A Hybrid 2D Method for Sparse Matrix Partitioning

Rob Bisseling, Tristan van Leeuwen

Utrecht University

Ümit Çatalyürek

Ohio State University

Support from BSIK-BRICKS/MSV and NCF

Universiteit Utrecht

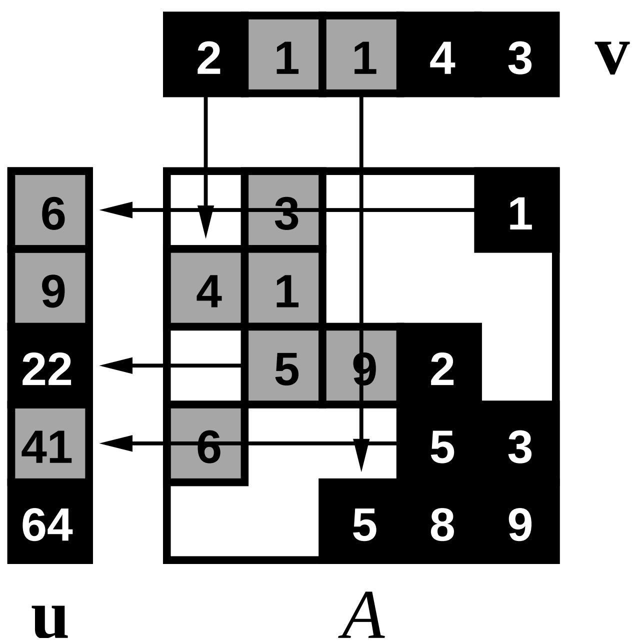
Outline

1. Introduction
   - Mondriaan 2D matrix partitioning
   - Fine-grain 2D partitioning
2. New: hybrid method for 2D partitioning
   - The difficulty of hybrids
   - Combining the Mondriaan and fine-grain methods
3. Experimental results
   - PageRank matrices: Stanford, Stanford-Berkeley
   - Other sparse matrices: term-by-document, linear programming, polymers
4. Conclusions and future work

Parallel sparse matrix–vector multiplication $\mathbf{u} := A\mathbf{v}$

A sparse $m \times n$ matrix, $\mathbf{u}$ dense m-vector, $\mathbf{v}$ dense n-vector

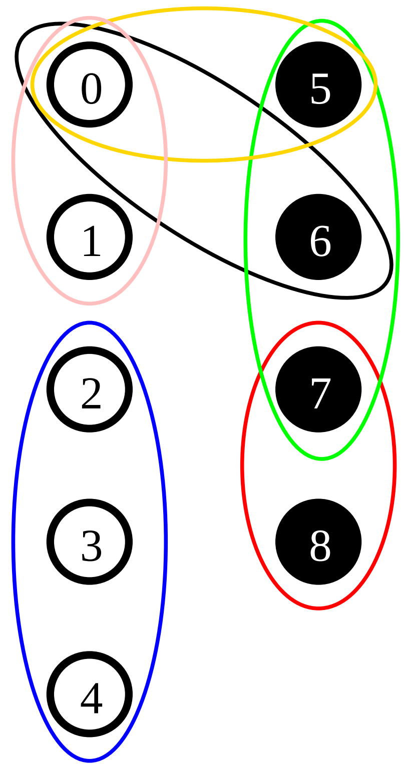
$$u_i := \sum_{j=0}^{n-1} a_{ij} v_j$$



4 phases: communicate, compute, communicate, compute

Universiteit Utrecht

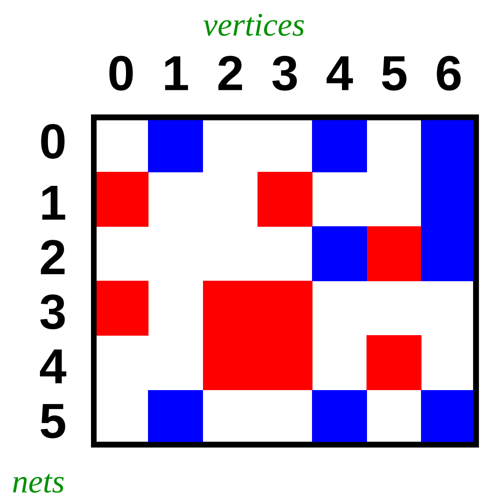
Hypergraph



Hypergraph with 9 vertices and 6 hyperedges (nets), partitioned over 2 processors

Universiteit Utrecht

1D matrix partitioning using hypergraphs

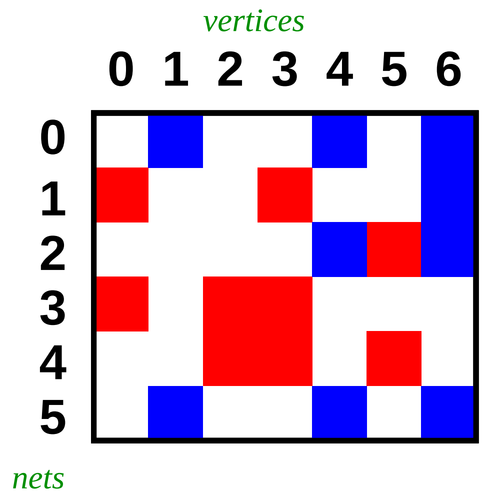


Column bipartitioning of $m \times n$ matrix

- Hypergraph $\mathcal{H} = (\mathcal{V}, \mathcal{N}) \Rightarrow$ exact communication volume in sparse matrix–vector multiplication.
- Columns $\equiv$ Vertices: $0, 1, 2, 3, 4, 5, 6$.
  Rows $\equiv$ Hyperedges (nets, subsets of $\mathcal{V}$):

$$n_0 = \{1,4,6\}, \quad n_1 = \{0,3,6\}, \quad n_2 = \{4,5,6\},$$
$$n_3 = \{0,2,3\}, \quad n_4 = \{2,3,5\}, \quad n_5 = \{1,4,6\}.$$

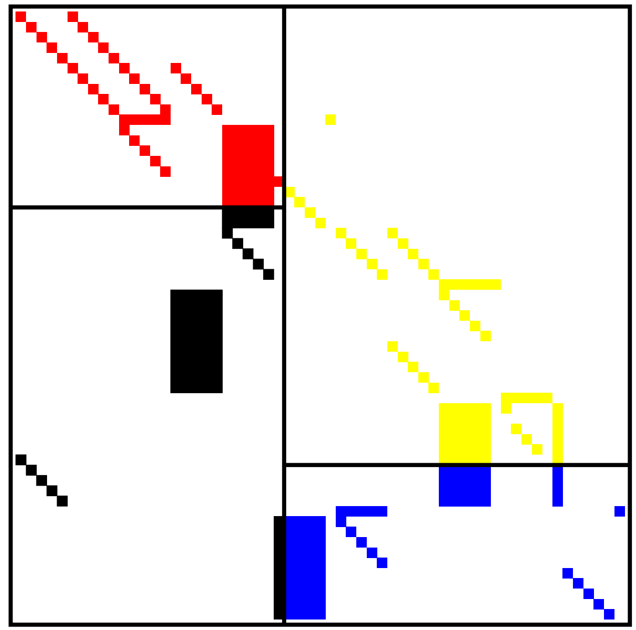
Minimising communication volume



- Broken nets: n_1, n_2 cause one horizontal communication
- Use Kernighan–Lin/Fiduccia–Mattheyses for hypergraph bipartitioning
- Multilevel scheme: merge similar columns first, refine bipartitioning afterwards
- Used in PaToH (Çatalyürek and Aykanat 1999) for 1D matrix partitioning.

