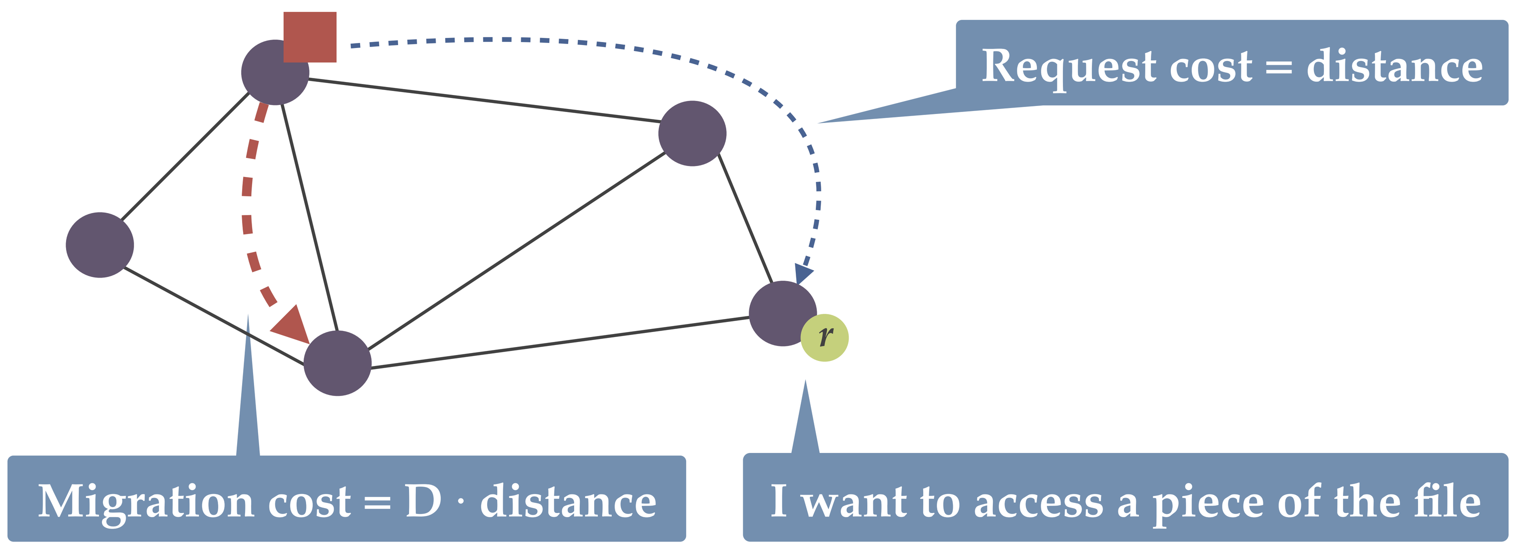
On phase-based algorithms for file migration

Marcin Bieńkowski, Jarosław Byrka, Marcin Mucha

File migration problem

(Originally: *page migration*.) An online problem defined in a weighted graph. There is a single file of size D stored at one node. In a single round:

- a node requests access to the file,
- an algorithm may migrate (the file) to a new node.



Goal: minimize the total cost (request costs + migration costs).
Performance metric: competitive ratio (online-to-optimal-offline ratio).

Our focus: deterministic algorithms for large D (i.e., $1/D$ is negligible).

Previous results

	lower bound	upper bound
trees and uniform metrics	3 [1]	3 [1,2]
arbitrary graphs	3.0000074 [3]	**4.086 [4]**

[1] Black, Sleator '89
[2] Chrobak, Larmore, Reingold, Westbrook '97
[3] Matsubayashi '15
[4] Bartal, Charikar, Indyk '97 (algorithm MTLM)

Many other results for randomized algorithms and for small D values.

Move-To-Local-Min (MTLM)

Parameterized by constants c and α.
Main idea: balance request costs and migration costs.

