Simpler Compression Based Algorithms for the Nonbipartite Matching Problem Using Superconcentrators

Martin Löhnertz

Institute of Computer Science V
University of Bonn

26.1.2005

Overview

1. Introduction
2. Graph Compression
3. Superconcentrators
4. Main Result

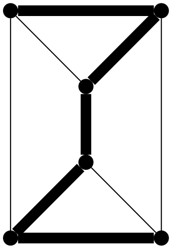
1 Introduction
Matching Problem
Matching Algorithms

2 Graph Compression

3 Superconcentrators

4 Main Result

Basic Concepts



- Graph $G = (V, E)$
- Matching $M \subset E$: Set of independent edges
- Augmenting path: Alternating path that starts and ends with a free vertex
- Maximum Matching

Kuhn, Hall, Berge (?)	bipartite	$\mathcal{O}(nm)$
Edmonds (1965)	nonbipartite	$\mathcal{O}(n^4)$
Hopcroft, Karp (1972)	bipartite	$\mathcal{O}(\sqrt{n}m)$
Even, Kariv (1975)	nonbipartite	$\mathcal{O}(n^{2.5})$
Micali, Vazirani (1980,1990), Blum(1990,1999)	nonbipartite	$\mathcal{O}(\sqrt{n}m)$
Feder,Motwani (1990)	bipartite	$\mathcal{O}(\sqrt{n}m\frac{\log\frac{2n^2}{m}}{\log n})$
Goldberg,Karzanov (1995,2002), Jungnickel, Paeger(2001), Löhnertz (2001,2004)	nonbipartite	$\mathcal{O}(\sqrt{n}m\frac{\log\frac{2n^2}{m}}{\log n})$

1 Introduction

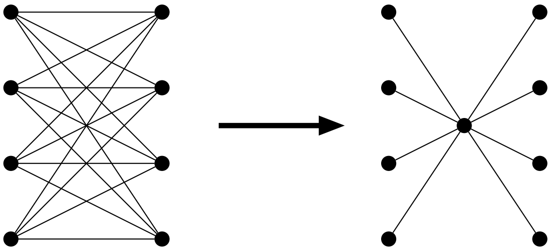
2 Graph Compression
The Bipartite Case
The skew-symmetric Case
Matching-Equivalent Subgraphs

3 Superconcentrators

4 Main Result

Graph Compression

- Replace homogeneous subgraphs by sparser structures
- Goal is not maximum compression but conservation of certain properties e.g. connectedness
- Replace cliques by stars until m is small.



Existence of a Large Clique

Theorem 1 (Feder,Motwani 1990)

A bipartite Graph $G = (A \dot\cup B, E)$ with $n^{2-\delta}$ edges contains a $(n^{1-\delta}, \frac{\delta \log n}{\log \frac{2n^2}{m}})$ clique.

Notation:

$k := \frac{\delta \log n}{\log \frac{2n^2}{m}}$

$n^{\underline{k}} := n * (n-1) * \cdots * (n-k+1)$

$d_i :=$ Degree of a_i

Proof of Theorem 1

- Idea: Show there is an ordered set of size k of vertices from B which have $n^{1-\delta}$ common neighbors in A.
- Number all of these sets from S_1 to $S_{n^{\underline{k}}}$
- Create $n \times n^{\underline{k}}$ matrix M with $M_{i,j} = 1$ if all vertices in set j are neighbors of a_i

$$a_1, \cdots, a_n \left\{ \overbrace{\begin{matrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \end{matrix}}^{S_1, \cdots, S_{n^{\underline{k}}}} \right.$$

$$a_1, \cdots, a_n \left\{ \overbrace{\begin{matrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \end{matrix}}^{S_1, \cdots, S_{n\underline{k}}} \right.$$

- $N_1(S_j) = \sum_{i=0}^{n} M_{ij}$
- $N_2(a_i) = \sum_j M_{ij} = d_i^{\underline{k}}$
- $C = \sum_{j=1}^{n\underline{k}} N_1(S_j) = \sum_{i=0}^{n} N_2(a_i)$
- $= \sum_{i=0}^{n} d_i^{\underline{k}}$

- Pigeonhole Principle: There is a set S_{j_0} with $N_1(S_{j_0}) \geq \frac{C}{n^{\underline{k}}}$
- C is minimal if all $d_i = \frac{m}{n}\ \forall i$
- $\frac{C}{n^{\underline{k}}} = \frac{\sum_{i=1}^{n} d_i^{\underline{k}}}{n^{\underline{k}}} \geq \frac{n*(\frac{m}{n})^{\underline{k}}}{n^{\underline{k}}}$
- $\geq \frac{n*(\frac{m}{n}-k)^k}{n^k}$ note: $k = \frac{\delta \log n}{\log \frac{2n^2}{m}} \leq \frac{m}{2n}$
- $\geq n^{1-\delta}$

