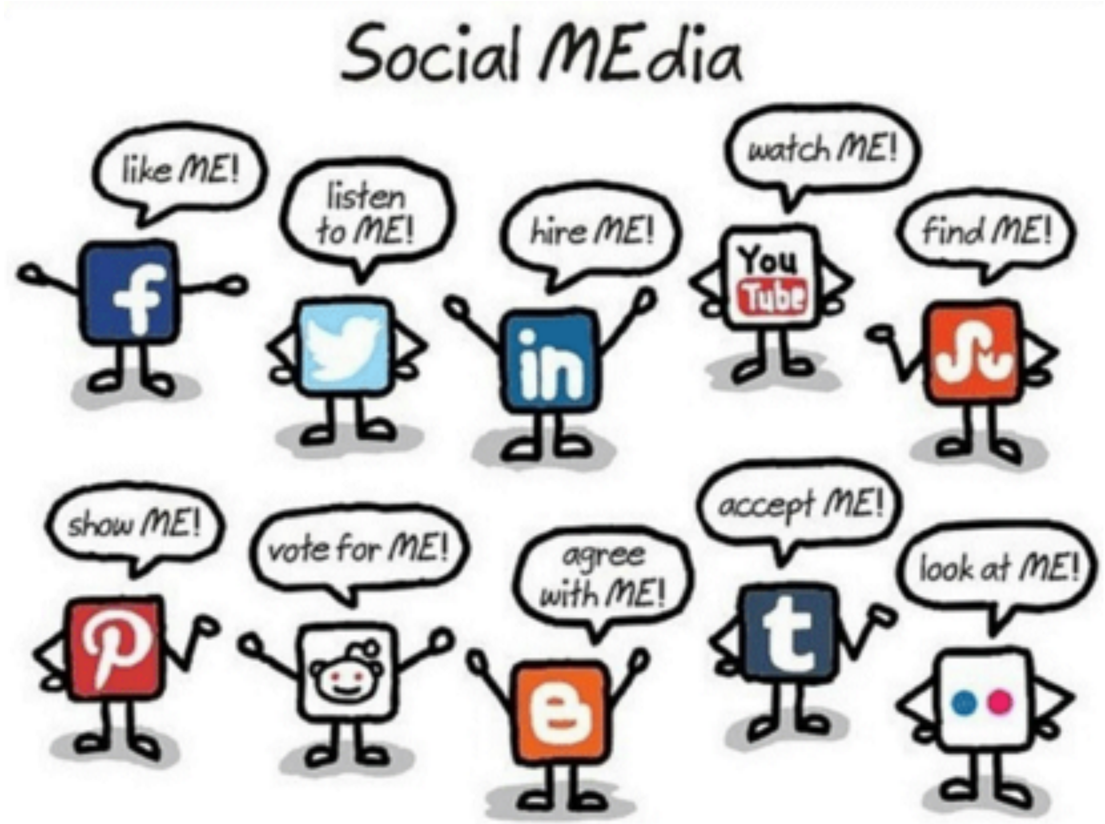
THE POWER OF GRAPHS IN SPEEDING UP ONLINE LEARNING AND DECISION MAKING

Michal Valko

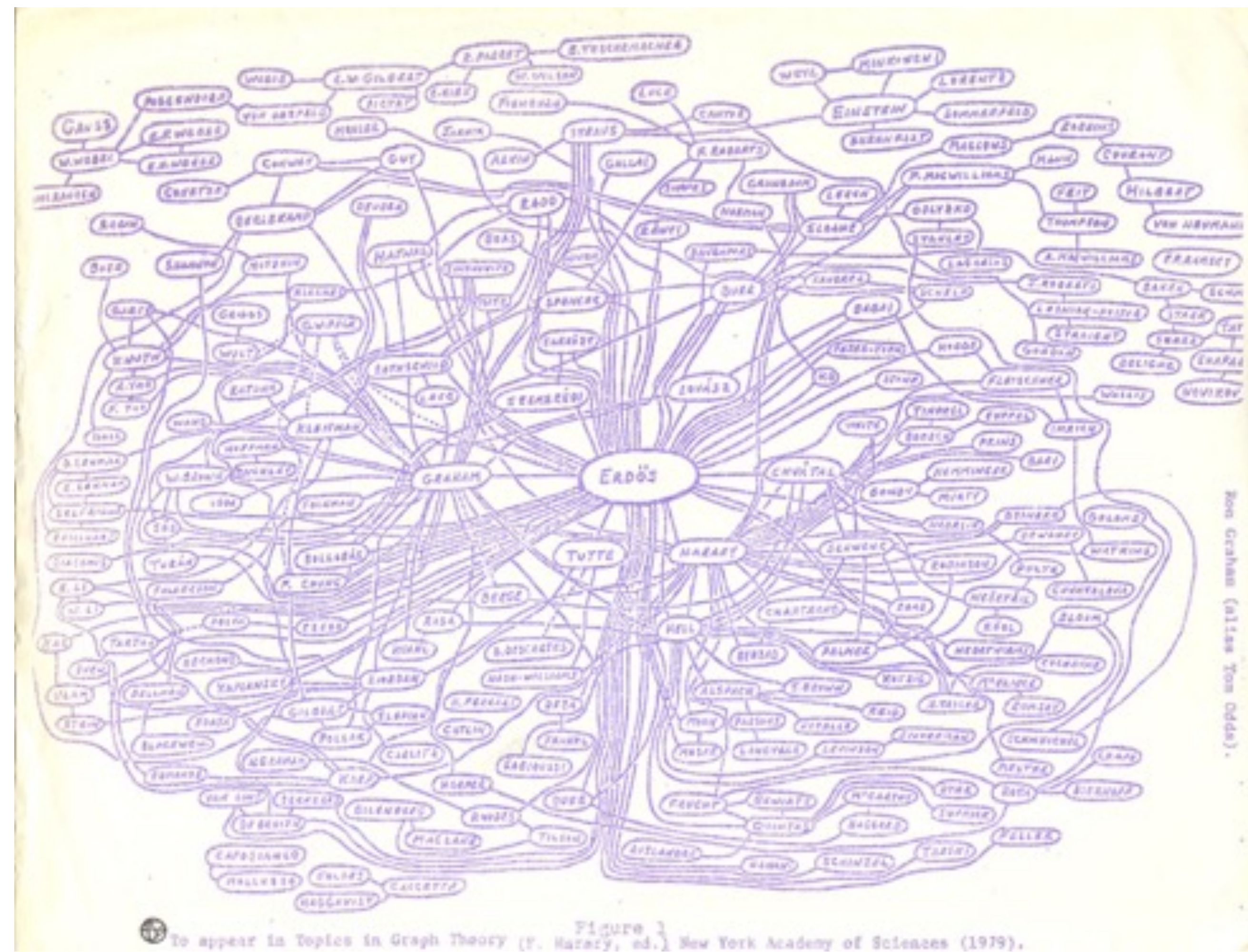
SequeL @ Inria Lille — Nord Europe

inventors for the digital world



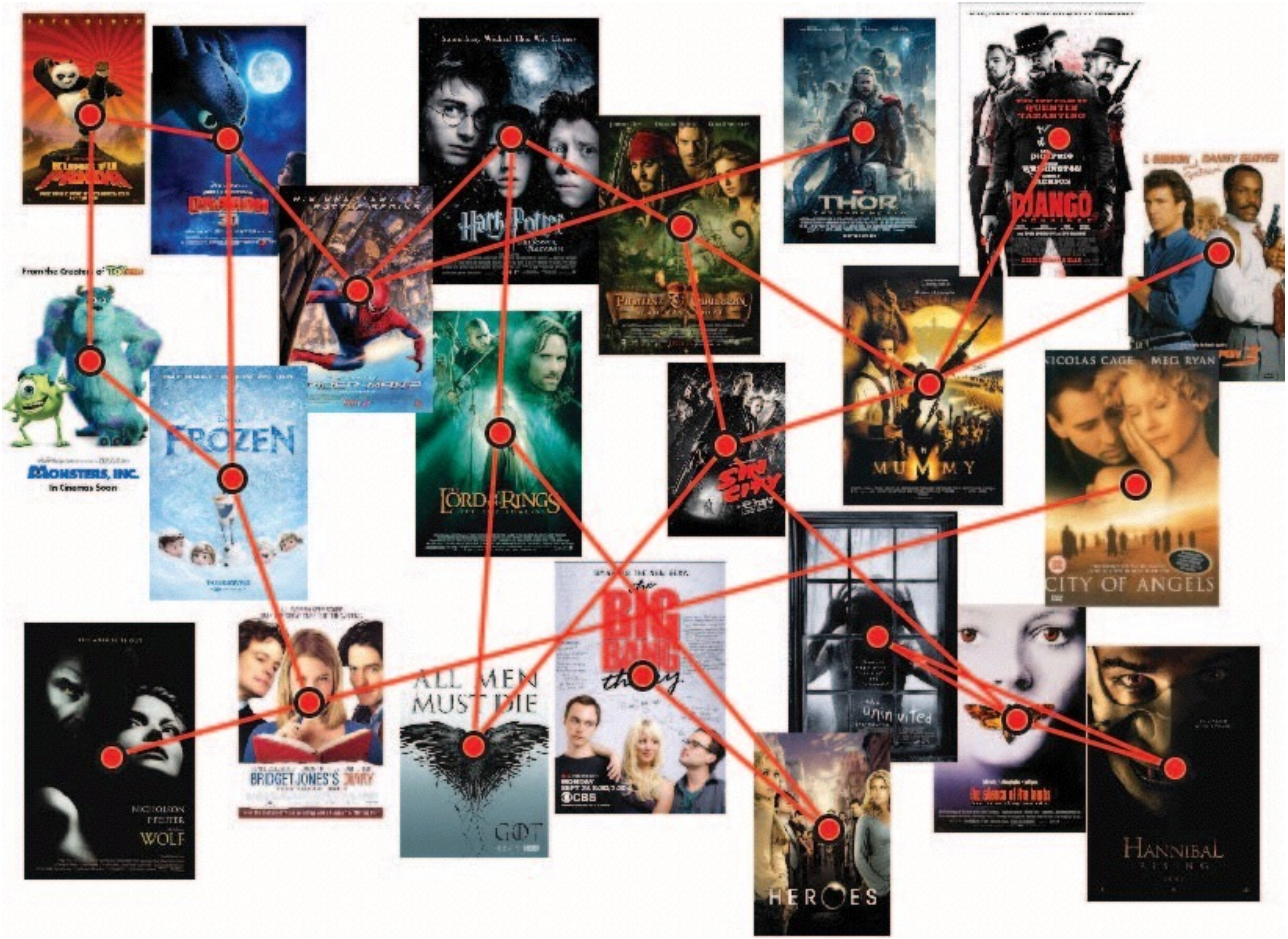
online social networks

Berkeley's floating sensor network

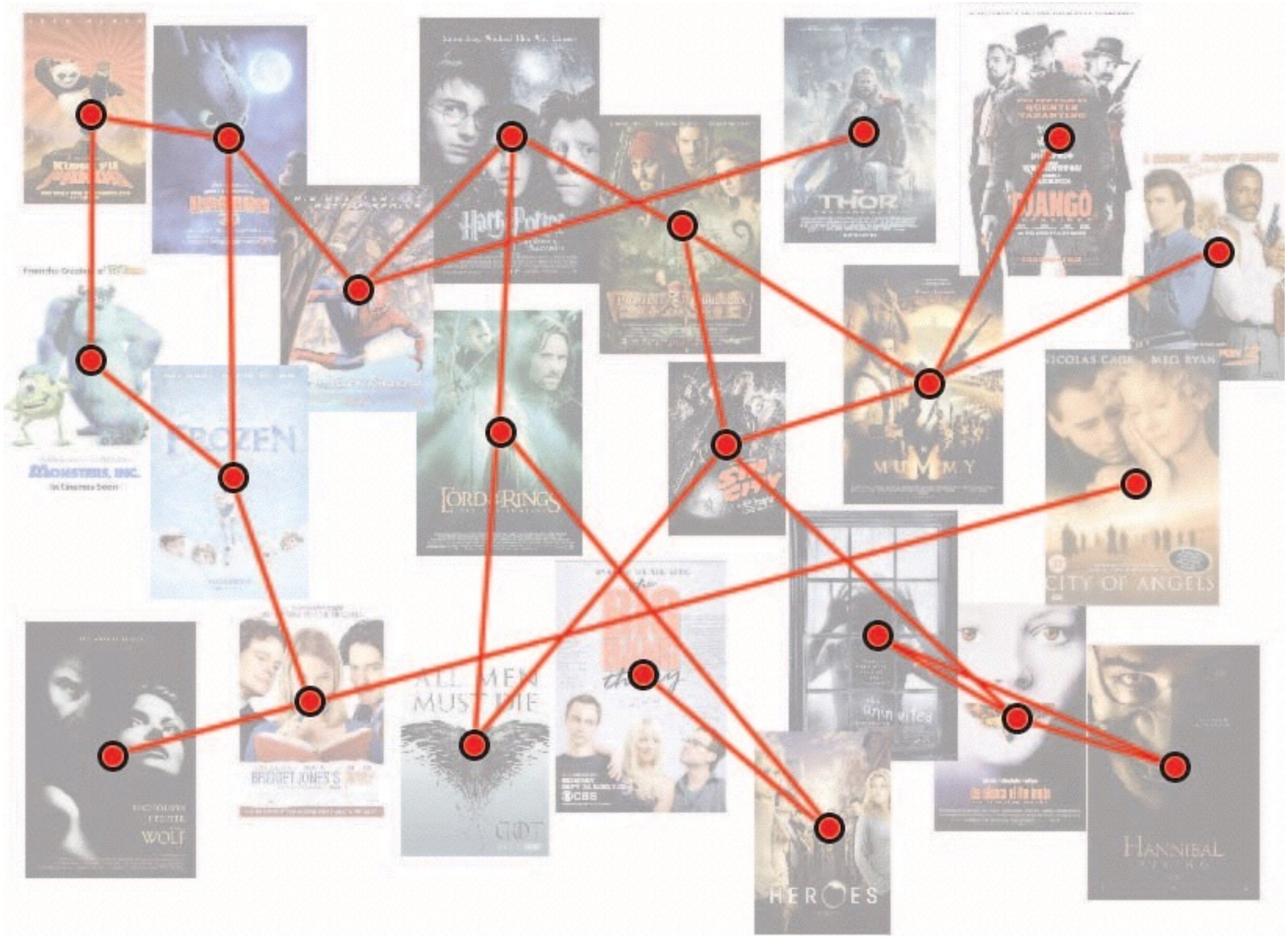


Erdös number project


KUNG FU PANDA
From the Creators of TOY STORY
MONSTERS, INC.
In Cinemas Soon
FROZEN
Harry Potter
THOR
DJANGO
LORD OF THE RINGS
SIN CITY
MUMMY
NICOLAS CAGE
MEG RYAN
CITY OF ANGELS
NICHOLSON
PFEIFFER
WOLF
BRIDGET JONES'S DIARY
ALL MEN MUST DIE
GOT
the BIG BANG theory
CBS
the uninvited
HEROES
HANNIBAL RISING

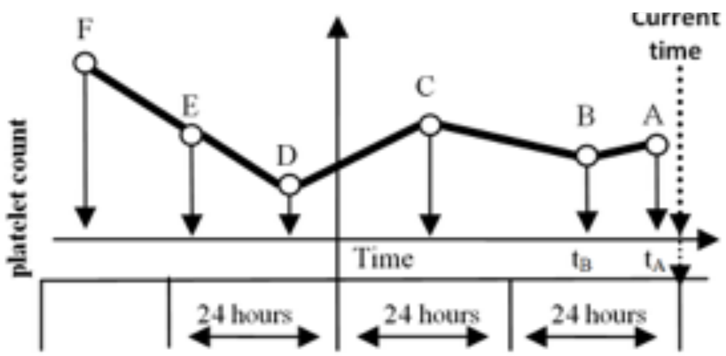


From the Creators of
MONSTERS, INC.
In Cinemas Soon
FROZEN
Harry Potter
THOR
DJANGO
NICOLAS CAGE
MEG RYAN
CITY OF ANGELS
MUMMY
ALL MEN
MUST DIE
GOT
BRIDGET JONES'S DIARY
NICHOLSON
PFEIFFER
WOLF
HEROES
HANNIBAL


THOR
QUENTIN TARANTINO
DJANGO
FROZEN
MONSTERS, INC.
Harry Potter
NICOLAS CAGE MEG RYAN
CITY OF ANGELS
MUMMY
ALL MEN MUST DIE
GOT
HEROES
HANNIBAL
WOLF
BRIDGET JONES'S
CBS

MY PAST 10 YEARS WITH GRAPH IN ML

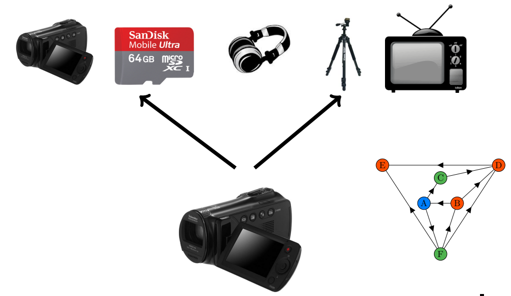
Online anomaly detection for medical decisions



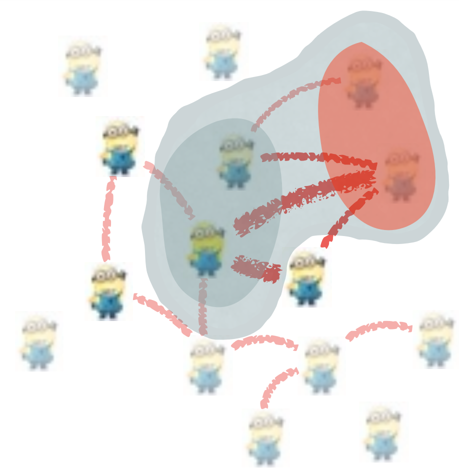
Building good models takes time and they are often unavailable

online decision-making

side observations

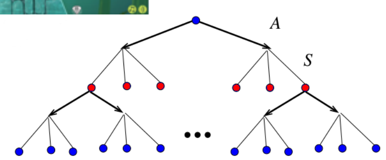


smoothness of rewards



online influence maximization

adaptive structural exploration



Monte-Carlo tree search