Universiteit Utrecht

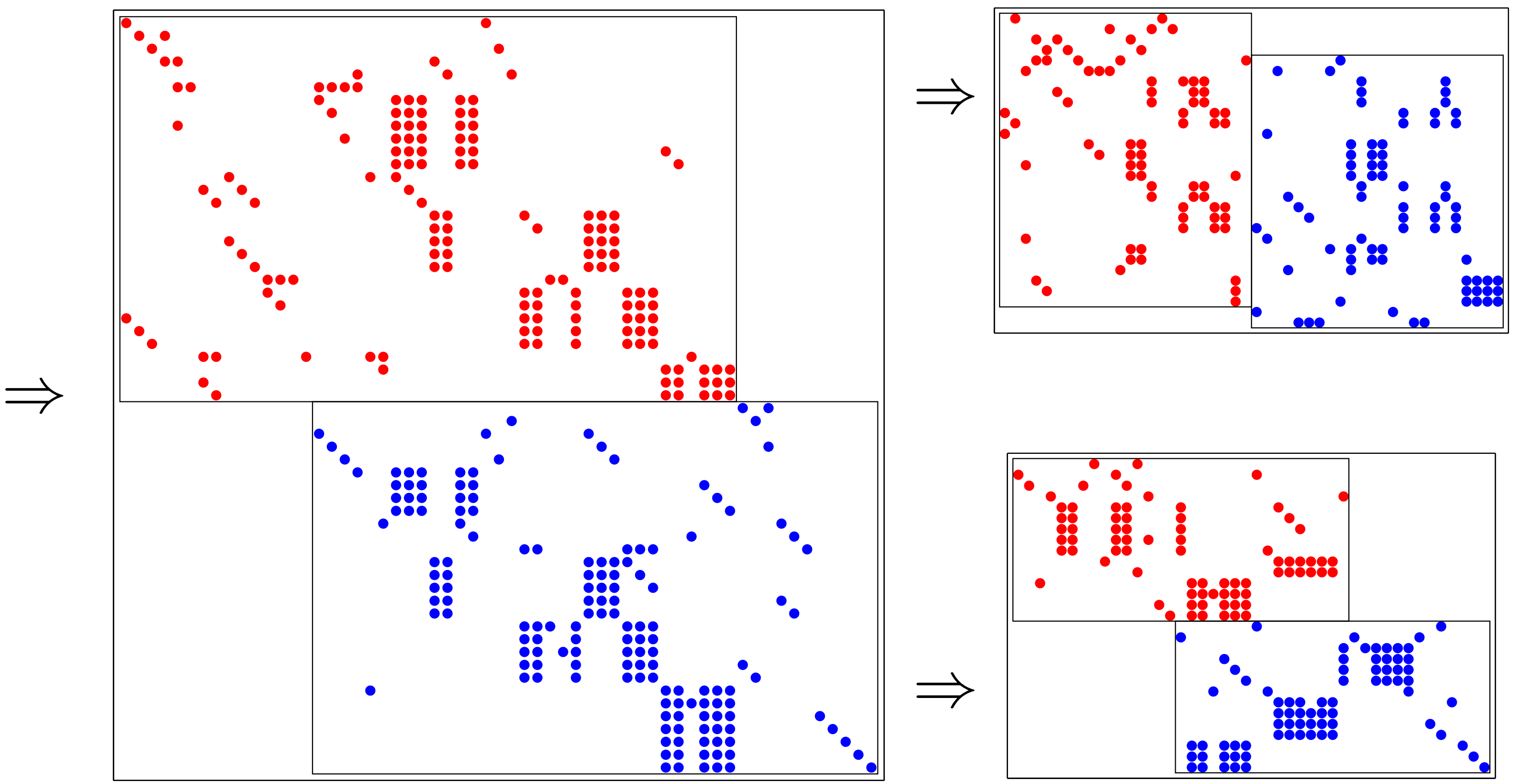
Mondriaan 2D matrix partitioning



- Block distribution (without row/column permutations) of 59×59 matrix `impcol_b` with 312 nonzeros, for $p = 4$
- Mondriaan package v1.0 (May 2002). Originally developed by Vastenhouw and Bisseling for partitioning term-by-document matrices for a parallel web search machine.