- Operates in phases of length $c \cdot D$.
- Within a phase, keeps the file at A_{init} and serves requests $r_1, r_2, \ldots, r_{cD}$.
- After a phase, migrates to a minimizer of the function

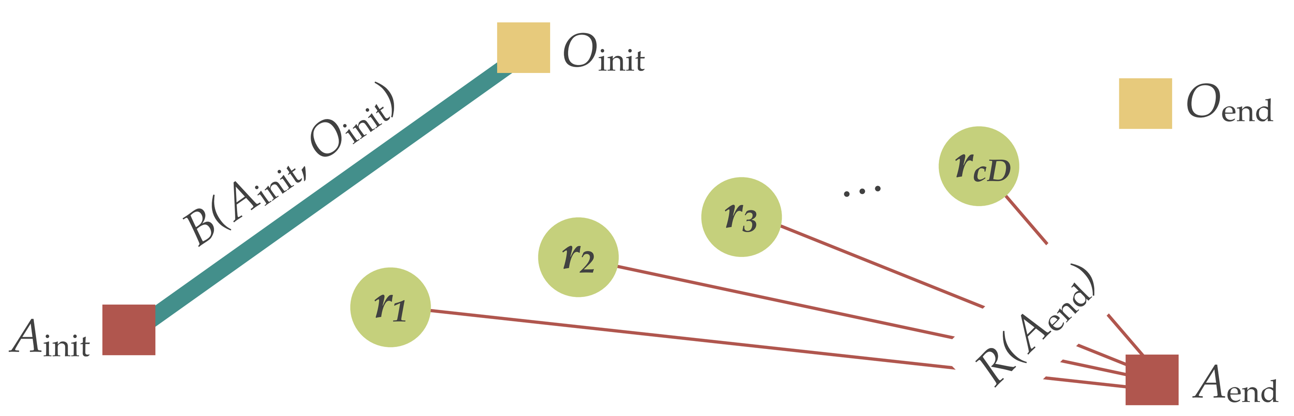
$$f(x) = D \cdot d(A_{\text{init}}, x) + \alpha \cdot \sum_{i=1}^{c \cdot D} d(x, r_i)$$

Make close migrations — Migrate towards requests

Original proof = long stream of creative applications of triangle inequality.

Recreating the analysis for MTLM

Fix values of c and α and consider a single phase.
Create the variables: $B(x,y) := D \cdot d(x,y)$ and $R(x) := \Sigma_{i=1..cD}\, d(x, r_i)$



Consider the following LP:
maximize $C_{\text{MTLM}} = R(A_{\text{init}}) + B(A_{\text{init}}, A_{\text{end}})$
subject to:

- $C_{OPT} = 1$
- $f(A_{\text{end}}) \leq f(x)$ for $x \in \{A_{\text{init}}, O_{\text{init}}, O_{\text{end}}\}$
- $C_{OPT} \geq B(O_{\text{init}}, O_{\text{end}})$
- $2\, C_{OPT} \geq R(O_{\text{start}}) + R(O_{\text{end}}) + (2-c) \cdot B(O_{\text{init}}, O_{\text{end}})$ — technical claim
- $B(\cdot,\cdot)$ and $R(\cdot)$ satisfy triangle inequalities

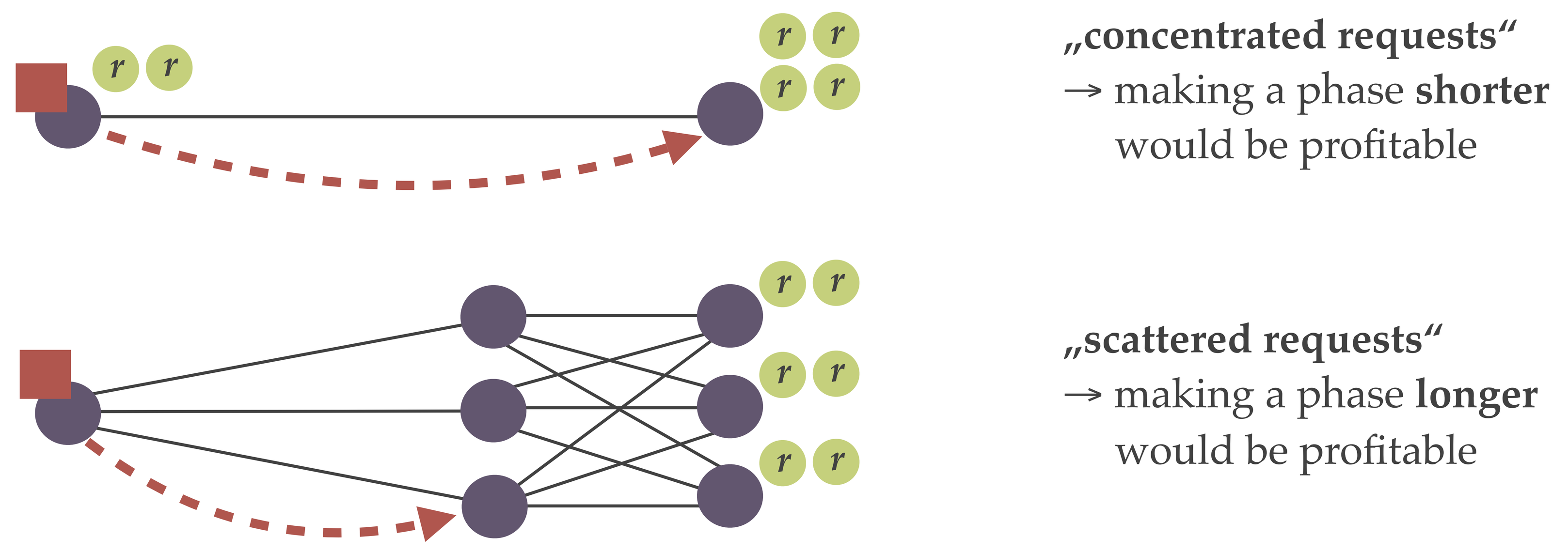
An input satisfies all LP constraints
→ LP value ≥ competitive ratio of MTLM.