2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019

Labels are often costly or unavailable

Online semi-supervised learning for personalization

new master course Graphs in ML at ML MSc program in Paris

Bandits and MDPs with discrete and continuous variables

online graph sparsification

JOINT WORK WITH…


Akram Erraqabi
U Montréal


Alessandro Lazaric
FAIR Paris


Alexandra Carpentier
U. Magdeburg


Azin Ashkan
Google


Branislav Kveton
Google Research


Daniele Calandriello
IIT, Genova


Gergely Neu
U Pompeu Fabra

Ilias Flaounas
Atlassian


Jean-Bastien Grill
DeepMind Paris


Julien Audiffren
U of Fribourg


Manjesh Hanawal
IIT Bombay


Mohammad Ghavamzadeh
FAIR California


Nello Cristianini
University of Bristol


Odalric-Ambrym Maillard
SequeL, Inria


Philippe Preux
SequeL, Inria


Rémi Munos
DeepMind Paris


Shipra Agrawal
Columbia U


Tomáš Kocák
SequeL, Inria


Venkatesh Saligrama
Boston University

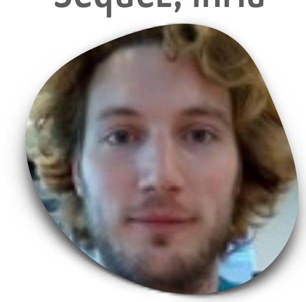

Yaakov Engel
Rafael


Zheng Wen
Adobe Research


Vianney Perchet
CMLA


Edouard Oyallon
CentraleSupelec


Rémi Bardenet
CNRS


Victor Gabillon
QUT Brisbane

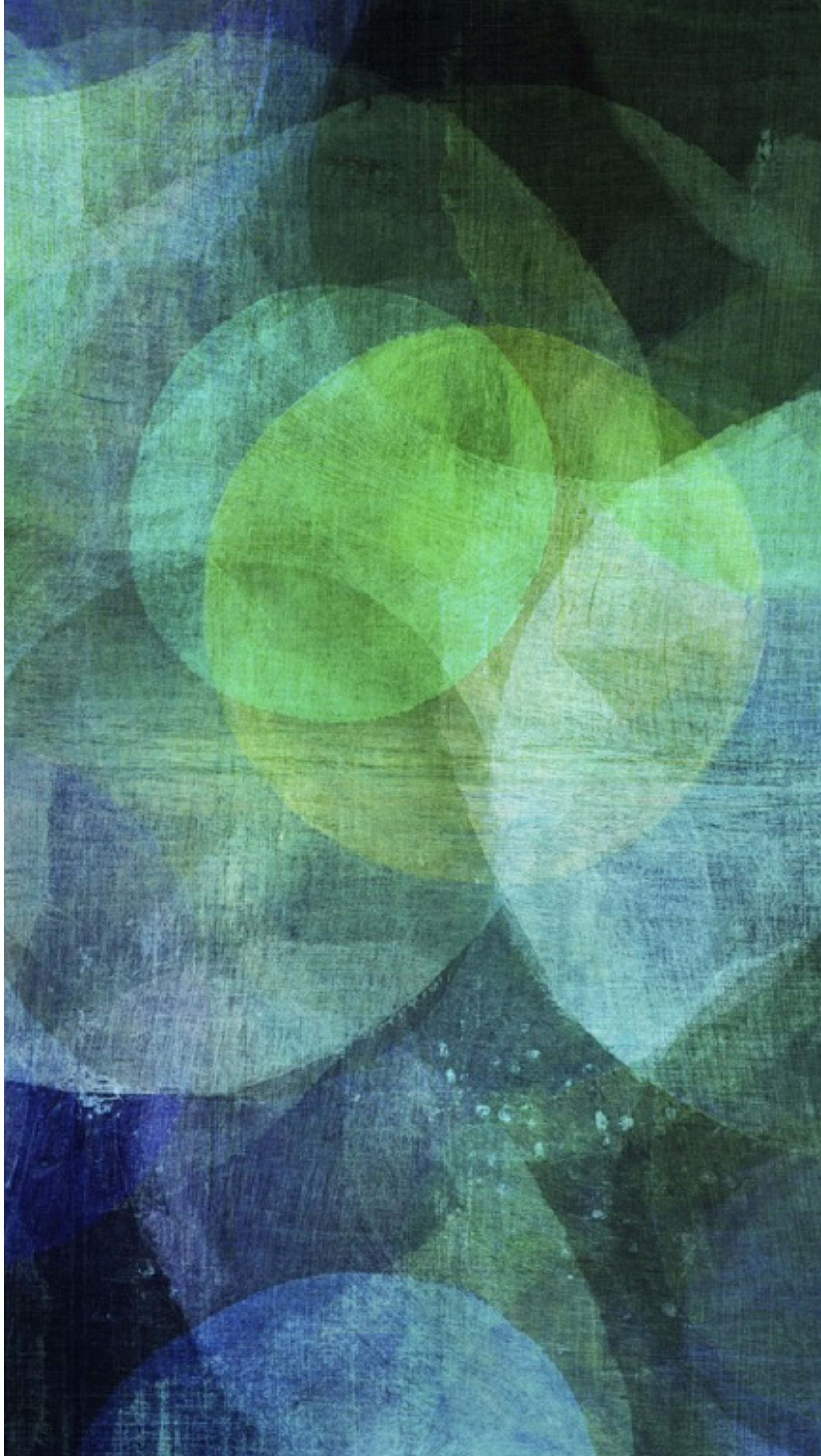

Yasin Abbasi-Yadkori
DeepMind London


Emilie Kaufmann
CNRS


Eugene Belilovsky
MILA

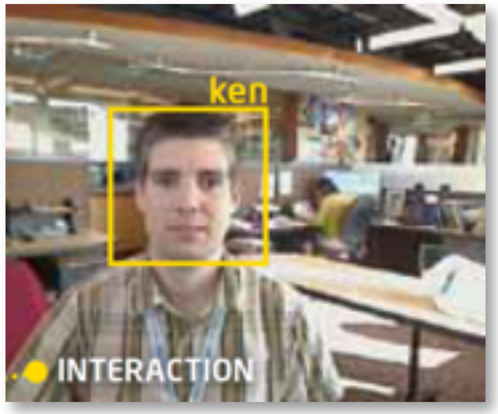
ONLINE LEARNING

when we reason on the fly



IN 2007 IT ALL STARTED WITH AN IDEA…

- Develop **sequential machine learning** recognition system
- System with **minimal feedback**
- 90% accurate over 90% of time
- With **theory** that guarantee's its performance
- **Efficient** (e.g., mobile device)