Universiteit Utrecht

Mondriaan 2D partitioning



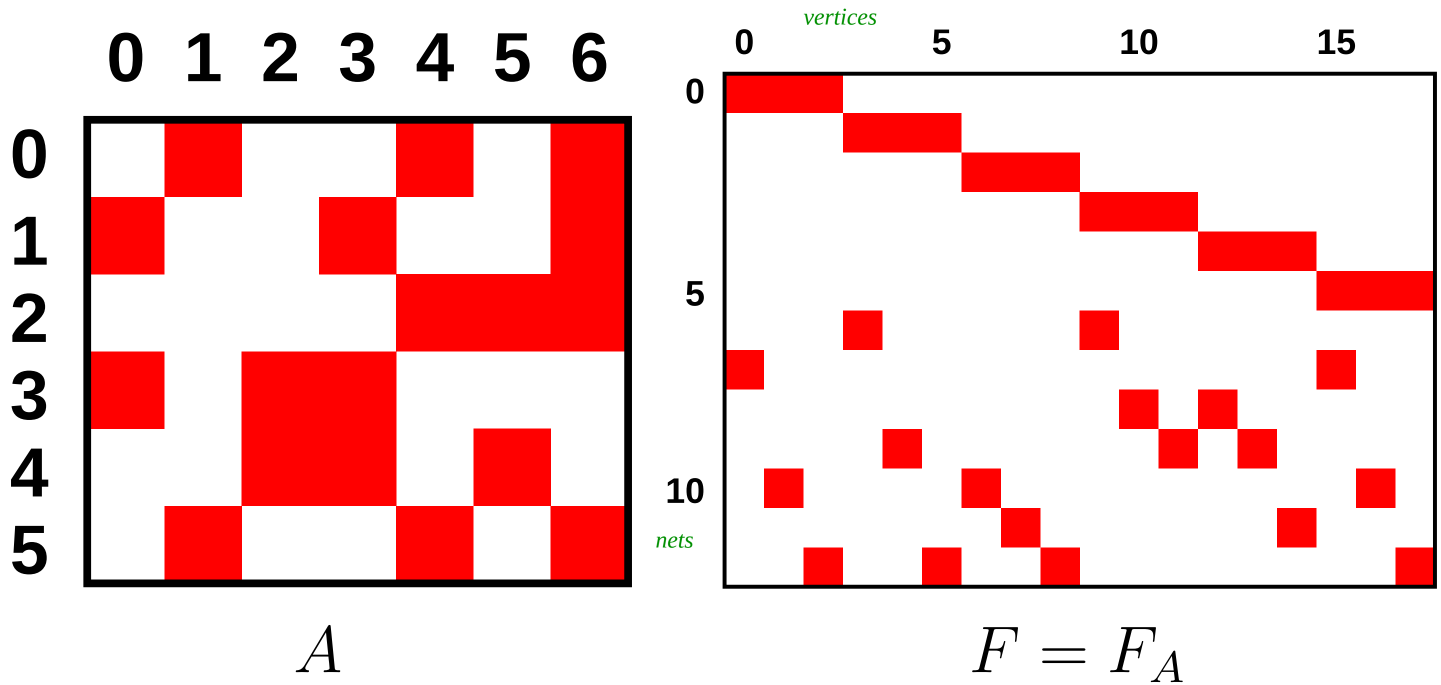
- Recursively split the matrix into 2 parts
- Try splits in row and column directions, allowing permutations. Each time, choose the best direction

Fine-grain 2D partitioning

- Assign each nonzero of A individually to a part.
- Each nonzero becomes a vertex; each matrix row and column a hyperedge.
- Hence $nz(A)$ vertices and $m + n$ hyperedges.
- Proposed by Çatalyürek and Aykanat, 2001.

Universiteit Utrecht

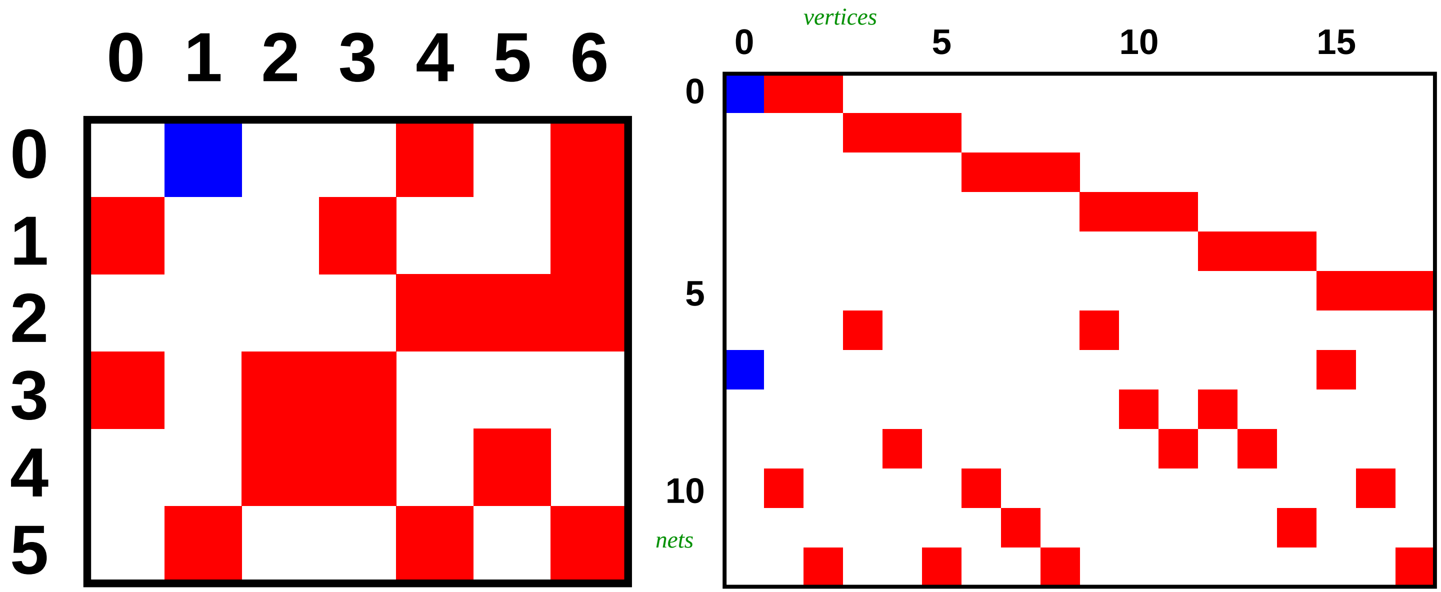
Matrix view of fine-grain 2D partitioning



- $m \times n$ matrix A with $nz(A)$ nonzeros
- $(m+n) \times nz(A)$ matrix $F = F_A$ with $2 \cdot nz(A)$ nonzeros
- a_{ij} is kth nonzero of A $\Leftrightarrow$ f_{ik}, $f_{m+j,k}$ are nonzero in F

