2$ steps. Hence,

$$d(u,v) \leq k_u + k_v + 2$$

And from equations (2) and (3) it follows:

$$d(u,v) \leq 2log_{1+\frac{\alpha}{d}} n + 2 - 2\log_{1+\frac{\alpha}{d}} 2$$

Since d is the maximum degree, $\frac{\alpha}{d} \leq 1$. Hence, $\log_{1+\frac{\alpha}{d}} 2 \geq 1$ and therefore we obtain that for any two vertices u,$v \in V$,

$$d(u,v) \leq 2log_{1+\frac{\alpha}{d}} n$$

and therefore we conclude that:

$$\text{Diam}(G) \leq 2log_{1+\frac{\alpha}{d}} n.$$

This concludes the proof. ■

Note that the factor of $\frac{\alpha}{d}$ occurs instead of α because the value of α tells us about the minimum size of any cut $(S, V-S)$ in the graph but in this case we are interested in the number of vertices in $V-S$ which are adjacent to vertices in S. The node expansion of a graph covers precisely this property.

Definition 2.9 *Let $G=(V,E)$ be any graph. Then node expansion of G denoted as α_v is defined as:*

$$\alpha_v = \min_{|U| \leq \frac{|V|}{2}} \frac{|\Gamma(U)|}{|U|}$$

where $\Gamma(U)$ is the set of vertices in $V-U$ that are adjacent to vertices of U.

Note that node expansion of a graph is never greater than 1 because for any $U \subseteq V$ such that $|U| = \frac{|V|}{2}$, $|\Gamma(U)| \leq |U|$.

Now in Theorem 2.8, we can replace $\frac{\alpha}{d}$ by α_v and the theorem still holds. The proof will proceed on the same lines and instead of using $|S|\frac{\alpha}{d}$ as the lower bound for number of vertices in $V-S$ that are adjacent to vertices in S , we use the lower bound $|S|\alpha_v$ for any subset S of V statisfying $|S| \leq \frac{n}{2}$. Hence, for any graph G of size n,

$$\text{Diam}(G) \leq 2log_{(1+\alpha_v)} n$$

2.2.2 Edge expansion is achieved by connected sets

Theorem 2.10 *Given a graph $G=(V,E)$ with edge expansion α, there exists $S \subseteq V$ such that $G[S]$ and $G[V-S]$ are connected and $\alpha(S)=\alpha$, where $G[S]$ denotes the subgraph of G induced by S.*

Proof. We will prove the above theorem by contradiction. So let us assume that there exists no such subset S. Let U be the set whose $\alpha(U) = \alpha$. Without loss of generality, $|U| \leq \frac{|V|}{2}$.

If U and its complement are connected, we have found the subset S. If that is not the case, then the following cases arise:

Case 1: $G[U]$ isn't connected. Then, let U_1,$U_2 \cdots U_k$ be the components of $G[U]$. So for any distinct i,j there's no edge between U_i and U_j. Now

$$C(U, \bar{U}) = \bigcup_{i=1}^{k} C(U_i, \bar{U})$$

Since there's no edge between U_i and U_j for $i \neq j$, it follows:

$$\begin{aligned} C(U, \bar{U}) &= \bigcup_{i=1}^{k} C(U_i, \bar{U_i}) \\ \alpha &= \frac{|C(U, \bar{U})|}{|U|} \end{aligned}$$

Since $C(U, \bar{U})$ is a disjoint union of $C(U_1, \bar{U_1})$, $C(U_2, \bar{U_2}) \cdots C(U_k, \bar{U_k})$, it follows that:

$$\alpha = \frac{\sum_{i=1}^{k} |C(U_i, \bar{U_i})|}{\sum_{i=1}^{k} |U_i|}$$

So there exists i such that:

$$\begin{aligned} \frac{|C(U_i, \bar{U_i})|}{|U_i|} &\leq \alpha \\ \alpha(U_i) &\leq \alpha \end{aligned}$$

Since α is the minimum among all sparsities, it follows that $\alpha(U_i) = \alpha$. Also, $G[U_i]$ is connected. Hence we can always find a subset U such that ,$|U| \leq \frac{|V|}{2}$, $G[U]$ is connected and $\alpha(U) = \alpha$. Hence this case isn't valid.