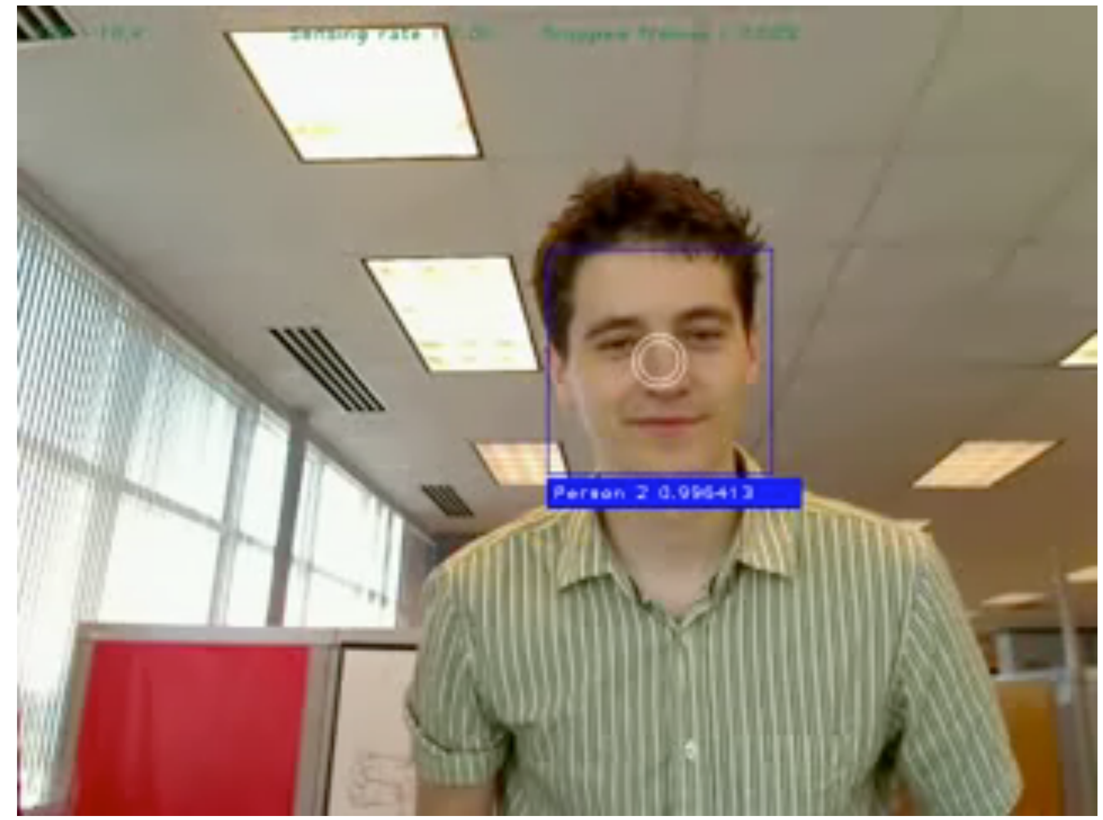


from B. Kveton

… AND RESULTED IN A REAL SYSTEM IN 2009

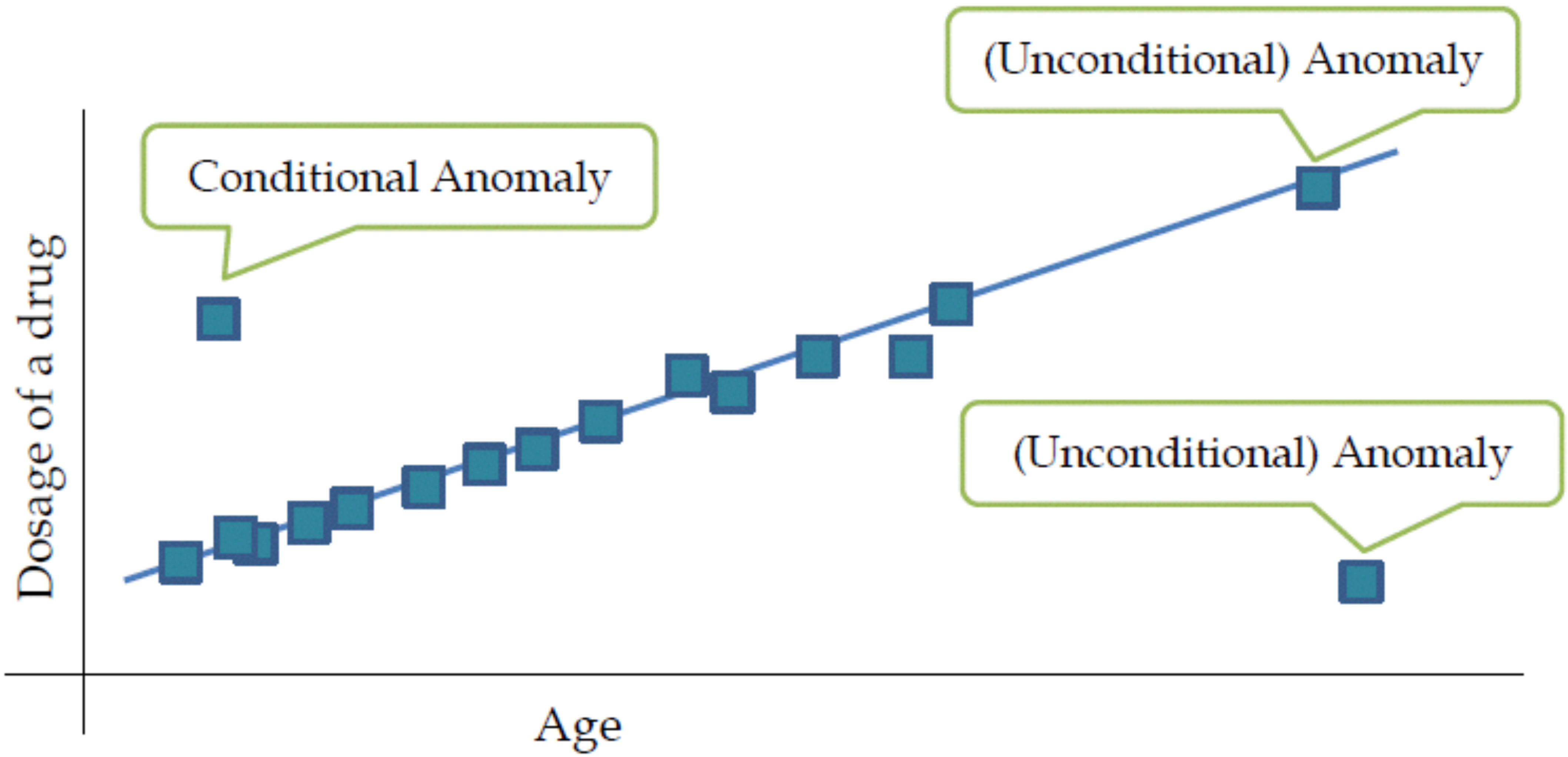
- **adaptive graph-based** recognition system
  - highly accurate
  - trained from a **small amount** of labeled data
  - real-time running time
  - robust to outliers
  - theoretical analysis

$$\frac{1}{n}\sum_{t}(\ell_t^{q}[t]-y_t)^2 \leq \frac{1}{n_l}\sum_{i\in l}(l_i^{*}-y_i)^2 + O(n^{-\frac{1}{2}})$$



