Matching in Data Streams

Morteza Monemizadeh
Charles University, Prague

3/2/2016 at 11:00 am
Core A (Room 301)

Abstract

As noted by Lovasz and Plummer in their classic book:

"Matching Theory is a central part of graph theory, not only because of its applications, but also because it is the source of important ideas developed during the rapid growth of combinatorics during the last several decades."

We consider variants of matching in data streams when the insertion and the deletion of edges are revealed in a streaming fashion. In particular,

When the input graph is planar, we present a simple and elegant streaming algorithm that with high probability estimates the size of a maximum matching within a constant factor using $O(n^{2/3})$ space, where $n$ is the number of vertices. The approach generalizes to the family of graphs that have bounded arboricity, which include graphs with an excluded constant-size minor.

We further reduce the required memory size to $O(n)$ for three restricted settings: (i) when the input graph is a forest; (ii) when we have 2-passes and the input graph has bounded arboricity; and (iii) when the edges arrive in random order and the input graph has bounded arboricity.

We present a simple but powerful subgraph sampling primitive that is applicable in a variety of computational models including dynamic graph streams (where the input graph is defined by a sequence of edge/hyperedge insertions and deletions) and distributed systems such as MapReduce. In the case of dynamic graph streams, we use this primitive to prove that there exists an $O(k^2)$ space algorithm that returns the edges of a maximum matching on the assumption the cardinality is at most $k$. The best previous algorithm used $O(kn)$ space we prove our result is optimal up to logarithmic factors. Our algorithm has $O(1)$ update time.

We also show that there exists an $O(n^{2/3})$ space algorithm that returns an $\alpha$-approximation for matchings of arbitrary size. In independent work, Assadi et al.~(SODA 2016) proved this approximation algorithm is optimal and provided an alternative algorithm. We generalize our exact and
approximate algorithms to weighted matching. For graphs with low arboricity such as planar graphs, the space required for constant approximation can be further reduced. While there has been a substantial amount of work on approximate matching in insert-only graph streams, these are the first non-trivial results in the dynamic setting.

Organizer(s): Eric Allender, Pranjal Awasthi, Michael Saks and Mario Szegedy