Universiteit Utrecht

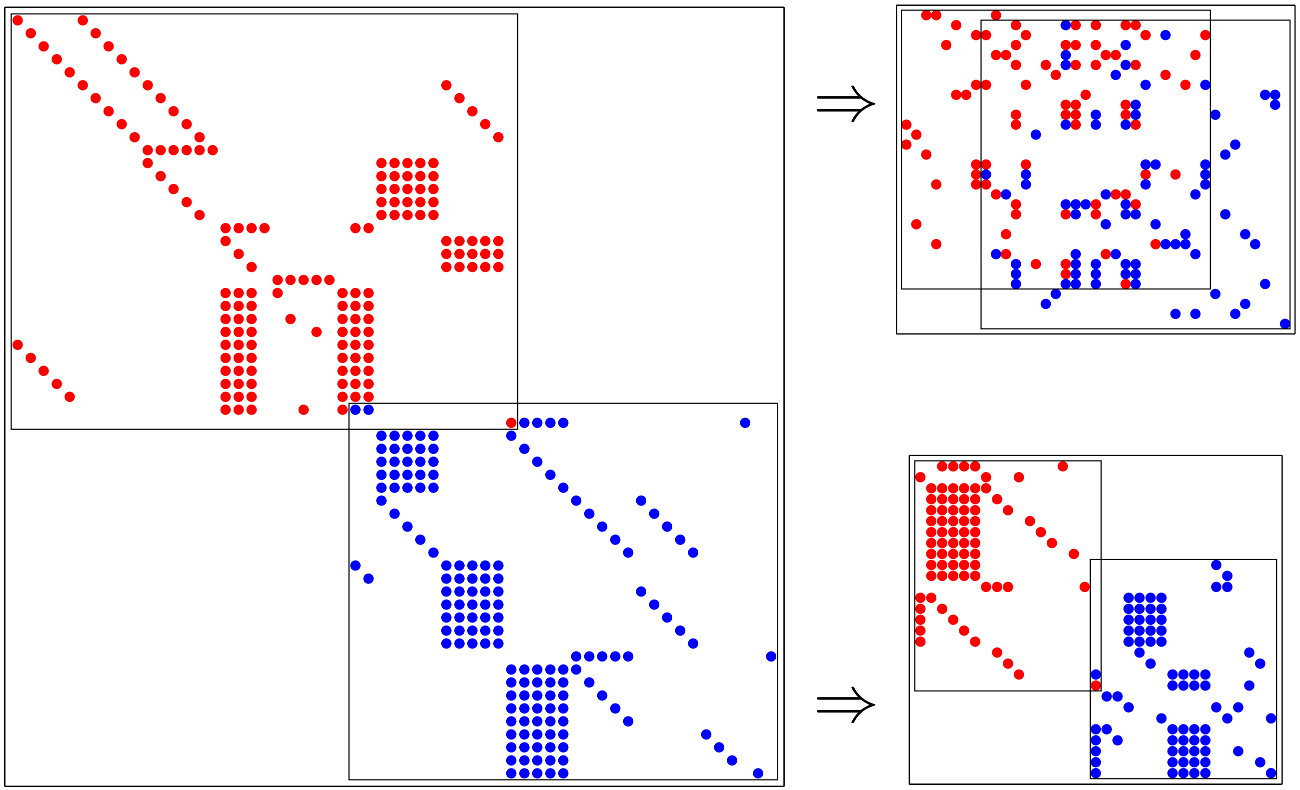
Communication for fine-grain 2D partitioning



A $\quad$ $F = F_A$

- Broken net in first m nets of hypergraph of F: nonzeros from row a_{i*} are in different parts, hence horizontal communication in A.
- Broken net in last n nets of hypergraph of F: vertical communication in A.

Fine-grain 2D partitioning



- Recursively split the matrix into 2 parts
- Assign individual nonzeros to parts

Universiteit Utrecht

The difficulty of hybrids — a story

The beautiful American dancer Isadora Duncan (1878–1927) suggested to the Irish writer George Bernard Shaw (1856–1950) that they should have a child together:

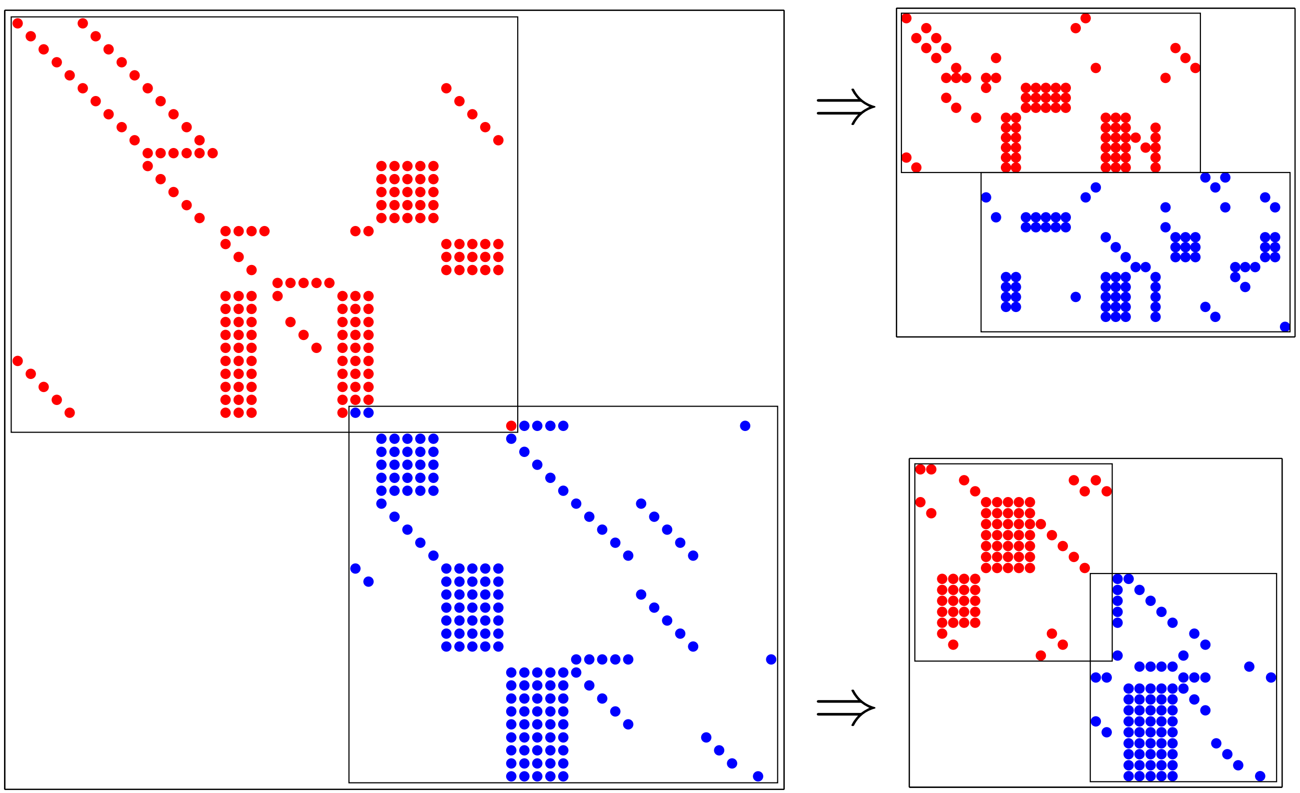
> “Think of it! With your brains and my body, what a wonder it would be."

Shaw’s reply:

> “Yes, but what if it had my body and your brains?"

Source:
http://www.chiasmus.com/mastersofchiasmus/shaw.shtml
Many different versions exist. Story may be apocryphal.