from B. Kveton

HEALTH: **CONDITIONAL** ANOMALY DETECTION

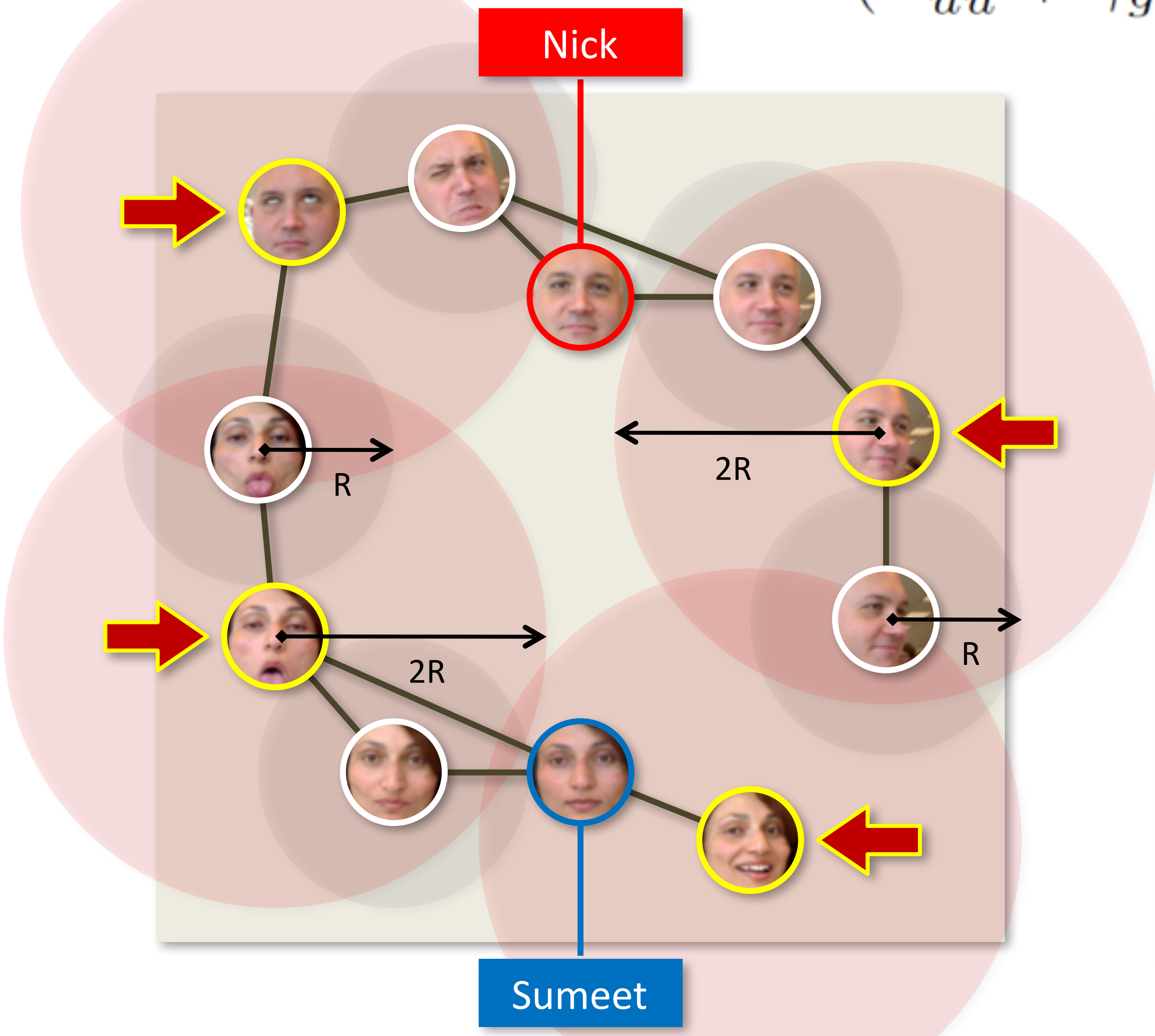


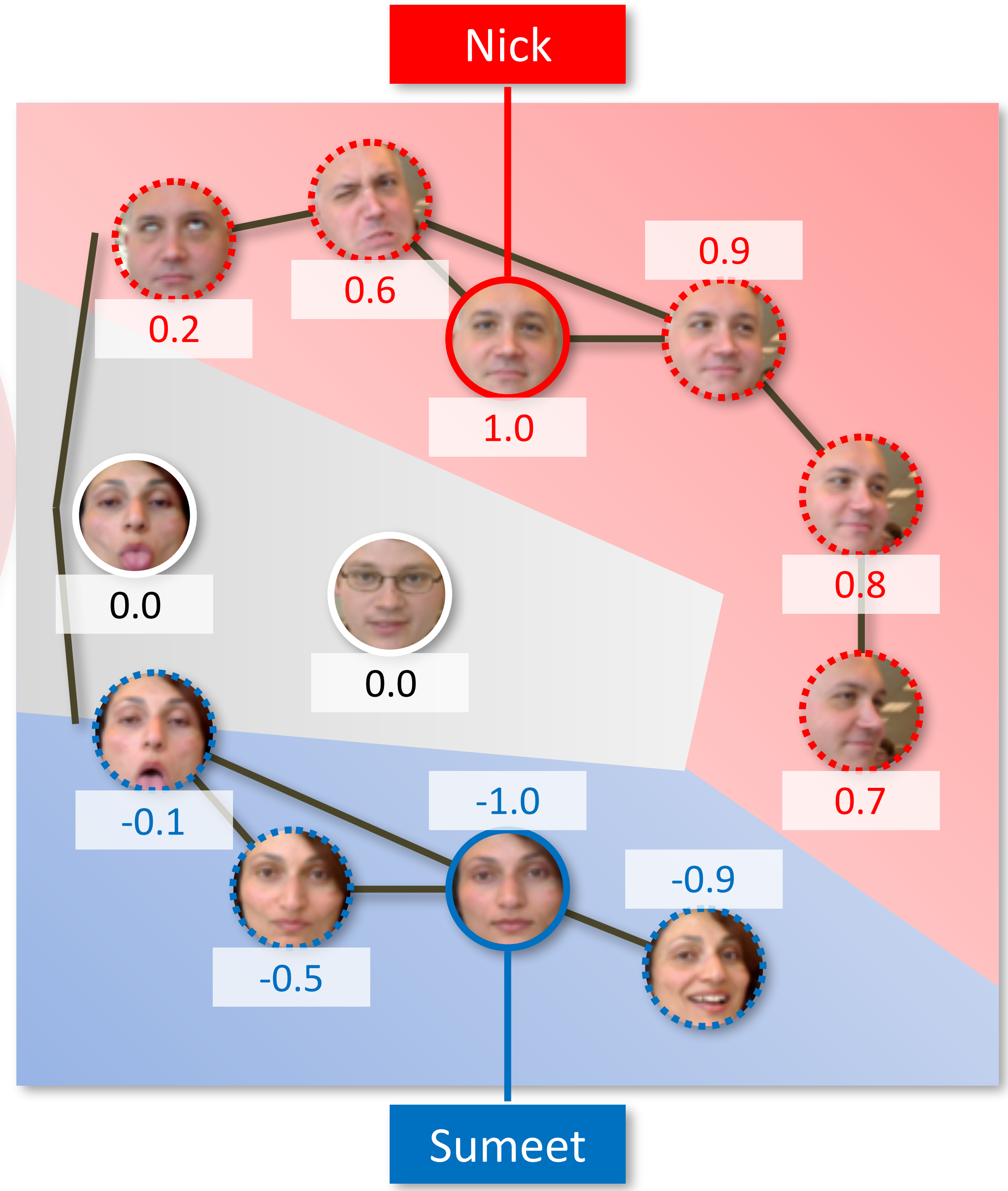
Conditional anomalies are often medical errors.

"Medical errors account for 200 000 preventable deaths a year. "

(HealthGrades study, Wall Street Journal, July 27th 2004)

$$\ell^q = (L^q_{uu} + \gamma_g V)^{-1} W^q_{ul} \ell_l$$



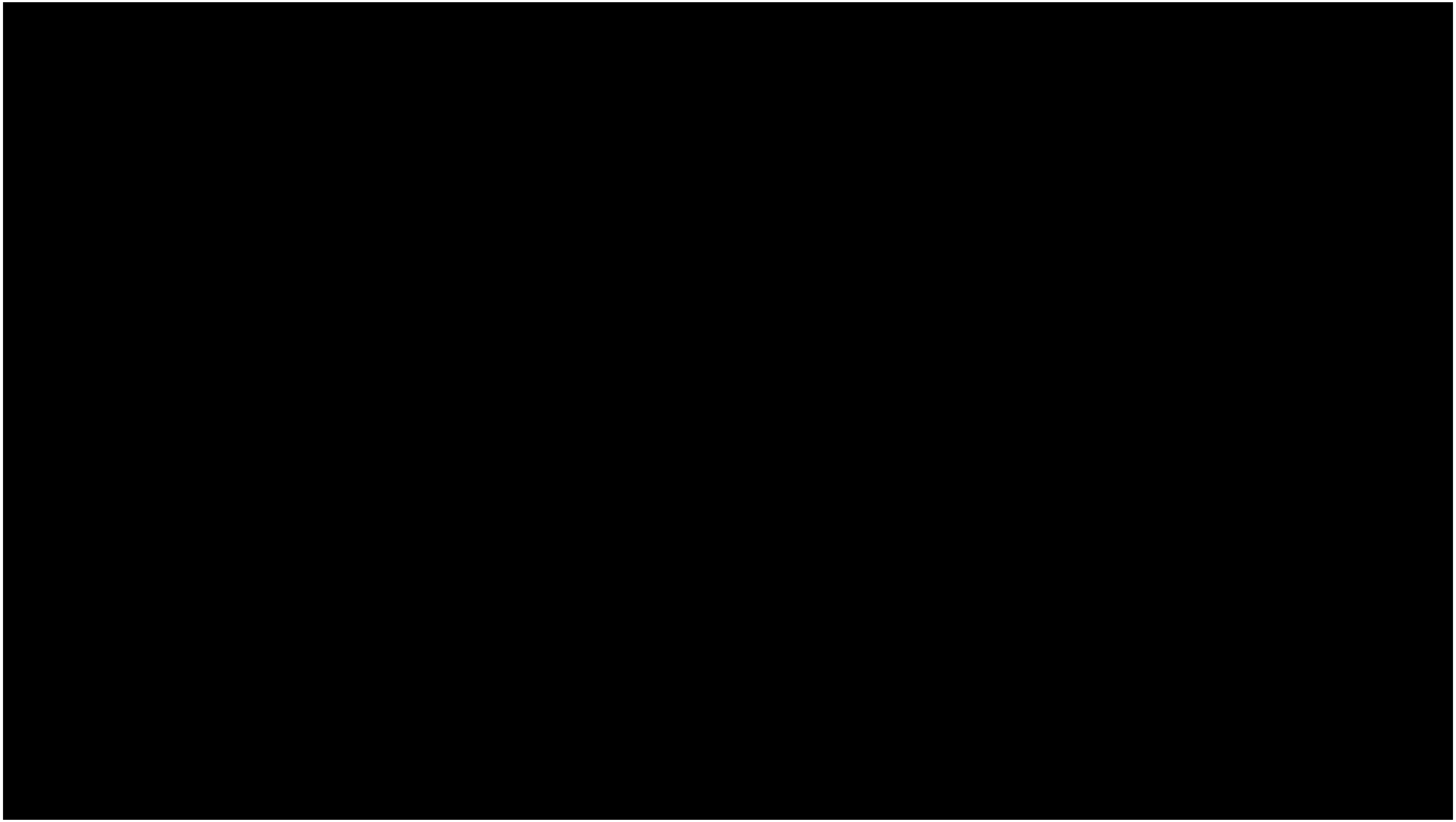


$$\frac{1}{n}\sum_t (\ell^q_t[t] - y_t)^2 \le \frac{3}{n}\sum_t (\ell^*_t - y_t)^2 + \frac{3}{n}\sum_t (\ell^o_t[t] - \ell^*_t)^2 + \frac{3}{n}\sum_t (\ell^q_t[t] - \ell^o_t[t])^2$$

Error of our solution	Offline learning error	Online learning error	Quantization error

Online Semi-Supervised Learning on Quantized Graphs. In Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence, California, July 2010.

