



# Graphs in Machine Learning

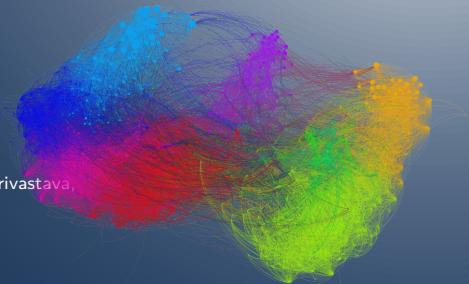
## Distributed Graph Processing

Challenges and Requirements

Michal Valko

*Inria & ENS Paris-Saclay, MVA*

Partially based on material by: Rob Fergus, Nikhil Srivastava,  
Yiannis Koutis, Joshua Batson, Daniel Spielman



# Distributed graph processing

Large graphs do not fit in memory

Get more memory

Many challenges

# Distributed graph processing

Large graphs do not fit in memory

Get more memory

↳ Either slower but larger memory

Many challenges

# Distributed graph processing

Large graphs do not fit in memory

Get more memory

- ↳ Either slower but larger memory  
Or fast memory but divided among many machines

Many challenges

# Distributed graph processing

Large graphs do not fit in memory

Get more memory

- ↳ Either slower but larger memory  
Or fast memory but divided among many machines

Many challenges

Needs to be scalable

# Distributed graph processing

Large graphs do not fit in memory

Get more memory

- ↳ Either slower but larger memory  
Or fast memory but divided among many machines

Many challenges

Needs to be scalable

- ↳ minimize pass over data / communication costs

# Distributed graph processing

Large graphs do not fit in memory

Get more memory

- ↳ Either slower but larger memory  
Or fast memory but divided among many machines

Many challenges

Needs to be scalable

- ↳ minimize pass over data / communication costs

Needs to be consistent

# Distributed graph processing

Large graphs do not fit in memory

Get more memory

- ↳ Either slower but larger memory  
Or fast memory but divided among many machines

Many challenges

Needs to be scalable

- ↳ minimize pass over data / communication costs

Needs to be consistent

- ↳ updates should propagate properly



# Distributed graph processing

Different choices have different impacts: for example splitting the graph according to nodes or according to edges.

Many computation models (academic and commercial) each with its pros and cons

- MapReduce

- MPI

- Pregel

- Graphlab**

# Michal Valko

`michal.valko@inria.fr`

Inria & ENS Paris-Saclay, MVA

`https://misovalko.github.io/mva-ml-graphs.html`