Output of factor-revealing LP

For c = 1.841 and α = 0.324, the LP value is indeed 4.086.

- Computer-based proof for MTLM performance.
- Mathematical proof can be obtained by looking at the dual solution.

LP returns tight examples:



Sidenote: We used these tight examples to show that no phase-based algorithm operating in phases of fixed-length can beat ratio 4.086.

Improvement idea

Choose parameters $c_S \leq c \leq c_L$ and some (linear) functions f_S and f_L.

If after $c_S \cdot D$ steps, the requests are "*rather concentrated*", an algorithm ends a (short) phase and migrates to a minimizer of f_S. Otherwise, the requests are "*rather scattered*" and after $c_L \cdot D$ steps the requests remain „*somewhat scattered*". An algorithm ends a (long) phase there and migrates to a minimizer of f_L.

Problem: We do not know how to define "rather concentrated"!

Our result: 4-competitive algorithm DLM

Choose value R_S.
Short phase: migrate to a minimizer of f_S if DLM-to-OPT ratio is at most R_S.
Long phase: migrate to a minimizer of f_L.

Proof framework:

- The ratio in a short phase is at most R_S by the definition.
- Bounding the ratio R_L in a long phase:
  - Write a maximization LP (similar to the LP for MTLM).
  - **Add additional constraints** stating that if DLM migrated already after a short phase, then its competitive ratio would be greater than R_S.
- The competitive ratio of DLM is max $\{R_S, R_L\}$.

Result: 4-competitive algorithm (non-constructive computer-based proof).

- For $R_S = 4$, it is possible to choose parameters c_S, c_L and coefficients for linear functions f_S and f_L, so that $R_L \leq 4$.
- The functions f_S and f_L split requests of a phase into several parts, with different emphasis on different parts.
- Parameters and coefficients were found by a computer-aided local search.

Constructive statement of DLM

Again, mathematical proof can be obtained by looking at the dual solution. This time however, it only guarantees that migrating either at the end of a short phase or at the end of a long one yields 4-competitiveness.

We managed to extract an actual constructive rule from the LP output:

$$f_S(x) = D \cdot d(A_{\text{init}}, x) + 2 \cdot \sum_{i=1}^{D} d(x, r_i) + \frac{3}{4} \sum_{i=D+1}^{(7/4) \cdot D} d(x, r_i)$$

$$f_L(x) = D \cdot d(A_{\text{init}}, x) + \sum_{i=1}^{D} d(x, r_i) + \frac{5}{16} \sum_{i=D+1}^{(7/4) \cdot D} d(x, r_i) + \frac{3}{8} \sum_{i=(7/4) \cdot D+1}^{(9/4) \cdot D} d(x, r_i)$$

If $\arg\min_x f_S(x) \leq \frac{9}{8} \sum_{i=D+1}^{(7/4) \cdot D} d(x, r_i)$, migrate to a minimizer of f_S.