FACE-RECOGNITION FOR INTEL



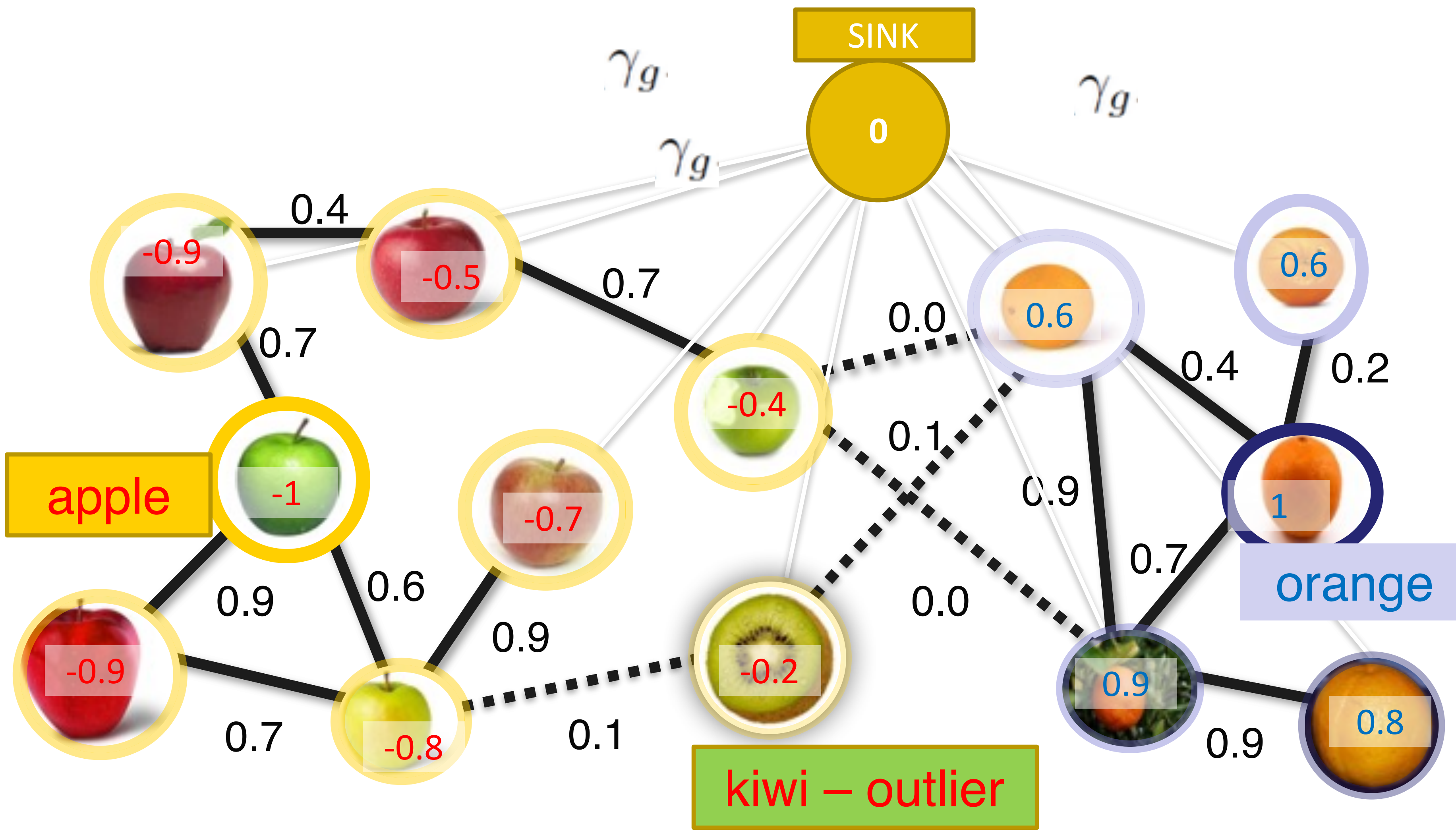
THIS CAN'T SCALE: CONNECTED CAR

Personalization

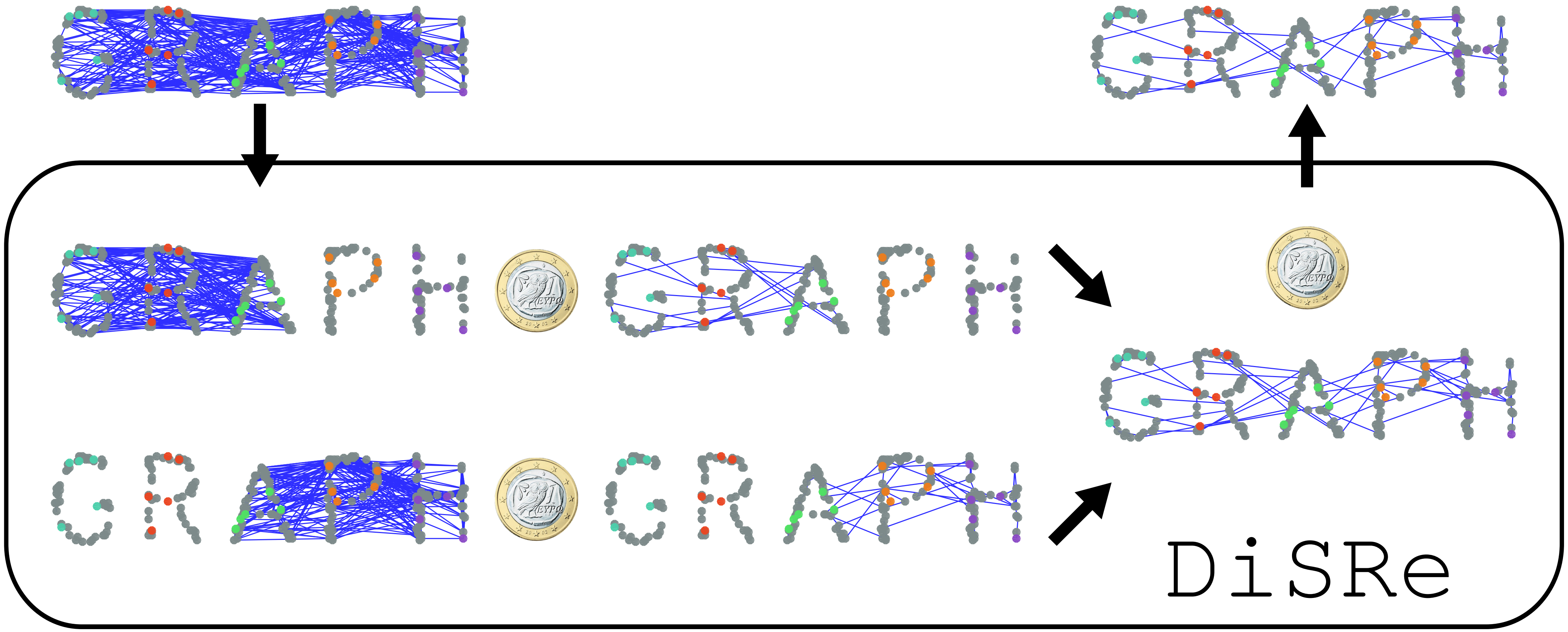
2 BIG REAL-WORLD ISSUES

- SIZE and SPEED
- ANOMALIES

$$\mathbf{f}_u = \left(\mathbf{L}_{uu} + \gamma_g \mathbf{I}\right)^{-1} \left(\mathbf{W}_{ul}\mathbf{f}_l\right)$$



HUGE AND/OR ONLINE



MV, Kveton, Huang, Ting: **Online Semi-Supervised Learning on Quantized Graphs** UAI 2010

Kveton, MV, Rahimi, Huang: **Semi-Supervised Learning with Max-Margin Graph Cuts** AISTATS 2010

Calandriello, Lazaric, MV: **Distributed sequential sampling for kernel matrix approximation** AISTATS 2017

Calandriello, Lazaric, MV: **Second-order kernel online convex optimization with adaptive sketching**, ICML 2017

Calandriello, Lazaric, MV: **Efficient second-order online kernel learning with adaptive embedding**, NIPS 2017

Calandriello, Koutis, Lazaric, MV: **Improved large-scale graph learning through ridge spectral sparsification**, ICML 2018

code: http://researchers.lille.inria.fr/~valko/hp/publications/squeak.py

Industry transfer to intel

- **Context Aware Vehicle**
  - recognizes when your face is turned to the side
- **Everyday Sensing and Perception**
  - health monitoring and assisted living
- **Google TV project**
  - Personalized advertisement
- **Connected Cars**
  - Ford, Toyota, Audi/VW Group, Nissan
- **Intel Phone (marketed in 2015)**
  - adaptive logging in

6870 lines of code in C++ using OpenCV library
2-3 years of research + development

Technology transfer to UPMC (2011)

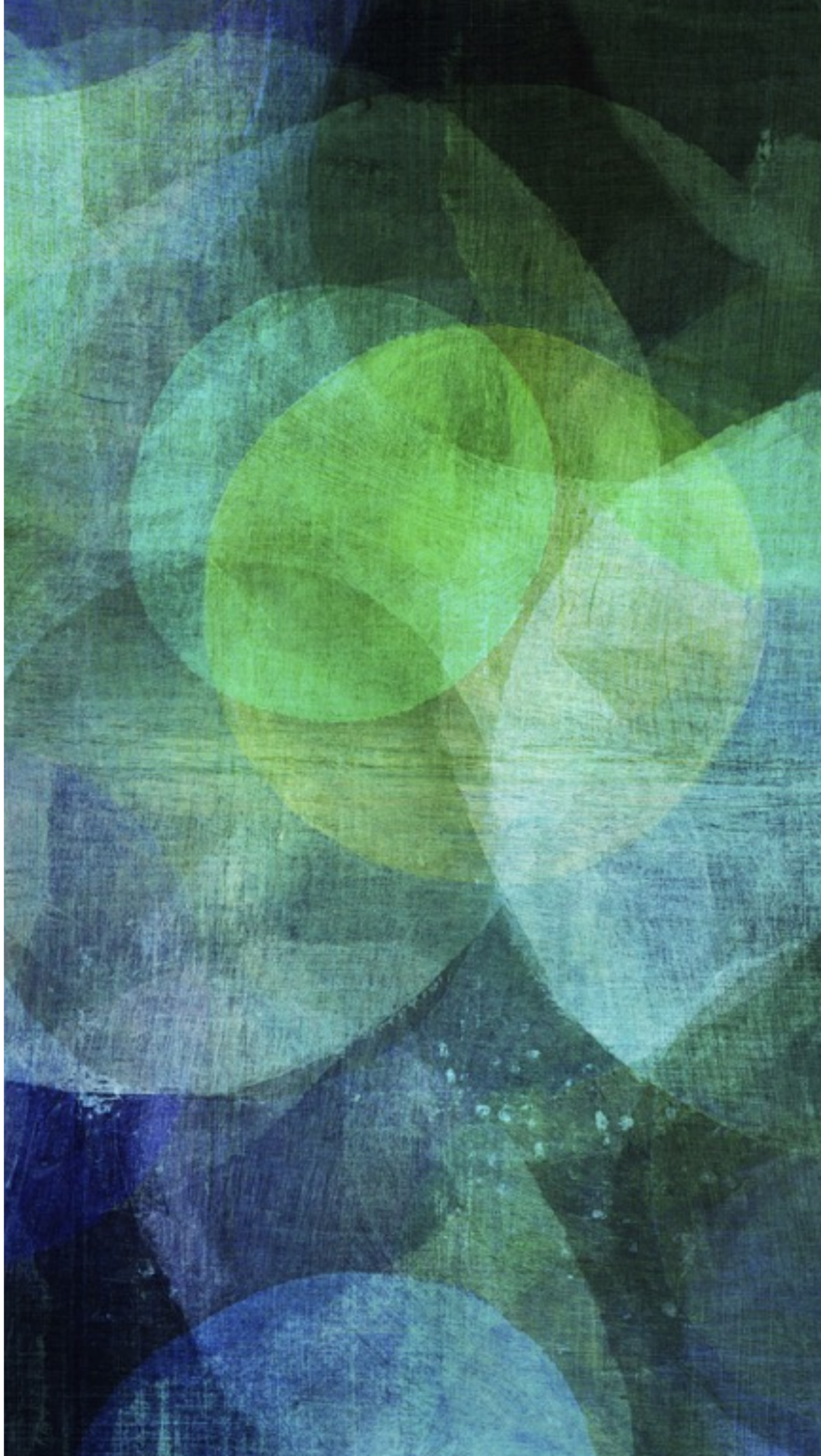
- 3 NIH grants $2,961,032

  14 GB of data, 27667 lines of code, 2007-2011.

  Homer Warner Award 2010
- **Example:** Heparin Induced Thrombocytopenia
  - BEFORE: **about 10 years of creating the rule**
  - BEFORE: **Rule definition has 5 pages**
  - BEFORE: **Every adjustment takes 3 months**
  - AFTER: **5 years of historical data (no supervision needed)**
  - AFTER: **Better performance (prediction/recall) than for the rule**
- **Large study:** 734 decisions (orders) for 40 000 cases
- **Evaluation:** 54.5% of alerts found useful
- Used by Department of Clinical Care
- **Explainability**

ONLINE DECISION-MAKING

when we want to act



Example of a graph bandit problem

movie recommendation

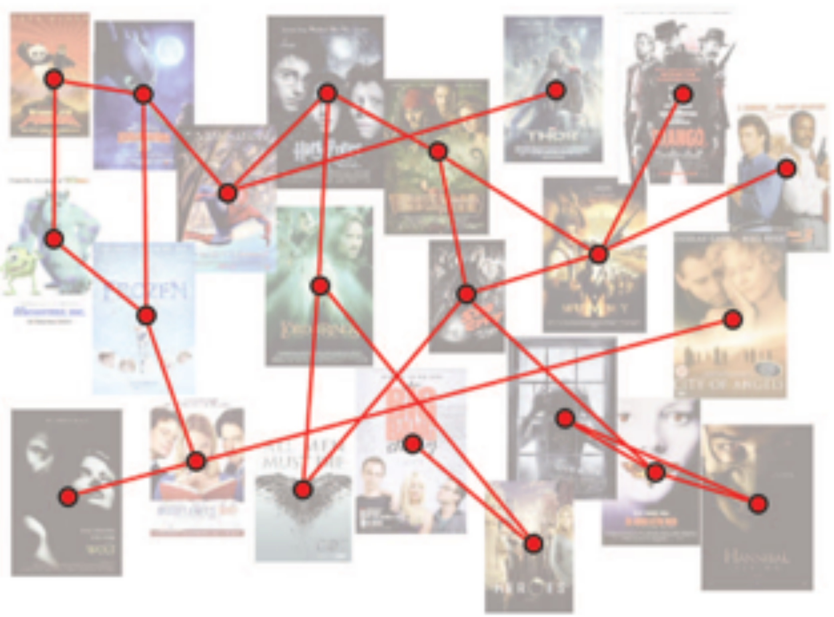
- recommend movies to a **single** user
- **goal:** maximise the sum of the ratings (minimise regret)
- good prediction after just a few steps

$$T \ll N$$

- extra information
  - ratings are **smooth** on a graph
- main question: can we learn **faster**?

GETTING REAL

Let's be lazy and ignore the structure



→

Multi-armed bandit problem!

Worst case regret (to the best fixed strategy)

$$R_T = \mathcal{O}\left(\sqrt{NT}\right)$$

#actions

#rounds

Matching lower bound (Auer, Cesa-Bianchi, Freund, Schapire 2002)

How big is N? Number of movies on **http://www.imdb.com/stats**: 5,310,913

Problem: Too many actions!

