Width Helps and Hinders Splitting Flows

Manuel Cáceres[1] Massimo Cairo[1] *Andreas Grigorjew*[1] Shahbaz Khan[2] Brendan Mumey[3] Romeo Rizzi[4] Alexandru Tomescu[1] Lucia Williams[3]

[1]University of Helsinki

[2]Indian Institute of Technology Roorkee

[3]Montana State University

[4]University of Verona

14 June, 2022

UNIVERSITY OF HELSINKI

Flow graphs

Given is an s-t DAG $G = (V, E)$ and a flow $X : E \to \mathbb{Y}$ (e.g., $\mathbb{Y} = \mathbb{N}$, or $\mathbb{Y} = \mathbb{Z}$).
Conservation of flow: $\sum_{e \in \delta^+(v)} X(e) = \sum_{e \in \delta^-(v)} X(e) \forall v \in V \setminus \{s, t\}$.

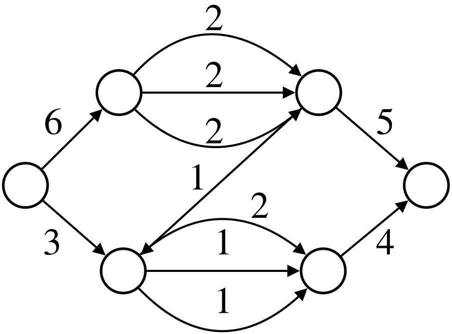


Figure: Simple flow graph.

Minimum Flow Decomposition (MFD$_{\mathbb{Y}}$)

Minimum Flow Decomposition (MFD$_{\mathbb{Y}}$) of (G, X): minimum sized set of s-t paths and weights $\{(P_1, w_1), \ldots, (P_k, w_k)\}$ ($w_i \in \mathbb{Y}$) with

$$X = \sum_{i=1}^{k} w_i P_i.$$

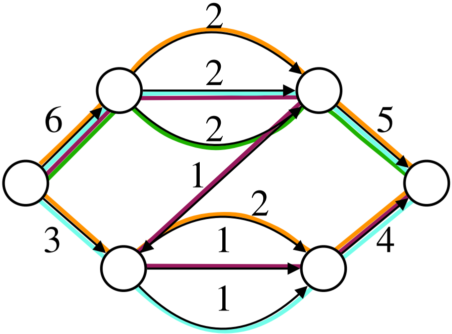


Figure: Flow decomposition in 6 s-t paths.

Minimum (Differential) Flow Decomposition

Use in biology

- $MFD_{\mathbb{N}}$: reconstructing sequences (RNA transcripts, viral strains), transportation planning
- $MFD_{\mathbb{Z}}$: so far unconsidered. Can explain the *difference* between two flows; e.g. differential expression in RNA-seq: minimize the number of up/down regulated genes
- $MFD_{\mathbb{Y}}$ is NP-hard. No good approximations known so far.

Results: Approximating MFD

- Approximating $\mathrm{MFD}_{\mathbb{Z}}$ using width: $O(\log \|X\|)$
- Width hinders greedy in $\mathrm{MFD}_{\mathbb{N}}$: $\Omega(\sqrt{m}) \rightarrow \Omega(m/\log m)$
- Width helps greedy in $\mathrm{MFD}_{\mathbb{N}}$: $O(\log |X|)$ on certain graphs

Results: Approximating MFD

- Approximating $\text{MFD}_{\mathbb{Z}}$ using width: $O(\log \|X\|)$
- Width hinders greedy in $\text{MFD}_{\mathbb{N}}$: $\Omega(\sqrt{m}) \to \Omega(m/\log m)$
- Width helps greedy in $\text{MFD}_{\mathbb{N}}$: $O(\log |X|)$ on certain graphs

Some notation.

- $||X|| = \max_{e \in E} |X(e)|$
- $|X| = \sum_{e \in \delta^+(s)} X(e)$ (total flow of X)

Results: Approximating MFD

- Approximating $\text{MFD}_{\mathbb{Z}}$ using width: $O(\log \|X\|)$
- Width hinders greedy in $\text{MFD}_{\mathbb{N}}$: $\Omega(\sqrt{m}) \to \Omega(m/\log m)$
- Width helps greedy in $\text{MFD}_{\mathbb{N}}$: $O(\log |X|)$ on certain graphs