Hybrid 2D partitioning



- Recursively split the matrix into 2 parts
- Try splits in row and column directions, and fine-grain
- Each time, choose the best of 3

Recursive, adaptive bipartitioning algorithm

MatrixPartition(A, p, ϵ)

input: ϵ = allowed load imbalance, $\epsilon > 0$.

output: p-way partitioning of A with imbalance $\leq \epsilon$.

if $p > 1$ **then**

$q := \log_2 p$;

$(A_0^{\mathrm{r}}, A_1^{\mathrm{r}}) := h(A, \mathrm{row}, \epsilon/q)$; hypergraph splitting

$(A_0^{\mathrm{c}}, A_1^{\mathrm{c}}) := h(A, \mathrm{col}, \epsilon/q)$;

$(A_0^{\mathrm{f}}, A_1^{\mathrm{f}}) := h(A, \mathrm{fine}, \epsilon/q)$;

$(A_0, A_1) :=$ best of $(A_0^{\mathrm{r}}, A_1^{\mathrm{r}})$, $(A_0^{\mathrm{c}}, A_1^{\mathrm{c}})$, $(A_0^{\mathrm{f}}, A_1^{\mathrm{f}})$;

$maxnz := \frac{nz(A)}{p}(1+\epsilon)$;

$\epsilon_0 := \frac{maxnz}{nz(A_0)} \cdot \frac{p}{2} - 1$; MatrixPartition($A_0, p/2, \epsilon_0$);

$\epsilon_1 := \frac{maxnz}{nz(A_1)} \cdot \frac{p}{2} - 1$; MatrixPartition($A_1, p/2, \epsilon_1$);

else output A;

Universiteit Utrecht

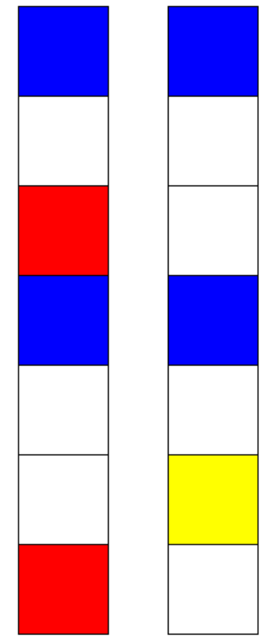
Similarity metric for column merging (coarsening)

Column-scaled inner product:

$$M(u,v) = \frac{1}{\omega_{uv}} \sum_{i=0}^{m-1} u_i v_i$$

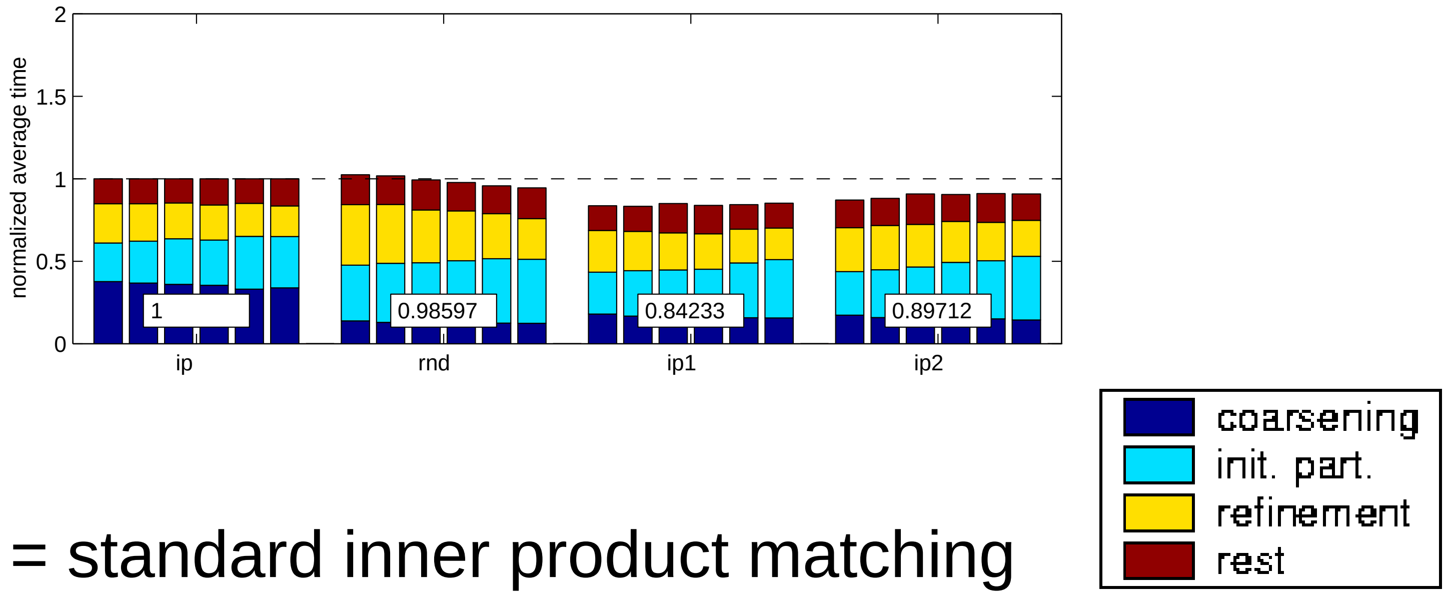
- $\omega_{uv} = 1$ measures overlap
- $\omega_{uv} = \sqrt{d_u d_v}$ measures cosine of angle
- $\omega_{uv} = \min\{d_u, d_v\}$ measures relative overlap
- $\omega_{uv} = \max\{d_u, d_v\}$

Here, d_u is the number of nonzeros of column u.