Binary Search Tree for Vertices in V

- binary tree of depth $\log n$
- each leaf corresponds to a vertex in V
- edges labeled with 0 and 1
- V_a $a \in \{0,1\}^{\leq \log n}$ denotes all vertices which are descendants of vertex reached from root by following edges as described by a

Finding a Large Clique

- binary search for each element y_i of $S^* = (y_1, \cdots, y_k)$
- start with $w = \epsilon, V' = V \backslash \{y_1, \cdots y_{i-1}\}$
- calculate $N_2(\mathcal{S}_{i,w.0})$ and $N_2(\mathcal{S}_{i,w.1})$
  with $(\mathcal{S}_{i,a})\{S = (s_1, \cdots s_n)|$
  - $s_l = y_l \ \forall l < i$
  - $s_i \in V'_a$
  - $s_l \in V' \backslash \{s_i\} \ \forall l > i$

  $\}$
- append 0 or 1 to w depending on which one was larger
  if $|w| < \log n$ goto step 3
- let y_i be the only vertex in V_a
  $V' := V' \backslash y_i$

Making the Nonbipartite Problem Bipartite

- replace each vertex v by two vertices $[v, A]$ and $[v, B]$
- connect $[u, A]$ to $[v, B]$ iff $(u, v) \in E$
- new Graph G_B is bipartite
- matching M_B in G_B can be transformed to matching in M if $([u, A], [v, B]) \in M_B \Leftrightarrow ([v, A], [u, B]) \in M_B$
- standard maxflow/matching algorithms do not guarantee this
- Graphs are called skew symmetric as $[u, A] \leftrightarrow [u, B]$ is involutoric automorphism

Finding Skew-Symmetric Matchings

- general skew symmetric flow algorithms (Jungnickel & Paeger 2001,Goldberg & Karzanov 2002)
- specialized matching algorithm (Blum 1999)

Common idea: choose only augmenting paths not containing $[u, A]$ and $[u, B]$ for any u

Symmetric Compression

Definition 2

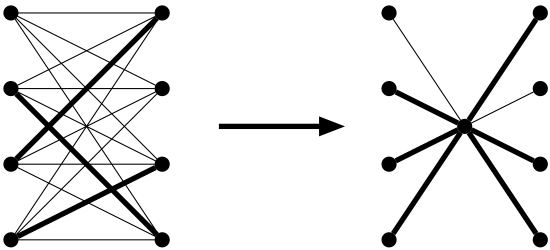
A clique decomposition is called “symmetric” if for each biclique C the symmetric biclique is also contained.

Lemma 3 (Goldberg, Karzanov 2002)

A symmetric decomposition can be found in the same time as a normal decomposition.

Stars Work for Flow Algorithms

- At most 1 unit of flow may pass each vertex that is not a center of a star.
- This cannot be transferred to matching algorithms.



CME-Graphs

Definition 4

A Graph $G_S = (V_L \dot\cup V_M \dot\cup V_R)$ is matching-equivalent to biclique (V_L, V_R) if

- for each matching in G_S covering V_M the number of matching covered vertices in V_L and V_R is equal
- for pair of sets $(V'_L \subset V_L, V'_R \subset V_R)$, $|V'_L| = |V'_R|$ there is a matching covering exactly $V_M \cup V'_R \cup V'_L$

i.e. one can replace a clique by a CME graph without disturbing the perfect matching properties of the rest of the graph

1 Introduction

2 Graph Compression

3 Superconcentrators
Definitions
Construction of Concentrators

4 Main Result

There are 'Magic' Graphs

Definition 5

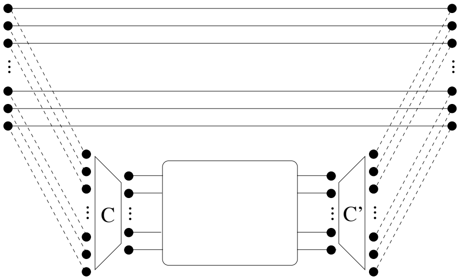
A m, n concentrator is a graph $(V_I \dot\cup V_M \dot\cup V_R, E), |V_I| = m, |V_R| = n$ where for every subset $A \subset V_I$ of size n there ar n node disjoint paths from A to V_R

Definition 6

A n, n-superconcentrator is a graph $G_{SU} = (V_I \dot\cup V_M \dot\cup V_O, E)$ in which for each pair of subsets $V_I' \subset V_I, V_O' \subset V_O$, $r = |V_I| = |V_O|$ there are r vertex disjoint paths from a vertex from V_I to a vertex from V_O

There are Linear Size Superconcentrators

- recursive construction
- sub-superconcentrators are smaller by a constant fraction



Construction of Concentrators

- Probabilistic Construction Pippenger (1972): $39n$
- Deterministic Construction: Gabber,Galil (1979): $504n$

$\rightarrow$ uses 7 simple permutations for the construction, but proof is measure-theoretic and applies residual-calculus

1 Introduction

2 Graph Compression

3 Superconcentrators

4 Main Result
CME Graphs from Superconcentrators
proof of the main theorem

Main Result

Lemma 7

There are linear sized CME-graphs.