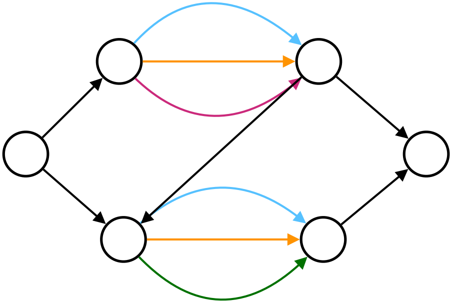


Figure: A graph of width 4, single path edges coloured

Results: Approximating MFD

- Approximating $\text{MFD}_{\mathbb{Z}}$ using width: $O(\log \|X\|)$
- Width hinders greedy in $\text{MFD}_{\mathbb{N}}$: $\Omega(\sqrt{m}) \rightarrow \Omega(m/\log m)$
- Width helps greedy in $\text{MFD}_{\mathbb{N}}$: $O(\log |X|)$ on certain graphs

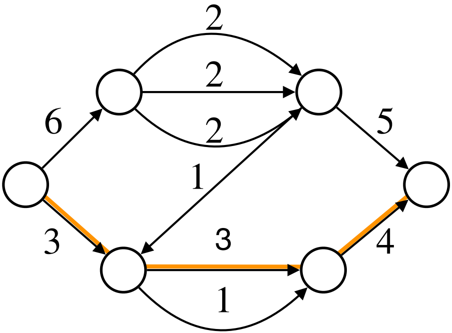


Figure: Orange path of weight 3 highest weight path (greedy-weight)

Approximating $\mathrm{MFD}_{\mathbb{Z}}$

Observation: Odd edges form (undirected) s-t paths and cycles

Power of two–approach

1. "Remove" the odd part: Given flow X, define flow Y with $\|Y\| \leq 1$ of the same parity as X
2. Decompose Y into at most width(G) paths
3. Divide the remaining flow by 2

Removing the odd part

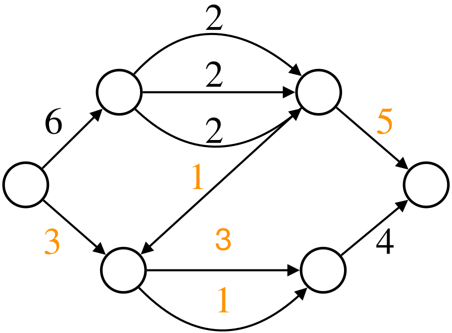


Figure: Example iteration.

Removing the odd part

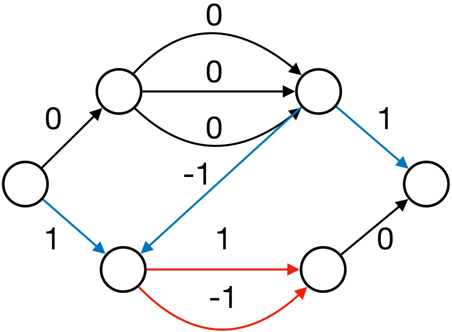


Figure: Example iteration.

Removing the odd part

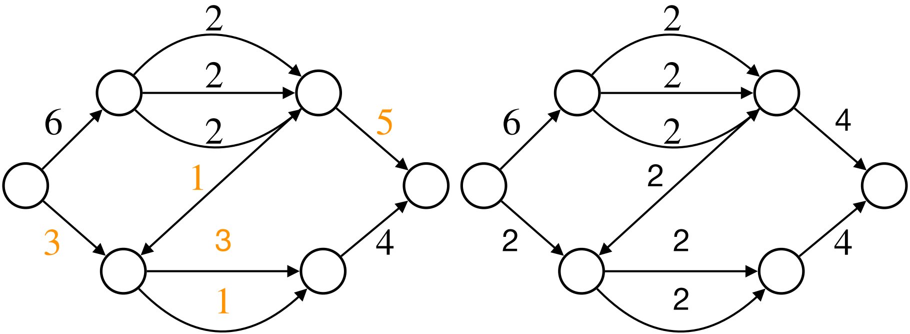


Figure: Example iteration. left: Initial graph, right: Odd part removed

Removing the odd part

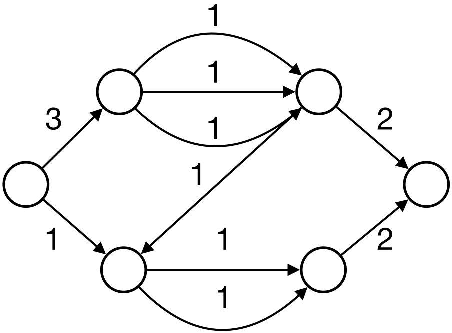


Figure: Example iteration.

Approximating $\mathrm{MFD}_{\mathbb{Z}}$

Corollary

For any flow $X : E \to \mathbb{Z}$ with $\|X\| \leq 1$, there exist paths $P_1, \dots, P_k$ with $k \leq \mathrm{width}(G)$ such that $X = P_1 + \dots + P_\ell - P_{\ell+1} - \dots - P_k$ (for some $0 \leq \ell \leq k$).

