Lecture 16 (Part II: Graph Streaming)

\[ G = (V, E) \text{ or weighted } G = (V, E, W) \]

Streaming models:

- **insertion only**: token inserts edges \( e \in V \times V \)
  - \( |V| = n \) is known in advance
  - \# edges is not known
  - promise: edges appear only once in stream

- **insertion for weighted graphs**:
  - token is \( (e, w) \), \( e \in E(G), w \text{ weight of } e \)

- **dynamic stream**:
  - token is \( (e, b) \in E(G) \times \{\pm 1, -1\} \)
    - \((e, +1) \Rightarrow \text{edge } e \text{ is inserted}\)
    - \((e, -1) \Rightarrow \text{edge } e \text{ is deleted}\)

- **turnstile** \( (e, m) \in E(G) \times \mathbb{Z} \)
  - \( m \) is the multiplicity of edge.
Connectivity

- Maintain a spanning forest greedily
  - Use union-find data structure to test if new edge makes a cycle with currently saved edges.

Space: $O(n \log n) = o(n^2)$ 2 bits

Bipartiteness

- Check if odd cycle.  

$O(n \log n)$ space.
Shortest path: query \((x, y) \forall x, y \in G\)

\(t\)-Spanner \(H\) for \(G\) is subgraph s.t.
\[
d(x, y) \leq d(x, y) \leq t \cdot d(x, y)
\]

**Correctness:** \(H\) is spanner.

\(H \subseteq \phi\)

- On edge \((u, v)\)
  \[
  \text{if } d_H(u, v) \geq t + 1
  \]
  \[
  H \leftarrow H \cup \exists (u, v) \in G
  \]

- On query \((x, y)\) output \(d_H(x, y)\).

Each edge \(e \in G\) is stretched by path of length \(t + 1\)
How large is the stored $H$?

Girth of $G$ is length of smallest cycle

Obs: Girth of $H$ above is $\geq t+2$

Alon, Hoory, Linial:

Any graph $G$ on $n$ vertices, of girth $> k$ must have $\leq n + n^{0(1/k)}$ edges.

For $t$-approx: $O(n \log n)$ space.