Universiteit Utrecht

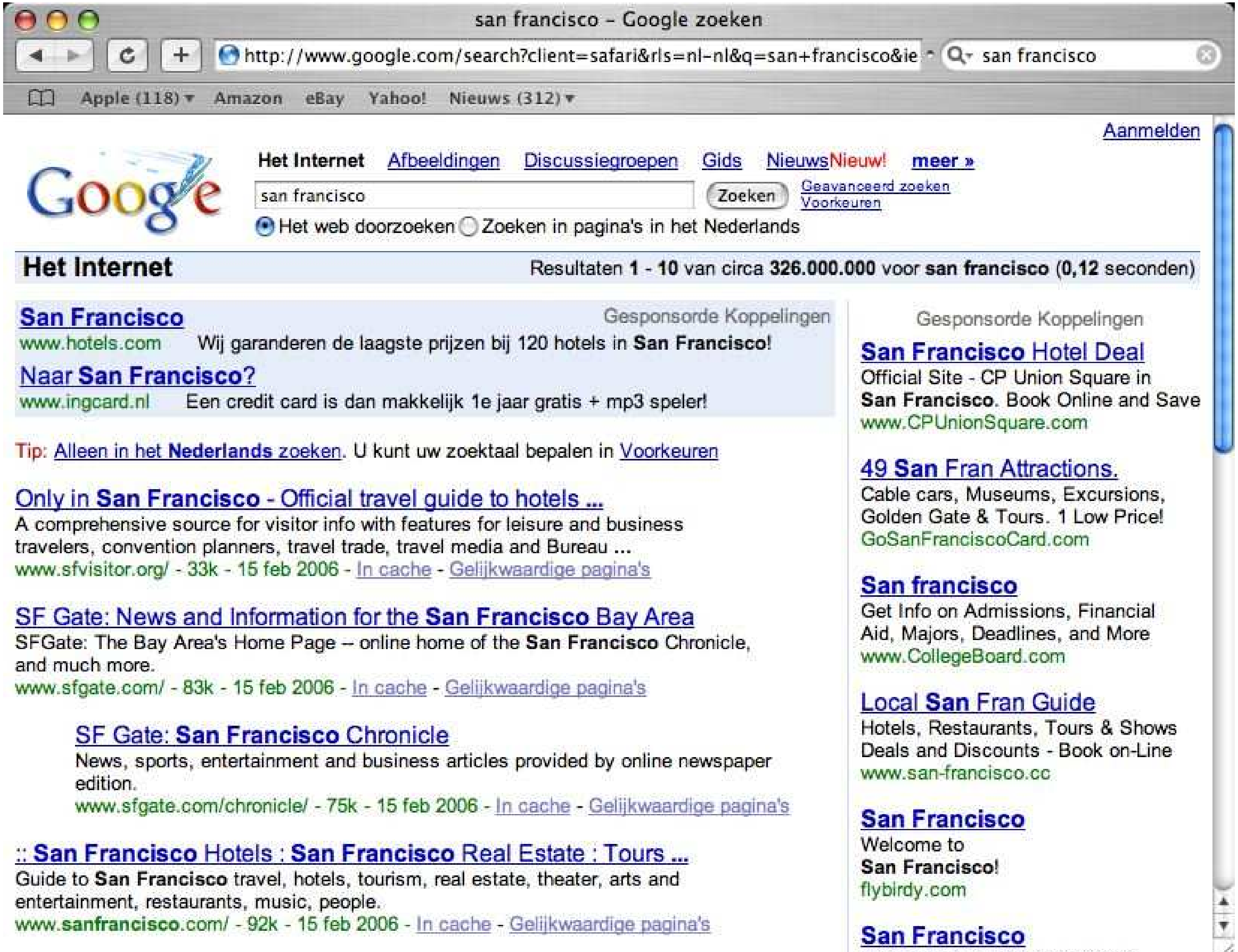
Speeding up the fine-grain method



- ip = standard inner product matching
- ip1 = inner product matching using an upper bound on the overlap, e.g. d_u to stop searching early.
  For fine-grain method, bound is sharper: 1 at first level.
- ip2 = alternate between matching with overlap in top and bottom rows.
- rnd = choose a random match with overlap ≥ 1

Universiteit Utrecht

Web searching: which page ranks first?



Universiteit Utrecht

The link matrix A

- Given n web pages with links between them.
  We can define the sparse $n \times n$ link matrix A by

$$a_{ij} = \begin{cases} 1 & \text{if there is a link from page } j \text{ to page } i \\ 0 & \text{otherwise.} \end{cases}$$

- Let $\mathbf{e} = (1, 1, \ldots, 1)^T$, representing an initial uniform importance (rank) of all web pages. Then

$$(A\mathbf{e})_i = \sum_j a_{ij} e_j = \sum_j a_{ij}$$

  is the total number of links pointing to page i.

- The vector $A\mathbf{e}$ represents the importance of the pages; $A^2\mathbf{e}$ takes the importance of the pointing pages into account as well; and so on.

The Google matrix

- A web surfer chooses each of the outgoing N_j links from page j with equal probability. Define the $n \times n$ diagonal matrix D with $d_{jj} = 1/N_j$.
- Let α be the probability that a surfer follows an outlink of the current page. Typically $\alpha = 0.85$. The surfer jumps to a random page with probability $1 - \alpha$.
- The Google matrix is defined by (Brin and Page 1998)

$$G = \alpha A D + (1-\alpha)\mathbf{e}\mathbf{e}^T/n.$$

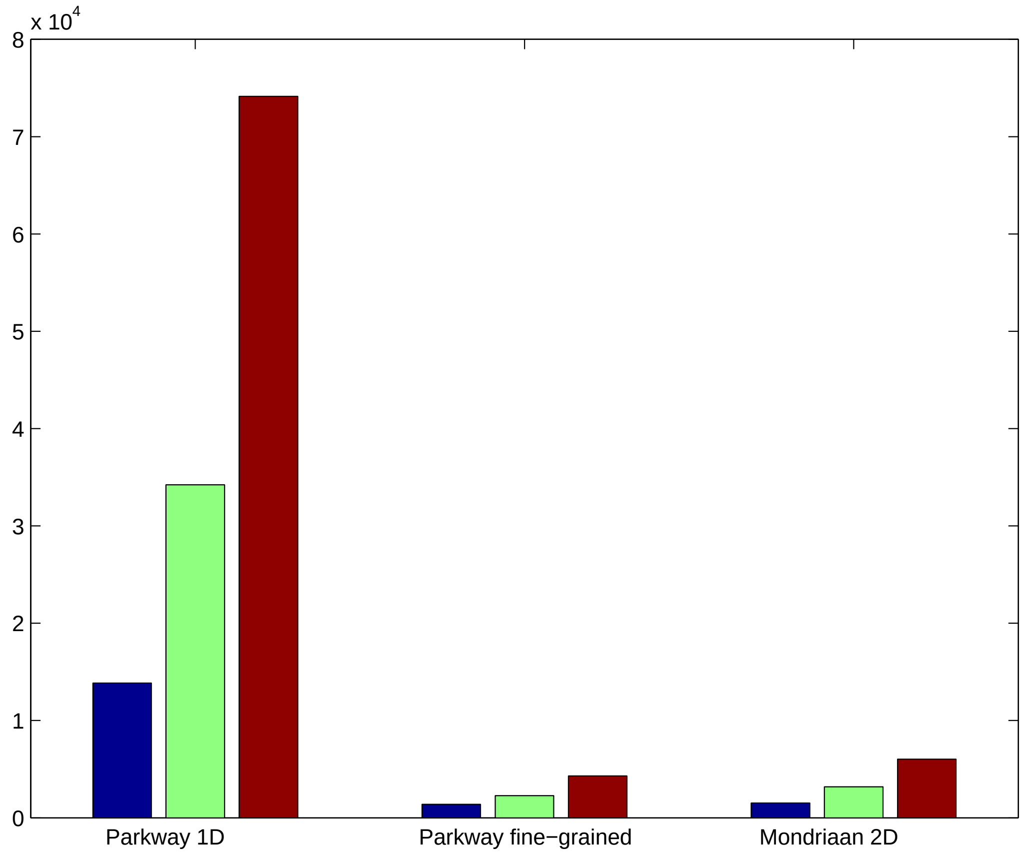
- The PageRank of a set of web pages is obtained by repeated multiplication by G, involving sparse matrix–vector multiplication by A, and some vector operations.

