There are at least 3 methods to explore a graph:

1) DFS (1970's)
2) BFS 1980, 2000, 2010's (FRT Trees)
3) Random Walks
Low Diameter Decomposition

Applications of Low Dia. Decomp

1) Spanners (Distance Preserving Sparse Graphs)
2) Hop Set (Added edges to decrease number of edges needed in shortest paths)
3) Low Stretch Spanning Tree (LSST) (preserve distances on average)

Applications of LSST

Fast Alg for
1) Linear Solvers
   2) Max Flow
   3) Image Processing

4) Metric embeddings
Low Diameter Decomposition

Let $G = (V, E)$ undirected (unweighted)

$d(v) \equiv \text{degree of } v \in V$

$\text{Vol}(W) = \sum_{v \in W} d(v) \quad W \subseteq V.$

Note $\text{Vol}(V) = 2m$

Def $\text{Boundary}(W) \equiv$

$\partial W = \{(x, y) \mid x \in W, y \notin W, (x, y) \in E\}$

Def $\text{Isoperimetric Number}(W) \equiv$

$\Phi(W) = \frac{|\partial W|}{\text{Vol}(W)}$

Prob: Given $G = (V, E)$, $x \in V$, $0 < \beta < 1$

Find $x \in W \subseteq V$ of nearby points

$s.t. \Phi(W) \leq \beta.$

We use BFS.
**Ball Growing**

\[ \text{Def: } B(x,r) = \{ y \in V \mid \text{dist}(x,y) \leq r \} \]

(Ball of radius \( r \) centered at \( x \))

**Procedure:** \text{Grow Ball}(G,x,\beta)

1. Set \( r = 1 \)
2. While \( E(B(x,r)) \geq \beta \) set \( r = r + 1 \)
3. Return \( B_r = B(x,r), r = r \)

**Claim:** \( R = O(\log n / \beta) \)

**Note:** If \( r < R \) then \( |\partial B_r| \geq \beta \text{Vol}(B_r) \)

**Def:** \( \overline{\partial(B_r)} = \{ y \mid (x,y) \in E, x \in B_r, y \notin B_r \} \)

**Note:** \( \text{Vol}(\overline{\partial(B_r)}) \geq \beta \text{Vol}(B_r) \) (why?)

Thus \( \text{Vol}(B_{r+1}) \geq (1+\beta) \text{Vol}(B_r) \)

These are new edges
\[(1 + \beta)^r \leq \text{Vol}(B_r) \leq 2m\]

Taking \(\log\):

\[r \log(1 + \beta) \leq \log 2m\]

\[\gamma = x - 1\]

\[\gamma = \log_2 x\]

\[\log_2 x \geq x - 1\] for \(1 \leq x \leq 2\)

\[\log_2 (x+1) \geq x\] for \(0 \leq x \leq 1\)

Thus:

\[\log_2 (1 + \beta) \geq \beta\]

for \(0 \leq \beta \leq 1\)

\[r \cdot \beta \leq \log_2 m + 1 \Rightarrow r \leq \frac{\log_2 m + 1}{\beta}\]
We now use GrowBall to get a partition of $V$.

Procedure: BallDecomp($G, \beta$)

1. While $V \neq \emptyset$
2. Pick $x \in V$
3. $B_x = \text{GrowBall}(G, x, \beta)$
4. Remove $B_x$ & $dB_x$ from $G$

Return Balls.

Note: $\text{dist}_G(v, w) \ll \text{dist}_B(v, w)$ for $v, w \in B \equiv \text{Ball}$

![Diagram of a graph with three labeled regions $B_1$, $B_2$, and a ball $B$ showing the partition process]
Ball Growing Using Exponential Delay

Procedure: \( \text{Exp Delay}(G, \beta) \)

1) Each vertex \( V \in V \) draws \( X_v \sim \text{Exp}(\beta) \)
2) Each \( V \in V \) computes \( S_v = X_{\text{max}} - X_v \)
3) Each \( V \in V \) starts a BFS at time \( S_v \)
   a) If \( V \) is not owned at time \( S_v \) then \( V \) owns \( V \).
   b) Each \( V \) is owned by first arrival vertex.

\[ \text{Def } u \in \text{cluster}(V) \text{ if} \]

1) \( V = \arg\min_{W} \{ \text{dist}(w, u) + S_w \} \)
   or equivalently:
2) \( V = \arg\max_{W} \{ X_w - \text{dist}(w, u) \} \)
What is maximum cluster radius? §

Note: At time $X_{\text{max}}$ all nodes are owned!