Case 2: $G[U]$ is connected and $G[V - U]$ is not. Then, let U_1,$U_2 \cdots U_k$ be the components of $G[V - U]$ (See figure 5).

We know that:

$$C(U, \bar{U}) = \bigcup_{i=1}^{k} C(U, U_i)$$

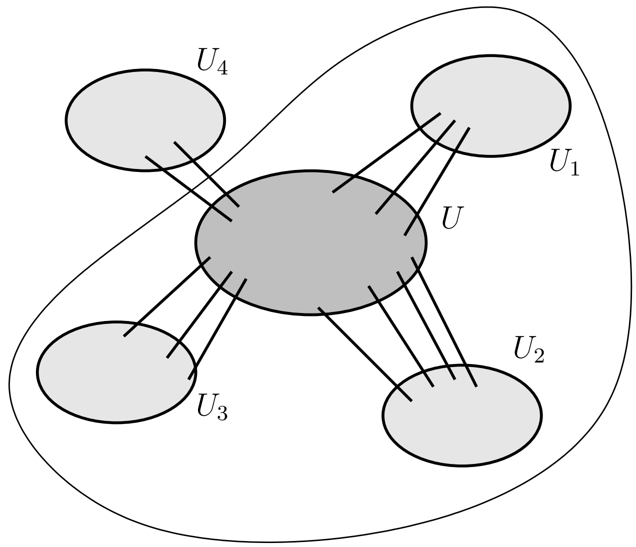

Figure 5: Theorem 2.10 case 2.2

Again since there's no edge between U_i and U_j for $i \neq j$, it follows:

$$C(U, \bar{U}) = \bigcup_{i=1}^{k} C(U_i, \bar{U}_i) \tag{4}$$

Note that the above union is disjoint. Now let us consider two subcases:
Case 2.1: Every component $U_1, U_2 \cdots U_k$ of $V - U$ has size less than or equal to $\frac{n}{2}$.

Now since $|U| \leq \frac{n}{2}$, $|U| \leq |\bar{U}|$. It follows that

$$\begin{aligned} \alpha &= \frac{|C(U, \bar{U})|}{|U|} \\ &\geq \frac{|C(U, \bar{U})|}{|\bar{U}|} \\ &\geq \frac{\sum_{i=1}^{k} |C(U_i, \bar{U}_i)|}{\sum_{i=1}^{k} |U_i|} \end{aligned}$$

So there exists i such that:

$$\frac{|C(U_i, \bar{U}_i)|}{|U_i|} \leq \alpha$$

Since $|U_i| \leq \frac{n}{2}$, it follows:

$$\alpha(U_i) \quad \leq \quad \alpha$$

Since α is the minimum among all sparsities, it follows that $\alpha(U_i) = \alpha$. Also, $G[U_i]$ and $G[V - U_i]$ are connected which contradicts our assumption that such a subset doesn't exist. Hence, this subcase isn't valid.
Case 2.2: There exists a component U_j of $V - U$ such that $|U_j| > \frac{n}{2}$.

$$\begin{aligned} \alpha &= \frac{|C(U, \bar{U})|}{|U|} \\ &= \frac{\sum_{i=1}^{k} |C(U_i, \bar{U_i})|}{|U|} \end{aligned} \tag{5}$$

Now consider the sparsity of $V - U_j$.

$$\begin{aligned} \alpha(V - U_j) &= \frac{|C(V - U_j, U_j)|}{|V - U_j|} \\ &= \frac{|C(U_j, \bar{U_j})|}{|U| + |U_1| + \cdots |U_{j-1}| + |U_{j+1}| \cdots |U_k|} \end{aligned} \tag{6}$$

From Equations (5) and (6), we observe that numerator of α is greater than or equal to the numerator of $\alpha(V - U_j)$ and the denominator of α is less than or equal to the numerator of $\alpha(V - U_j)$. So, $\alpha(V - U_j) \leq \alpha$. Moroever $G[U_j]$ and $G[V - U_j]$ are connected which contradicts our assumption that such a subset doesn't exist. Hence, this subcase isn't valid.

Hence, in all cases we obtained a contradiction. Therefore, there exists a subset S such that $G[S]$ and $G[V - S]$ are connected and $\alpha(S) = \alpha$. This concludes the proof. ∎