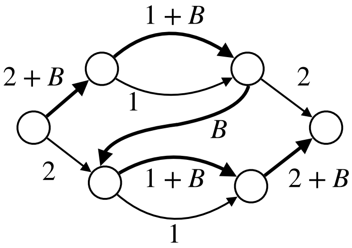
We can find flows $A \geq 0$ and $B \geq 0$ such that $|A| + |B| \leq \mathrm{width}(G)$ and $A - B = X$.

$O(\log \|X\|) \cdot \mathrm{width}(G)$-Approximation of $\mathrm{MFD}_{\mathbb{Z}}$

1. If $\|X\| > 1$, write $X = 2X' + X''$ ($\|X''\| \leq 1$), decompose X'' as above
2. Continue with $X = X'$
3. If $\|X\| \leq 1$, decompose X as above

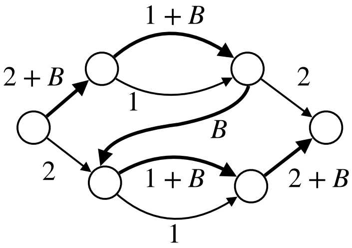
Width hinders greedy on $\mathrm{MFD}_{\mathbb{N}}$

Width hinders greedy on $MFD_{\mathbb{N}}$

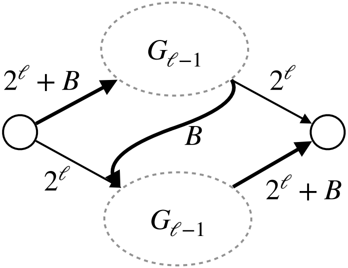


(a) The base case $(G_1, X_{1,B})$. Bold edges carry flow at least B.

Width hinders greedy on $MFD_{\mathbb{N}}$

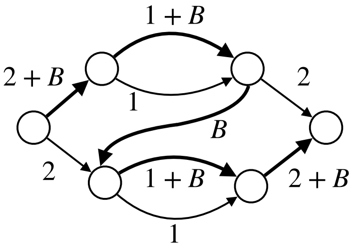


(a) The base case $(G_1, X_{1,B})$. Bold edges carry flow at least B.

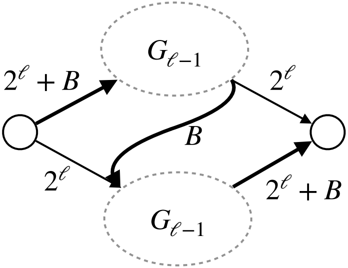


(b) Building $(G_\ell, X_{\ell,B})$ from two copies of $(G_{\ell-1}, X_{\ell-1,B})$ $(\ell > 1)$.

Width hinders greedy on $\text{MFD}_{\mathbb{N}}$



(a) The base case $(G_1, X_{1,B})$. Bold edges carry flow at least B.



(b) Building $(G_\ell, X_{\ell,B})$ from two copies of $(G_{\ell-1}, X_{\ell-1,B})$ $(\ell > 1)$.

$(G_\ell, X_{\ell,B})$ can be decomposed into $\Theta(\ell)$ paths. Greedy-weight uses $\Theta(2^\ell)$ paths.

$\rightarrow$ Approximation ratio for greedy-weight on $\text{MFD}_{\mathbb{N}}$ is $\Omega(m/\log m)$ for sparse graphs.

Width-stable graphs

Two properties

1. We say that G is *stable* if $\text{width}(G|_X) \leq \text{width}(G)$ for all non-negative flows X on G.
2. We say that G carries paths of large weight if $G|_X$ carries an s-t path of weight at least $|X|/\text{width}(G)$ for all non-negative flows X.

Many graphs in practice have these properties (e.g., series-parallel graphs).

Width-stable graphs

Two properties

1. We say that G is *stable* if $\text{width}(G|_X) \leq \text{width}(G)$ for all non-negative flows X on G.
2. We say that G carries paths of large weight if $G|_X$ carries an s-t path of weight at least $|X|/\text{width}(G)$ for all non-negative flows X.

Many graphs in practice have these properties (e.g., series-parallel graphs).

Lemma

Given both properties, greedy-weight uses at most $\left\lfloor \log |X| / \log \frac{\text{width}(G)}{\text{width}(G)-1} \right\rfloor + 1$ paths to decompose a non-negative flow X.

Proof: Using Property 2 c times induces a flow of $|X| \left(\frac{\text{width}(G)-1}{\text{width}(G)}\right)^c$.
Solve for c for the term to be smaller than 1.

Conclusions

- Width plays important role in $\mathrm{MFD}_{\mathbb{Y}}$
- Greedy approximations are worse than hoped
- But not necessarily in practice
- $\mathrm{MFD}_{\mathbb{Z}}$ as relaxation of $\mathrm{MFD}_{\mathbb{N}}$ gives theoretical hinsights

Thank you!

European Research Council

Established by the European Commission