Universiteit Utrecht

Comparing 1D, 2D fine-grain, and 2D Mondriaan

- The following 1D and 2D fine-grain communication volumes for PageRank matrices are published results from the parallel program Par$k$way v2.1 (Bradley, de Jager, Knottenbelt, Trifunović 2005).
- The 2D Mondriaan volumes are results with all our improvements (to be incorporated in v2.0), but using only row/column partitioning, not the fine-grain option.

Communication volume: PageRank matrix `Stanford`

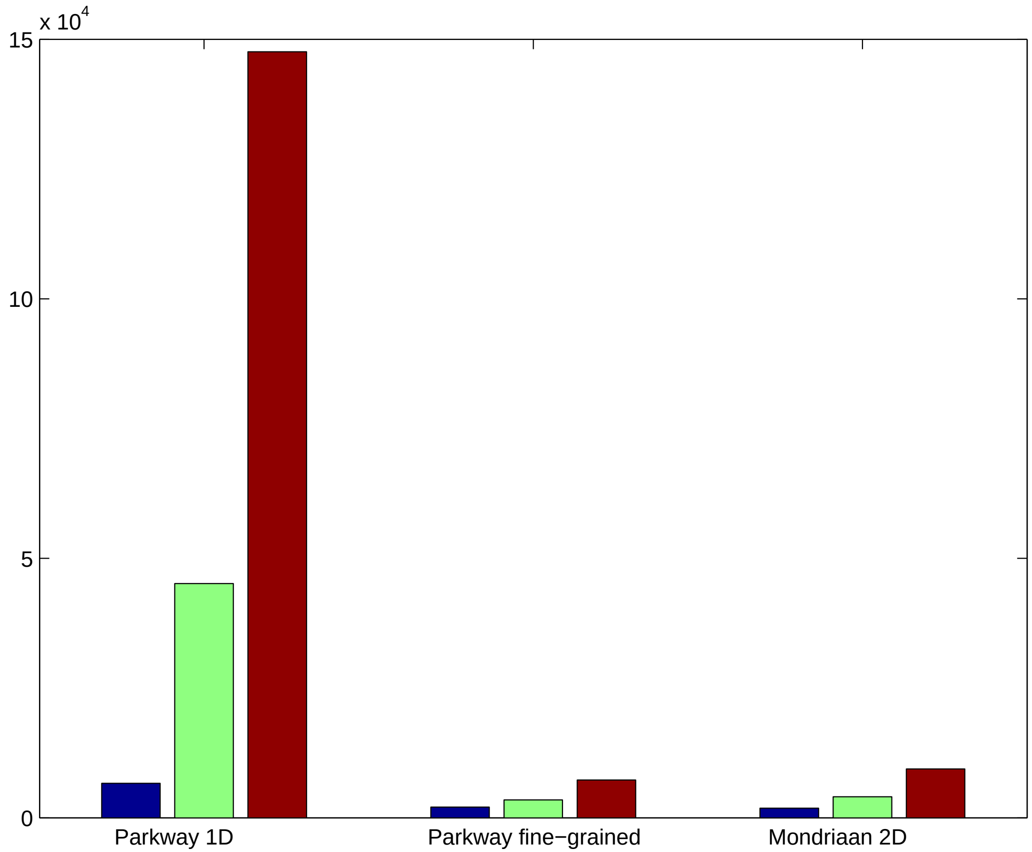


$p = 4, 8, 16$

- $n = 281,903$ (pages), $nz(A) = 2,594,228$ nonzeros (links).
- Represents the Stanford WWW subdomain, obtained by a web crawl in September 2002 by Sep Kamvar.

Universiteit Utrecht

Communication volume: `Stanford_Berkeley`



$p = 4, 8, 16$

- $n = 683,446$, $nz(A) = 8,262,087$ nonzeros.
- Represents the Stanford and Berkeley subdomains, obtained by a web crawl in Dec. 2002 by Sep Kamvar.