Thus max cluster radius $\leq X_{\text{max}}$

$$X_{\text{max}} \leq \frac{2 \ln n}{\beta}$$

with prob $\geq 1 - \frac{1}{n}$.

Question: What is prob an edge is inter-cluster?

i.e. prob an edge is cut!

Let $e = (u,v)$ be some edge and $c$ its midpoint.

We think of each vertex doing a BES of $G$,

starting at time $S_{vi} = S_i$

the arrival time at $c$ will be a

random variable

$$T_i = X_{\text{max}} - X_i + \text{dist}(V_i, c)$$

Def: $\overline{T}_i = X_{\text{max}} - T_i = X_i - \text{dist}(V_i, c)$

(early arrival) (Owner has max $\overline{T}_i$)
Horse Race & Photo Finish.

We think of $\text{dist}(V_i, C)$ as handy-cap and $x_i$ speed.

\[ \text{Def } \overline{T}_i = \text{ith order statistics of } \{\overline{T}_{(1)}, \ldots, \overline{T}_{(n)}\} \]

Note: $\overline{T}_i = \overline{T}_{(n)}$ then $V_i$ "own" C and either U or V.

\[ \text{Def } \text{GAP}_i = \overline{T}_{(n)} - \overline{T}_{(n-1)} \]

Note: If $\text{GAP}_i > 1$ then $V_i$ will own U & V.

Thus $E = (u, v)$ is not "cut".
By memory less property $\text{Gap}_i \sim \text{Exp}(\beta)$

Thus $\Pr[\text{Gap}_i < 1] = 1 - e^{-\beta}$

Claim: $1 - e^{-\beta} < \beta$

$e^{-\beta} = 1 - \beta + \beta^2/2! - \beta^3/3!$

$1 - e^{-\beta} = \beta - \beta^2/2! + \cdots < \beta$ (Taylor's Thm)

**Thm** All edges $e$ prob $e$ is cut or inter cluster $< \beta$.

**Note:** True for all edges!

Some edges may have even lower prob.
Exponential Delay

Thm ExpDelay generates a clustering

1) Max Radius ≤ \( \ln n / \beta \) expected.

2) Max radius ≤ 2 ln n / \( \beta \) with prob 1 - \( 1/n \).

3) Expected number of inter-cluster edges
   is at most \( \beta m \).

4) Run time is \( O(m+n) \)

5) (Strong Diameter Prop)

   If \( w \in \text{Cluster}_v \) then shortest
   path from \( v \) to \( w \) is in \( \text{Cluster}_v \).

There still may be a large number of
inter-cluster edges!

eg \( m = n^2 \) & \( \beta = 1/10 \) ⇒ \( n^{2/10} \) edge!

Will show that they are nice!
Question: How many clusters will a 12 vertex see (share an edge with)?

1) It will belong to one cluster.
2) How many edges to distinct clusters.

Back to horse racing.

Consider early arrivals to $V$.

Must arrive within 2 units to possibly own a neighbor of $V$.

eg $V$ sees 4 clusters
Possible Neighbor Clusters to \( V \).

We prove a more general thm:

Suppose \( B \) is a ball of \( G \) with
1) center \( V \).
2) diameter \( d \).

Consider random variable
\[
C_B = \text{Cluster}(B) = \left| \{ \text{cluster} \mid \text{cluster} \cap B \neq \emptyset \} \right|
\]
Theorem \( \text{Exp} [C_B] \leq e^{d\beta} \)

\[ A_B = \text{number of arrivals within d time of first} \]

Note \( C_B \leq A_B \)

Claim: \( \text{Prob} \left[ A_B \geq t \right] = \left(1 - e^{-\beta} \right)^{(t-1)} \)

Proof of claim.

Consider time of \( t \)th early arrival

\[ N \sim T_N \]

Where \( N \) is the number of arrivals before time \( t \).

Then, at time \( T_t \) there are \( t-1 \) memoryless exponentials still alive.

\[ \text{But prob of this is } (1-e^{-\beta})^{(t-1)} \]

\[ \square \]
\[ E[A_B] = \sum_{t=0}^{\infty} \text{Prob}[A_B \geq t] = \sum_{t=1}^{\infty} (1-e^{-d \beta})^{t-1} \]

\[ = \frac{1}{1-(1-e^{-d \beta})} = e^{d \beta} \]

QED

We use fact that \( \sum_{i=0}^{\infty} \alpha^i = \frac{1}{1-\alpha} \)