Algorithms for Streaming Tournaments

Sahil Kuchlous

Mentor: Prantar Ghosh

REU 2023, DIMACS
Supported by NSF grant CNS-2150186

The Streaming Model

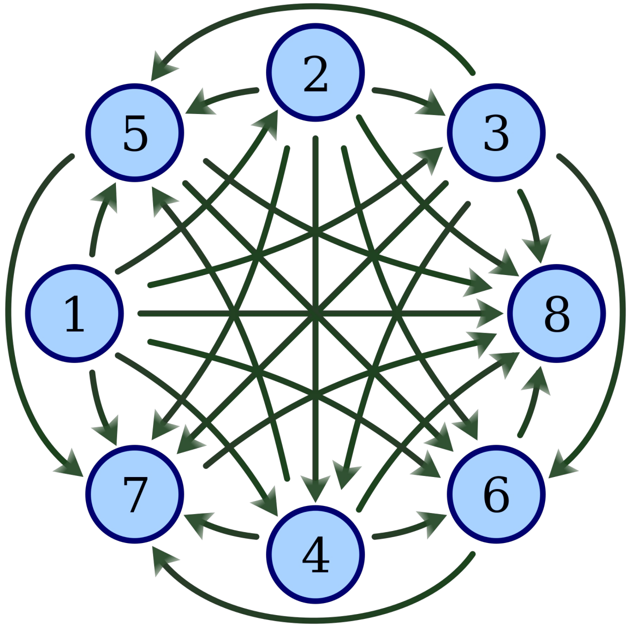
- Motivation: What if you have too much data to store in RAM?
- Sequential access to data; available memory: sublinear in the input size n
  - Single-pass is ideal, but $O(1)$ passes or more is also interesting
  - We don't care much about time
- Simple example: Mean in $O(\log n)$ space
- Harder example: # of distinct elements: $(1+\epsilon)$ approx. in $O(\epsilon^{-2} + \log n)$ space [KNW10]

Graph Streaming

- Motivation: Many graphs are too large to store (e.g., the internet, social networks)
- Input: Stream of edges in an n-vertex graph
  Goal: Answer questions about the graph without storing all edges
  - Most commonly $O(n \operatorname{polylog} n)$ memory (semi-streaming); $o(n^2)$ is still interesting
  - More complex models: edge deletions, sliding window
- Simple example: S-T connectivity in undirected graphs – $O(n \log n)$ space
- Harder example: Approximate min-cut – $(1+\epsilon)$ approx. in $O(n\epsilon^{-2})$ space [AG09]
- A good survey is [McG14]

Tournaments

- Digraphs in which all pairs of vertices have exactly 1 edge between them
  - Imagine a directed clique, or a competition where all pairs of players compete once
- Many streaming digraph problems are much easier on tournaments than on general digraphs [CGMV20]
  - SOTA: upper bounds on tournaments and lower bounds on general digraphs



Source: Wikepedia

Feedback Arc Set (FAS)

- Problem: What ordering of vertices minimizes number of back-edges?
- Gap between general digraphs and tournaments

FAS	Tournaments	Digraphs
Upper Bound	$(1+\epsilon)$ approx.: $\tilde{O}(n)$ space in 1 pass (exp. time) [CGMV20] $\tilde{O}(n^{1+1/p})$ space in p passes (poly. time) [BJW21]	**?**
Lower Bound	**?**	Any multiplicative approx.: $\Omega(n^2)$ space in 1 pass $\Omega(n^{1+\Omega(1/p)})$ space in p passes [CGMV20]

Our Main Goal

Our Project

- Goal: Investigate streaming algorithms on tournaments (upper and lower bounds)
- Problems we are interested in investigating:
  - Feedback arc set (FAS)
  - S-T connectivity – Given vertices s and t, is there a path from s to t?
  - Sink finding – Given a tournament with a sink, find one
  - …

Citations

- [AG09] Kook Jin Ahn and Sudipto Guha. Graph Sparsification in the Semi-streaming Model. In *International Colloquium on Automata, Languages, and Programming*, 2009.
- [BJW21] Anubhav Baweja, Justin Jia, and David P. Woodruff. An Efficient Semi-Streaming PTAS for Tournament Feedback Arc Set with Few Passes. *Innovations in Theoretical Computer Science*, 2022.
- [CGMV20] Amit Chakrabarti, Prantar Ghosh, Andrew McGregor, and Sofya Vorotnikova. Vertex Ordering Problems in Directed Graph Streams. In *Proc. 2020 ACM-SIAM Symposium on Discrete Algorithms (SODA),* 2020.
- [KNW10] Daniel M. Kane, Jelani Nelson, and David P. Woodruff. An optimal algorithm for the distinct elements problem. In Proc. 29th ACM Symposium on Principles of Database Systems, pages 41–52, 2010.
- [McG14] Andrew McGregor. Graph stream algorithms: a survey. In *SIGMOD Rec.,* 43(1), pages 9–20, 2014.