LEARNING FASTER

$$R_T = \mathcal{O}\left(\sqrt{NT}\right)$$

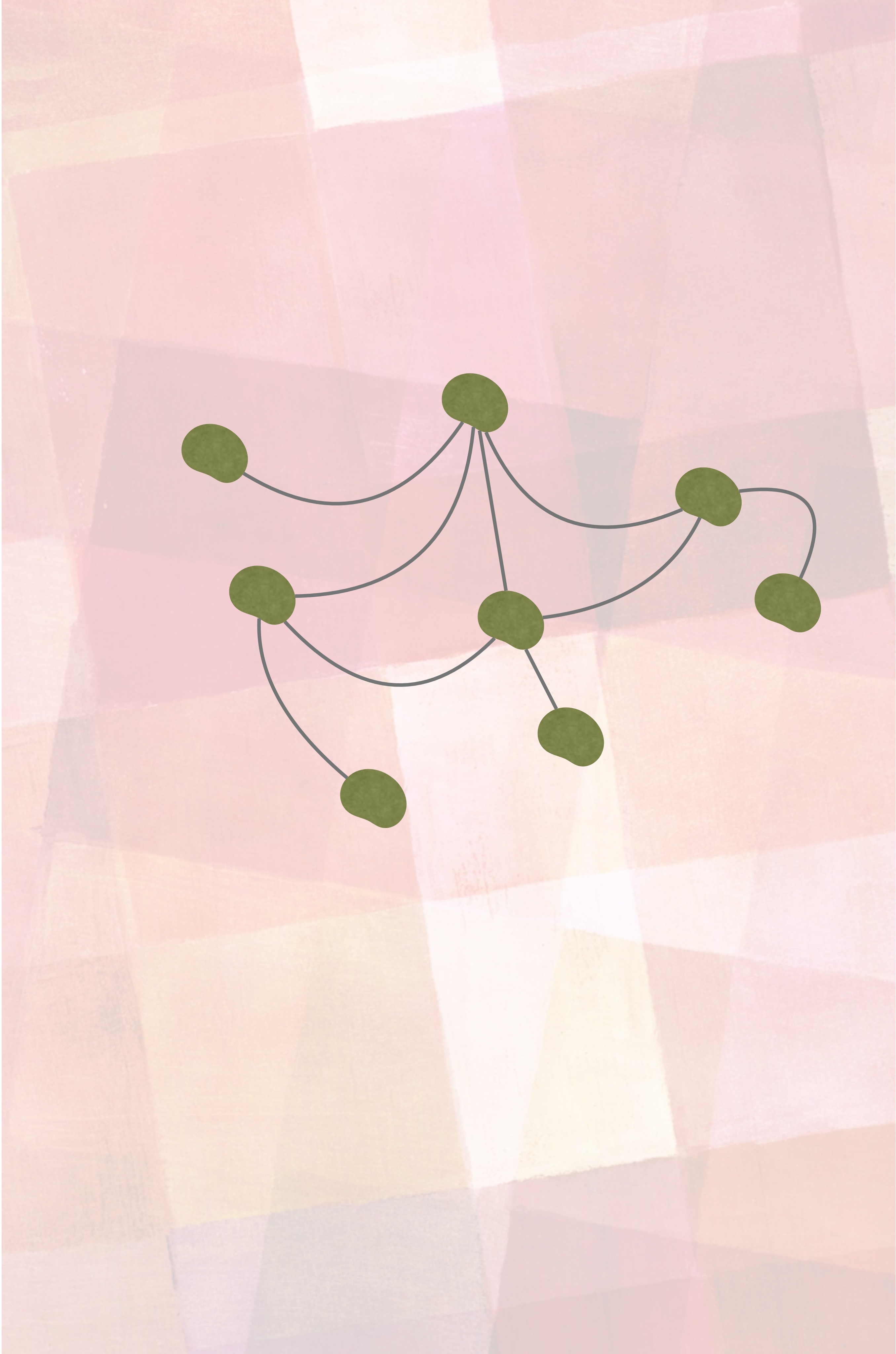
#actions

#rounds

- Arm independence is too strong and unnecessary
- Replace **N** with something much smaller
  - problem/instance/data dependent
  - example: linear design **N** to **D**

#dimensions

- **Here use Graphs to encode structure of decision making!**
  - sequential problems where **actions are nodes** on a graph
  - find strategies that replace **N** with a smaller **graph-dependent** quantity



GRAPH BANDITS: GENERAL SETUP

Every round **t** the learner

- picks a node $I_t \in [N]$
- incurs a loss ℓ_{t,I_t}
- optional feedback

The performance is total expected regret

$$R_T = \max_{i \in [N]} \mathbb{E}\left[\sum_{t=1}^{T} (\ell_{t,I_t} - \ell_{t,i})\right]$$

STRUCTURES IN ONLINE (RL/BANDIT) PROBLEMS

GRAPHS

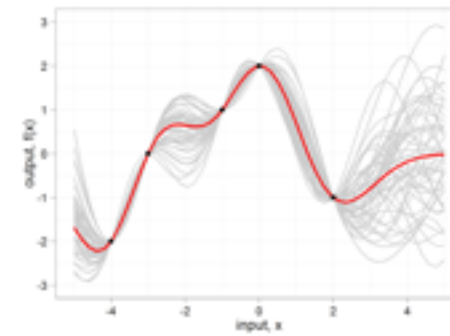
KERNELS

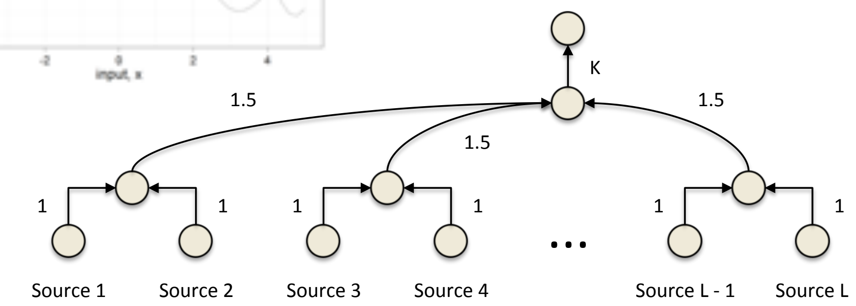
DISCOUNT FACTOR in MDPs

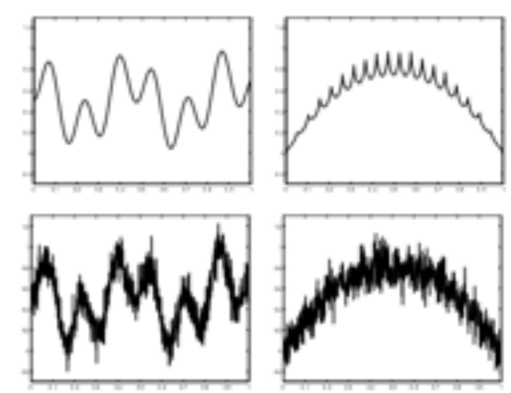
CONTINUOUS FUNCTIONS

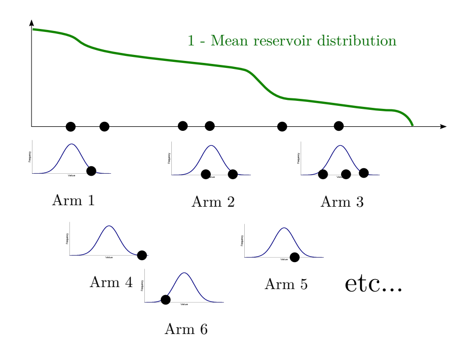
STRUCTURES WITHOUT TOPOLOGY

…









SPECIFIC GRAPH BANDIT SETTINGS

smoothness
spectral bandits
$R_T = \widetilde{\mathcal{O}}\left(d\sqrt{T \ln T}\right)$
#relevant eigenvectors

side observations
on graphs
$R_T = \widetilde{\mathcal{O}}\left(\sqrt{\overline{\alpha} T \ln N}\right)$
independence number

influence maximisation
revealing bandits
$R_T = \widetilde{\mathcal{O}}\left(\sqrt{r_* T D_*}\right)$
detectable dimension

noisy side observations
on graphs
$R_T = \widetilde{\mathcal{O}}\left(\sqrt{\overline{\alpha}^\star T \ln N}\right)$
effective independence number

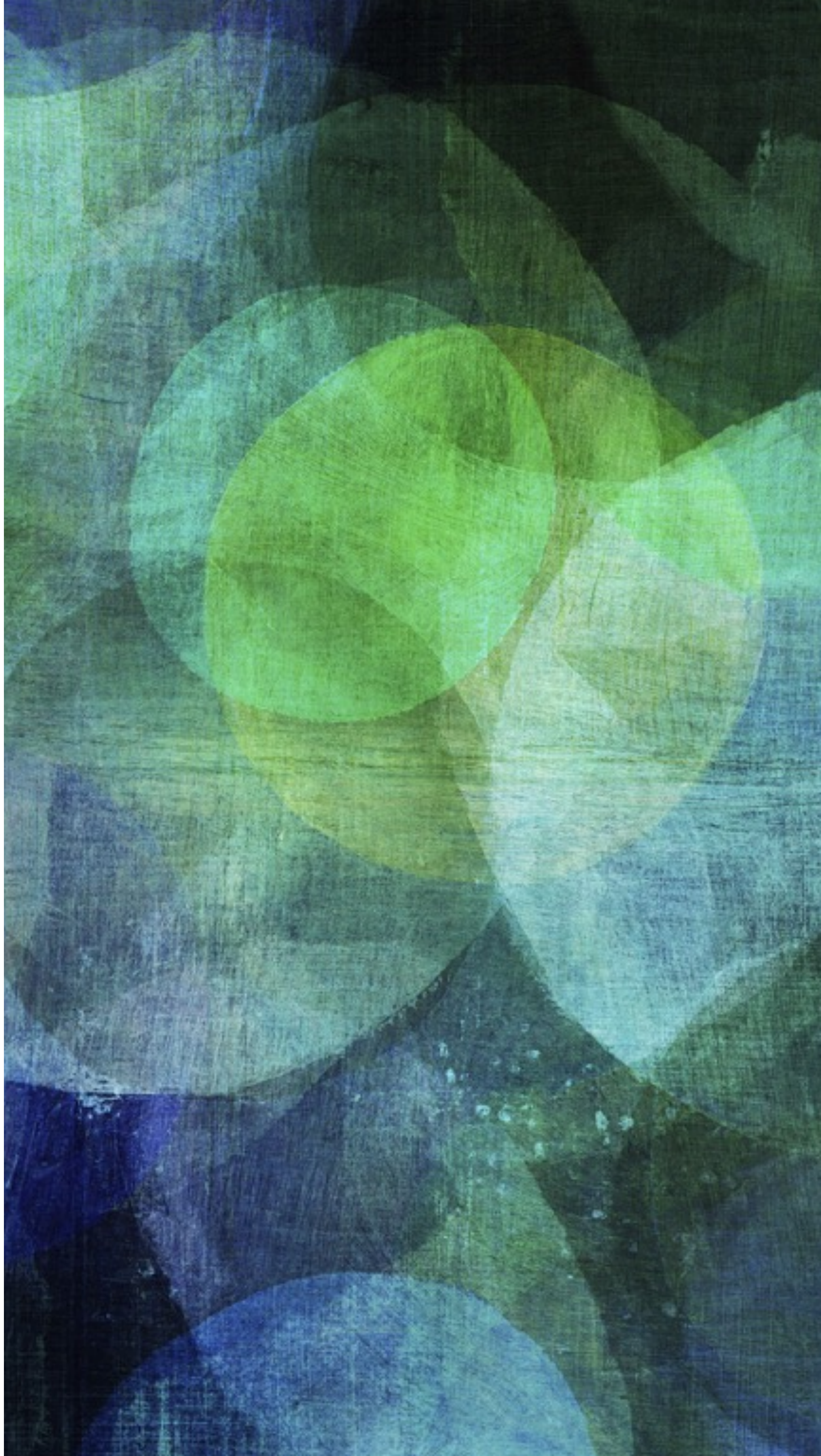
MV, Munos, Kveton, Kocák: **Spectral Bandits for Smooth Graph Functions**, ICML 2014

