

# Fast percolation

## Reminder

Percolation is the study of the global connection behaviour of a network when removing a certain fraction of its nodes, measured in the (relative) size of the largest connected component.

- Goal: learn a way of informed percolation which optimally disconnects a network (remove minimum amount of nodes with maximum effect)
- Useful for preventing disease spreading, or to find critical nodes in the network

## Setting up a model

We attach to each node a Bernoulli random variable of failure:

$$\hat{\mathbf{p}}_i = \sigma(\mathbf{w}^T \cdot \mathbf{info}_i)$$
$$\mathbf{Fail}_i(\varphi) \sim \text{Bernoulli} \left( \hat{\mathbf{p}}_i \frac{\varphi}{\frac{1}{n} \sum_i \hat{\mathbf{p}}_i} \right)$$

where  $\sigma$  is the logistic function, **info** is a vector of local information (eg., degree), and **w** the weighting of this information (a logistic model, sort of).

The value  $\varphi$  varies in  $[0; 1]$  and describes the fraction of failing nodes. The individual probabilities are scaled accordingly.

- Minimize the area under the percolation curve with respect to the weights  $\mathbf{w}$
- This involves sampling the curves (several times) at each parameter evaluation
- This is expensive, so I used simulated annealing – it requires the least amount of evaluations of the loss function
- Compare some different network types:

---

	Facebook		Traffic		Random	
	Train	Test	Train	Test	Train	Test
Vertices	572	347	7388	13389	2000	2000
Edges	3192	2519	10591	21246	5000	5000
Density	0.0195	0.0419	0.0004	0.0002	$\approx 0.0025$	

---



