Algorithms for Streaming Tournaments

Sahil Kuchlous
Mentor: Prantar Ghosh

REU 2023, DIMACS
Supported by NSF grant CNS-2150186
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 $O(n \text{ polylog } n)$ memory (semi-streaming); $o(n^2)$ is still interesting
  • 1 pass is ideal; can make a few more

• 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

Feedback Arc Set (FAS)

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

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

Our Main Goal
### Feedback Arc Set (FAS)

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

<table>
<thead>
<tr>
<th>FAS</th>
<th>Tournaments</th>
<th>Digraphs</th>
</tr>
</thead>
</table>
| Upper Bound | $(1 + \epsilon)$ approx.:  
$\mathcal{O}(n)$ space in 1 pass (exp. time) [CGMV20]  
$\mathcal{O}(n^{1+1/p})$ space in $p$ passes (poly. time) [BJW21] | ?                             |
| Lower Bound | Exact and $(1 + \epsilon)$ approx. for small $\epsilon$:  
$\Omega(n^2)$ space in 1 pass! | Any multiplicative approx.:  
$\Omega(n^2)$ space in 1 pass  
$\Omega(n^{1+\Omega(1/p)})$ space in $p$ passes [CGMV20] |

First $\Omega(n^2)$ lower bound on tournaments
Proof

Given a stream of edges in a tournament, followed by a pair of vertices $u$ and $v$, determining the orientation of the edge between $u$ and $v$ is hard.

If we can use the minimum FAS ordering to determine the orientation of the edge, it must also be hard.
Proof

What if we force $u$ and $v$ to be before $T$ in the optimal FAS ordering?

By adding a new large tournament $A$, we force the optimal FAS ordering to place $u$ and $v$ before $A$ and $T$ after $A$.

Now, if the ordering puts $u$ before $v$, we know the edge must be oriented from $u$ to $v$, and vice versa.

Thus, we can determine the orientation of the edge, so minimum FAS ordering must also be hard.

Proof formalized using communication complexity.
Results

• Single-pass lower bounds on tournaments:
  • $\Omega(n^2)$ space for exact FAS ordering (and $(1 + \epsilon)$ approx. for small $\epsilon$)
  • $\Omega(n^2)$ space for exact FAS size (and $(1 + \epsilon)$ approx. for small $\epsilon$)
  • $\Omega(n^2)$ space for s-t distance

• Multi-pass lower bounds on tournaments ($p$ passes):
  • $\Omega(n/p)$ space for s-t connectivity
  • $\Omega(n/p)$ space for acyclicity testing (simplified proof from [CGMV20])

• Algorithms on tournaments ($p$ passes):
  • $\tilde{O}(n/p)$ space algorithm for acyclicity testing (matches lower bound)
  • $\tilde{O}(n^{1+1/p})$ space algorithm for s-t connectivity

• Lower bounds on general digraphs ($p$ passes):
  • $\Omega(n/p)$ space for finding a sink in a general directed graph (matches upper bound)
Next Steps

• Better FAS upper bounds:
  • We don’t know any polynomial-time algorithms for FAS ordering or size that get better than a 5-approximation in a single pass

• FAS approximation lower bounds:
  • Can we rule out better space-approximation trade-offs?

• Deterministic FAS:
  • Can we get equally good algorithms or stronger lower bounds?

• Algorithms for s-t distance on tournaments

• Lower bounds for s-t connectivity on tournaments

• More interesting models (insertion-deletion, random order streams)
Acknowledgements

• Dr. Prantar Ghosh, for being an amazing mentor.
• Prof. Gallos, Caleb and the rest of the REU team, for organizing this program.
• The NSF, for sponsoring REU programs like these.
• Prof. Sudan, Prof. Hesterberg and my other mentors at Harvard, for encouraging me to try research this summer.
• My family and friends, some of whom are watching right now.
• All of the REU students, for their support and encouragement, and for making this summer fun.
Citations