Kocák, MV, Munos, Agrawal: **Spectral Thompson Sampling**, AAAI 2014

Hanawal, Saligrama, MV, Munos: **Cheap Bandits**, ICML 2015

SPECTRAL BANDITS

exploiting smoothness of rewards on graphs



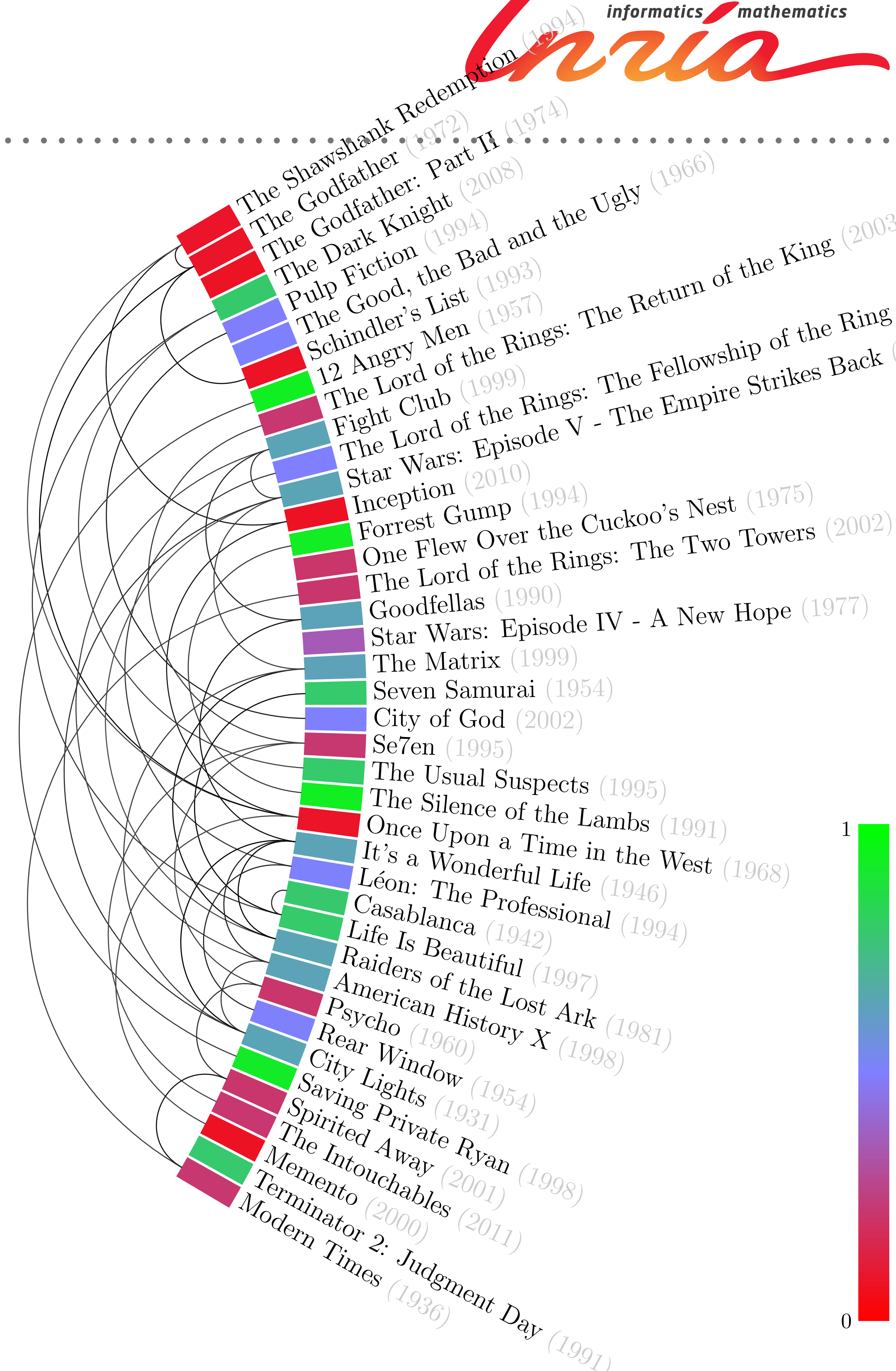
SPECTRAL BANDITS

Assumptions

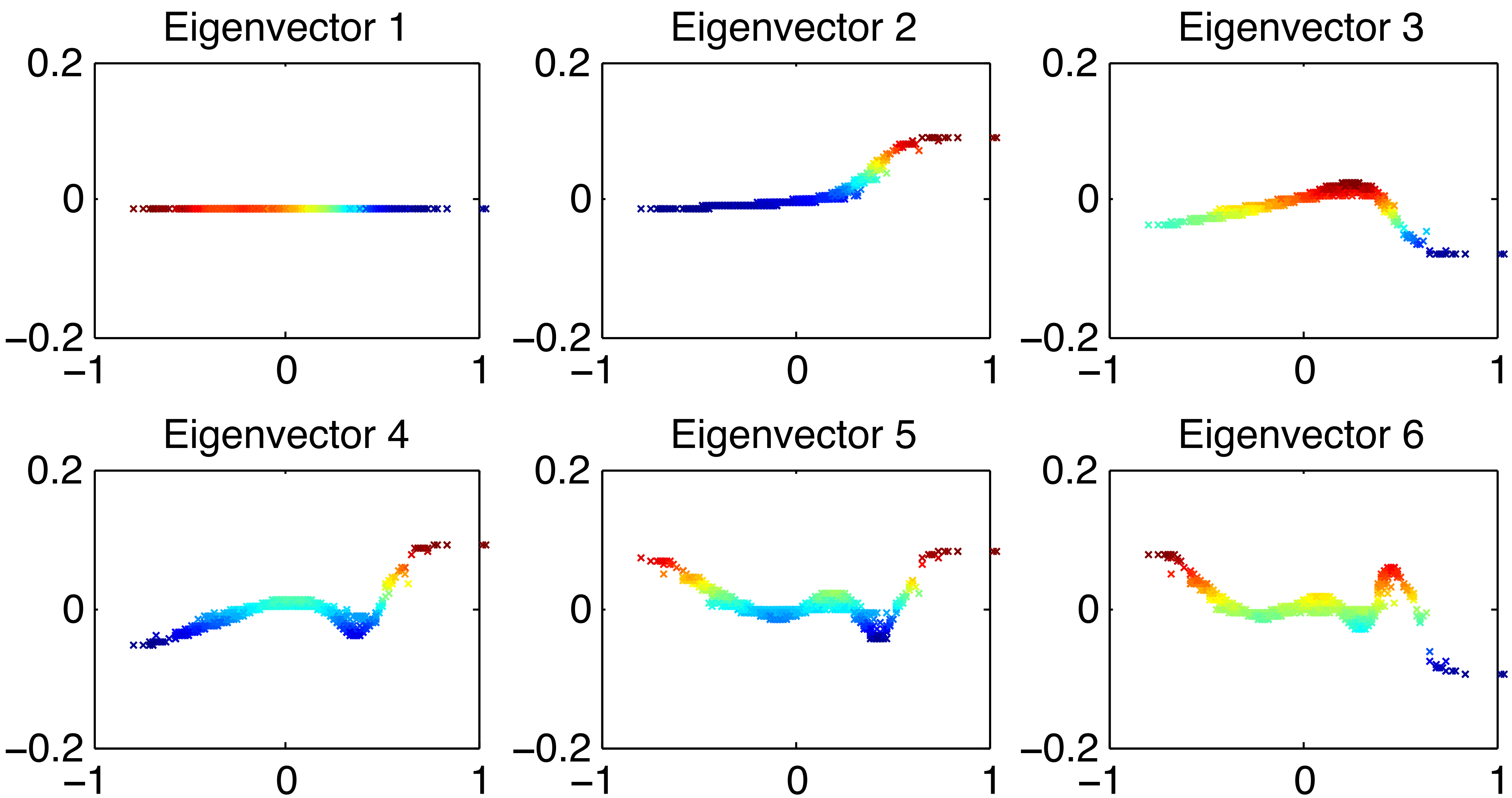
- ▶ Unknown reward function $f : V(G) \to \mathbb{R}$.
- ▶ Function f is **smooth** on a graph.
- ▶ Neighboring movies $\Rightarrow$ similar preferences.
- ▶ Similar preferences $\not\Rightarrow$ neighboring movies.

Desiderata

An algorithm useful in the case $T \ll N$!



FLIXSTER DATA



Eigenvectors from the Flixster data corresponding to the smallest few eigenvalues of the graph Laplacian projected onto the first principal component of data. Colors indicate the values.

LINEAR VS. SPECTRAL BANDITS

- **Linear bandit algorithms**
  - **LinUCB** (Li et al., 2010)
    - Regret bound $\approx D\sqrt{T \ln T}$
  - **LinearTS** (Agrawal and Goyal, 2013)
    - Regret bound $\approx D\sqrt{T \ln N}$

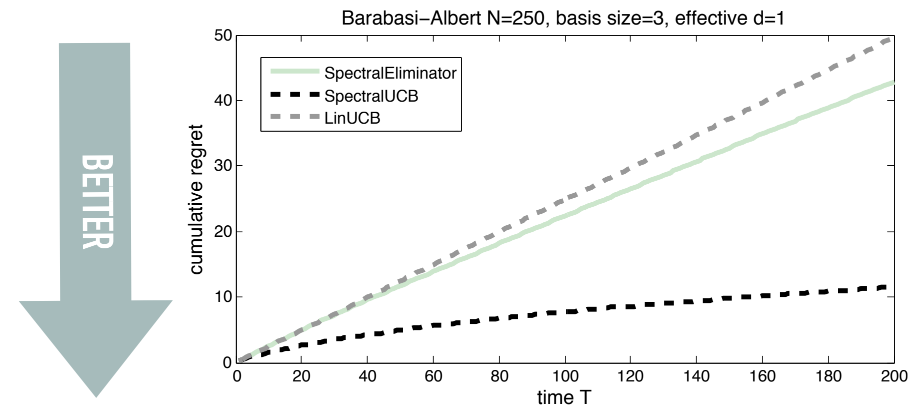
Note: D is ambient dimension, in our case N, length of x_i.
Number of actions, e.g., all possible movies $\rightarrow$ **HUGE!**

- **Spectral bandit algorithms**
  - **SpectralUCB** (Valko et al., ICML 2014)
    - Regret bound $\approx d\sqrt{T \ln T}$
    - Operations per step: $D^2 N$
  - **SpectralTS** (Kocák et al., AAAI 2014)
    - Regret bound $\approx d\sqrt{T \ln N}$
    - Operations per step: $D^2 + DN$

Note: d is effective dimension, usually much smaller than D.

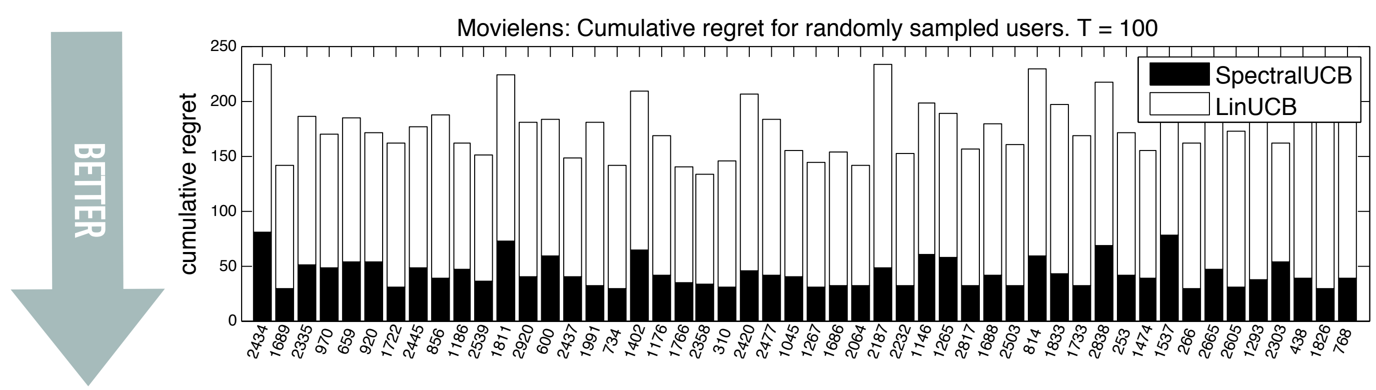
SPECTRALUCB REGRET BOUND

- d: Effective dimension.
- λ: Minimal eigenvalue of $\mathbf{\Lambda} = \mathbf{\Lambda_L} + \lambda\mathbf{I}$.
- C: Smoothness upper bound, $\|\boldsymbol{\alpha}^*\|_{\mathbf{\Lambda}} \leq C$.
- $\mathbf{x}_i^\mathsf{T}\boldsymbol{\alpha}^* \in [-1, 1]$ for all i.