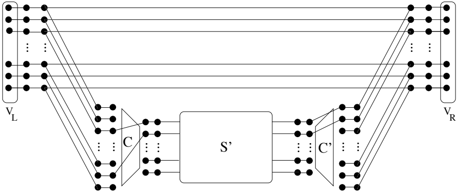
Theorem 8

In a general graph a maximum cardinality matching can be found in time

$$\mathcal{O}(\sqrt{n}m\frac{\log\frac{2n^2}{m}}{\log n})$$

Sketch of Proof of Lemma 7

- replace each vertex v in the superconcentrator by vertices v_i, v_o and name set of these nodes V_M
- v_i gets "incoming edges", v_o "outgoing" edges, add edges (v_i, v_o)
- Add set $V_L = \{v_1^L, \cdots v_n^L\}$ and connect each to the v_i of an input node
- Add set $V_R = \{v_1^R, \cdots v_n^R\}$ and connect each to the v_o of an output node



Sketch of Proof of Lemma 7

Lemma 9

The Construction yields an CME-Graph

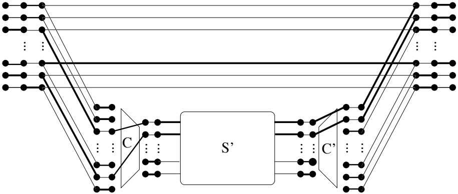
Property 1: for each matching in G_S covering V_M the number of matching covered vertices in V_L and V_R is equal

- Graph is bipartite with balanced partitions
- in such graphs the number of vertices covered in each partition are equal
- V_L and V_R belong to different partitions and are of same size
- V_M is evenly split and completely covered so in both partitions the number of covered vertices from V_M is equal.

Sketch of Proof of Lemma 9

Property 2: for each pair of sets $(V_L' \subset V_L, V_R' \subset V_R)$, $|V_L'| = |V_R'|$ there is a matching covering exactly $V_M \cup V_R' \cup V_L'$

- let $V_L' = \{v_1^L, \cdots, v_r^L\}$ and $V_R' = \{v_1^R, \cdots, v_r^R\}$
- There are r paths $P_1, \cdots P_r$ through the superconcentrator connecting $\{v_1^L, \cdots, v_r^L\}$ to $\{v_1^R, \cdots v_r^R\}$
- for each pair (v_i, v_o) not lying on a P_i match v_i to v_o
- for each $P_i = \{p_j^L, p_i^1, p_o^1, p_i^2, \cdots, p_o^{q-1}, p_i^q, p_o^q, p_h^R\}$ match p_o^j to p_i^{j+1} $\forall 1 \leq j < q$, p_j^L to p_i^1 and p_o^q to p_h^R.



Proof of the Main Theorem

Definition 10

Let G_C be the graph created by replacing each clique in a clique packing according to Feder and Motwani by an CME-subgraph.

Lemma 11

It is possible to reconstruct a maximum matching in the original graph G from a maximum matching in G_C in linear time.

Equal Deficiency

Lemma 12

The number of vertices not covered by a maximum matching M' in G_C is smaller or equal to the number of vertices not covered by a maximum matching M in G.

Proof: For each Clique there is a matching in the corresponding CME-Graph covering the whole ”interior” of this graph V_M and exactly the same number of vertices from its V_L and V_R. So there is a matching in G_C leaving at most as many vertices uncovered as M.

Reconstruction

Lemma 13

Given an CME-Graph G_S and a maximum Matching M' in this Graph it is possible to find a matching M'' of same size, that covers the same number of vertices in each V_L and V_R and does not cover other vertices in V_L or V_R than M'.

Proof: Let V_L^* and V_R^* be the vertices covered by M' and $|V_L^*| < |V_R^*|$. Choose any subset of size $|V_L^*|$ of V_R^* and create the matching M'' as the matching requested by property 2. Assume $|M'| > |M''|$. Look at $M' \Delta M''$. As M' was maximum there must be an augmenting path P^* starting and ending in V_R. Then $M'' \oplus P^*$ is a matching covering V_M and a different number of vertices in V_L and V_R. Contradiction.

Main Result

Theorem 14

In a general graph a maximum cardinality matching can be found in time

$$\mathcal{O}(\sqrt{n}m\frac{\log\frac{2n^2}{m}}{\log n})$$

5 Search Structures

Search Structures

- Compression based on deterministically created concentrators accelerates algorithm if $n > 2^{504}$
- Method does not work for nonsymmetric compression
- Use cliques only to accelerate searches and do other operations on original graph
- Broader concept: Search Structures, e.g. non disjoint cliques