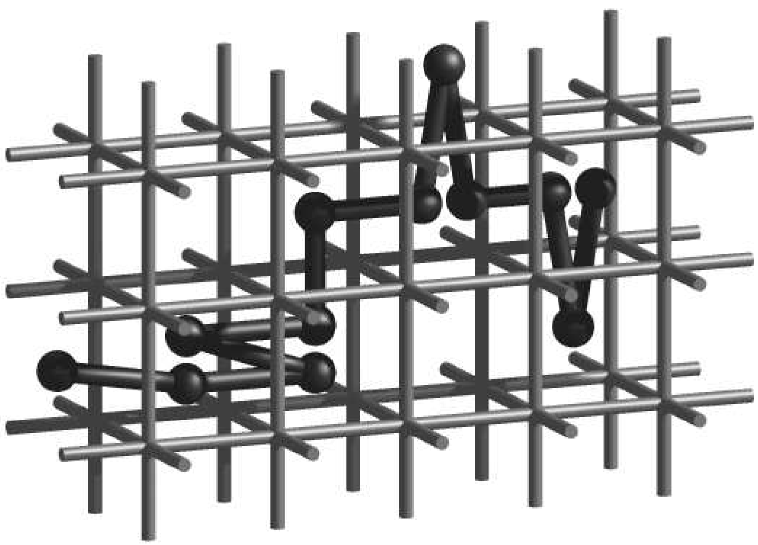
Universiteit Utrecht

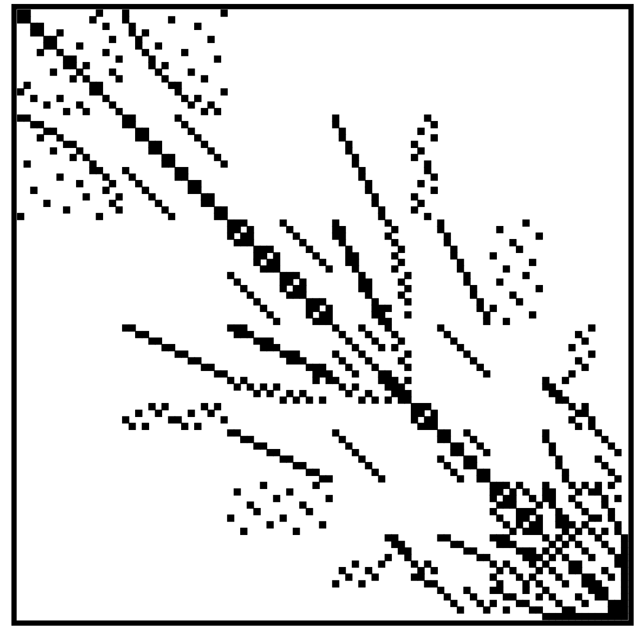
Meaning of results

- Both 2D methods save an order of magnitude in communication volume compared to 1D.
- Parkway fine-grain is slightly better than Mondriaan, in terms of partitioning quality. This may be due to a better implementation, or due to the fine-grain method itself. Further investigation is needed.
- 2D Mondriaan is much faster than fine-grain, since the hypergraphs involved are much smaller: 7×10^5 vs. 8×10^6 vertices for `Stanford_Berkeley`.

Universiteit Utrecht

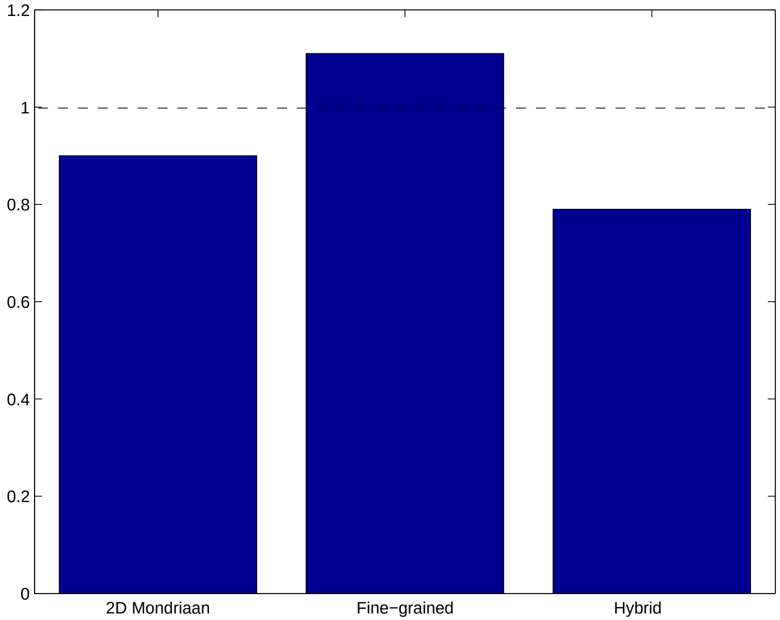
Transition matrix cage6 of Markov model





- Reduced transition matrix `cage6` with $n = 93$, $nz(A) = 785$ for polymer length $L = 6$.
- Larger matrix `cage10` is included in our test set of 18 matrices representing various applications: 3 linear programming matrices, 2 information retrieval, 2 chemical engineering, 2 circuit simulation, 1 polymer simulation, . . .

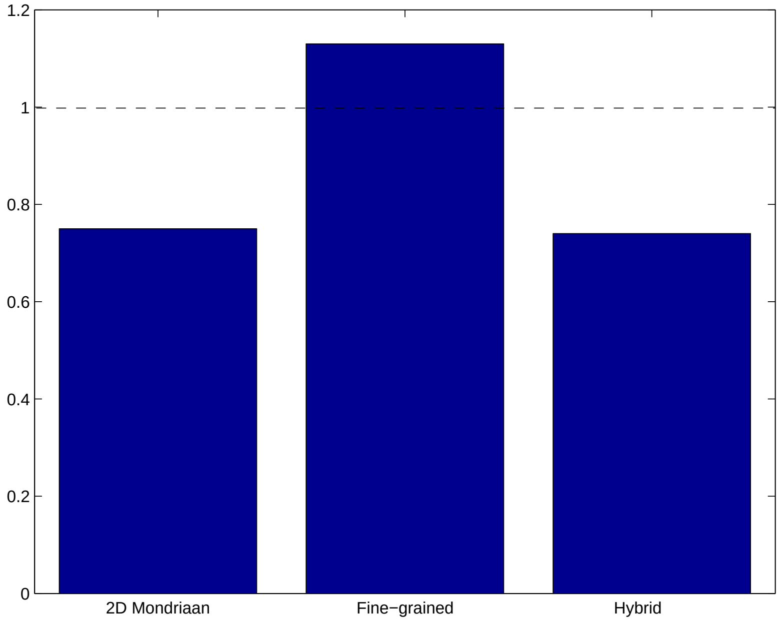
Average communication volume for 3 methods



- Test set of 18 matrices (smaller than PageRank matrices).
- Volume relative to original Mondriaan program, v1.02
- Implementation: Mondriaan's own hypergraph partitioner
- Fine-grained method has more freedom to find a good partitioning, but shows no gains on average

Universiteit Utrecht

Average communication volume for 3 methods



- Test set of 18 matrices.
- Volume relative to original Mondriaan program, v1.02
- Implementation: PaToH hypergraph partitioner.
  Highly optimised, and it shows.
- Hybrid method shows a little gain over 2D Mondriaan

Conclusions and ...

- We have presented a new hybrid method which combines two different 2D matrix partitioning methods: Mondriaan and fine-grain. The hybrid improves upon both.
- With a highly optimised hypergraph partitioner such as PaToH as the partitioning engine, the Mondriaan 2D method achieves almost the same quality as the hybrid method, but much faster.
- PageRank is a wonderful non-PDE application:
  - it affects our lives daily
  - it has embedded mathematical high technology
  - it uses the power method; only mathematicians and computer scientists know what this really means!
  - it exposes the power of 2D matrix partitioning methods

. . . *future work*

- We keep on improving the Mondriaan and PaToH hypergraph partitioners.
- New release of Mondriaan, v2.0, will incorporate all improvements.
- Mondriaan and PaToH are sequential.
- Soon, the parallel hypergraph partitioner Zoltan will be released by Sandia National Laboratories (Devine, Boman, Heaphy, Bisseling, Çatalyürek 2006), with many features from Mondriaan and PaToH, and a lot more.
- First parallel partitioner Par$k$way 2.1 (Knottenbelt, Trifunović 2005) is also publicly available.
- Partition PageRank in parallel!