The **cumulative regret** R_T of **SpectralUCB** is with probability $1 - \delta$ bounded as

$$R_T \leq \left(8R\sqrt{d \ln \frac{\lambda + T}{\lambda} + 2 \ln \frac{1}{\delta}} + 4C + 4\right) \sqrt{dT \ln \frac{\lambda + T}{\lambda}}.$$



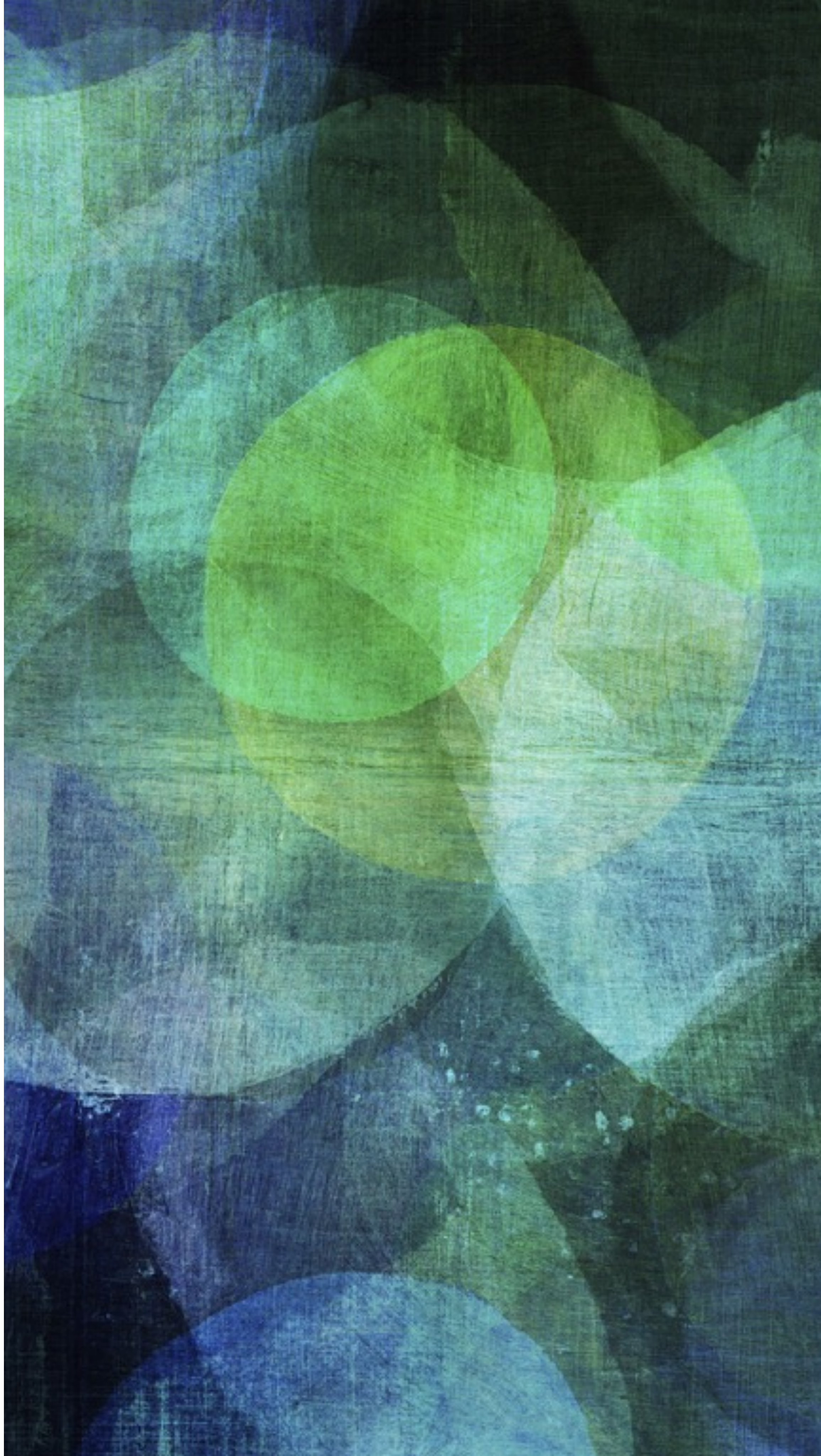
Kocák, Neu, MV, Munos: **Efficient learning by implicit exploration in bandit problems with side observations**, NIPS 2014

Kocák, Neu, MV: **Online learning with Erdos-Rényi side-observation graphs** UAI 2016

Kocák, Neu, MV: **Online learning with noisy side observations**, AISTATS 2016

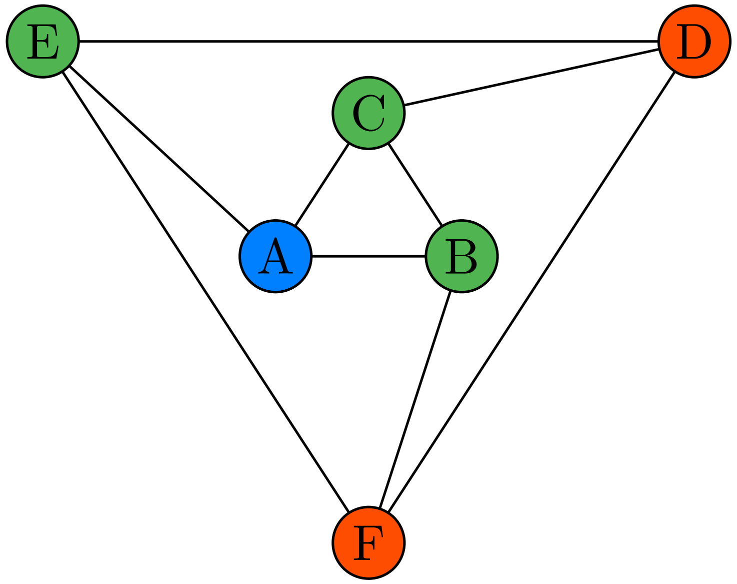
GRAPH BANDITS WITH SIDE OBSERVATIONS

exploiting free observations from neighbouring nodes

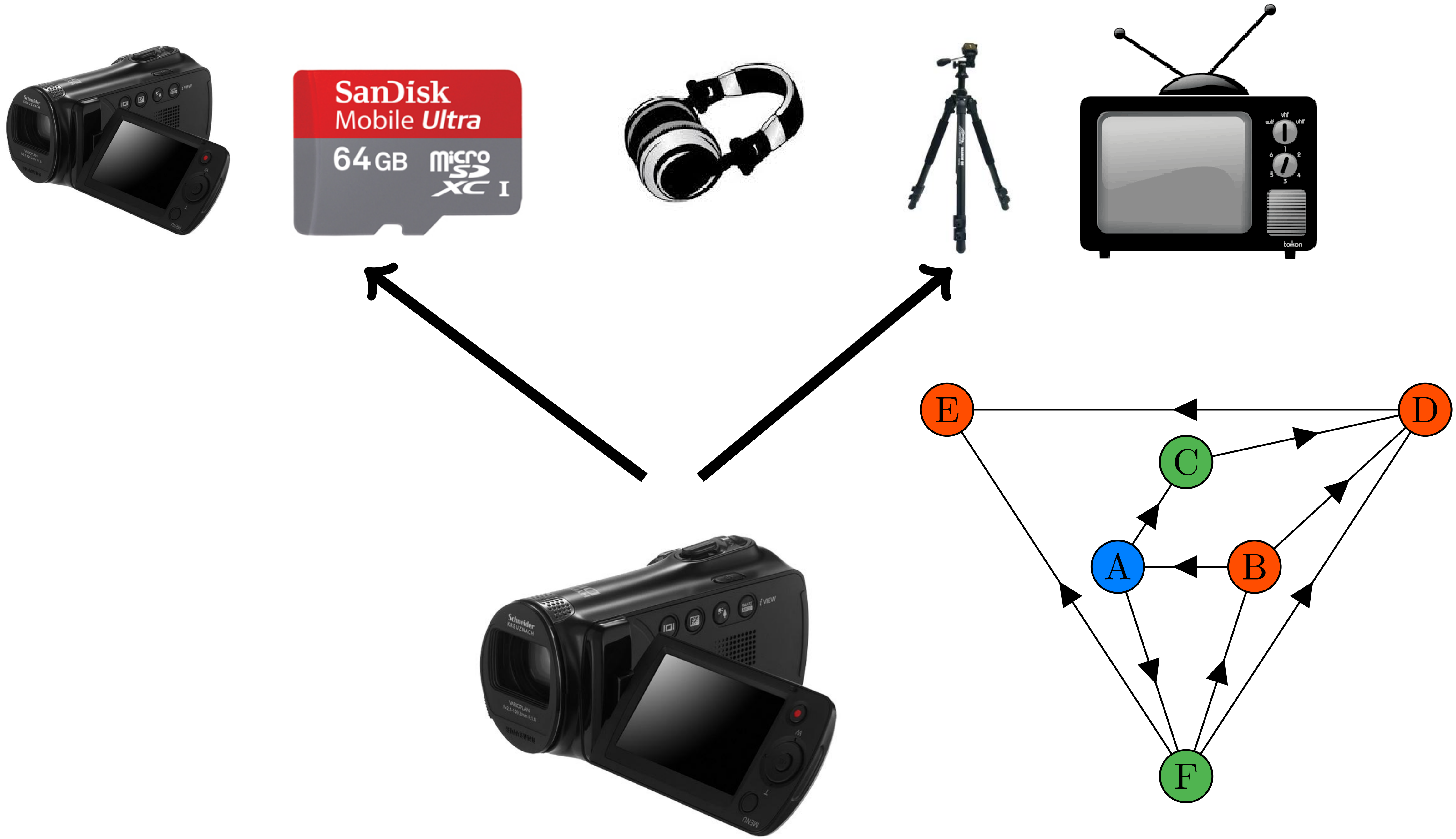


SIDE OBSERVATIONS: UNDIRECTED

Example 1: undirected observations



Example 2: Directed observation



SIDE OBSERVATIONS – AN INTERMEDIATE GAME

Full-information

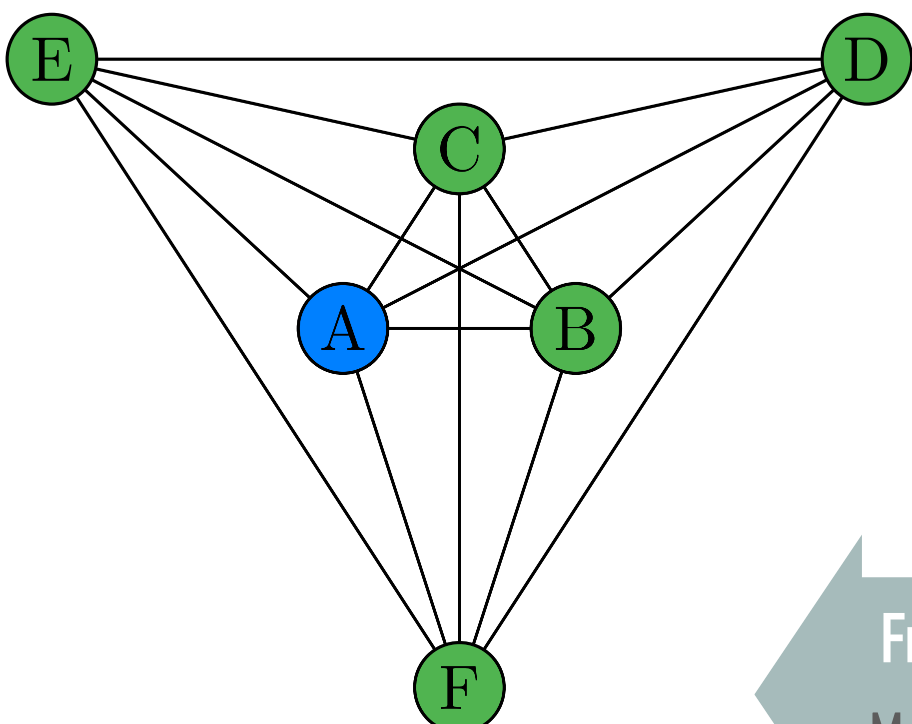
- observe losses of **all** actions
- example: Hedge

$$R_T = \widetilde{\mathcal{O}}(\sqrt{T})$$

Bandits

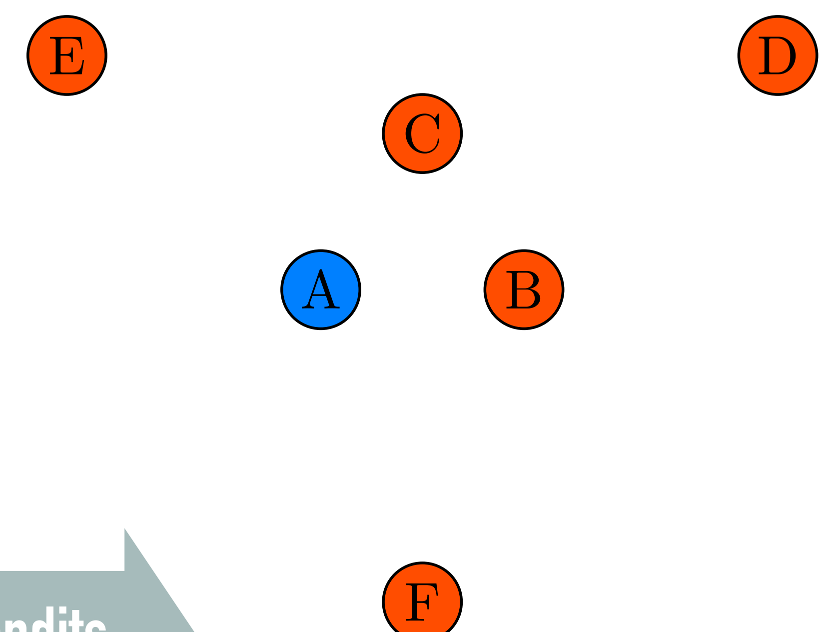
- observe losses of **the chosen** action
- example: EXP3

$$R_T = \widetilde{\mathcal{O}}(\sqrt{NT})$$



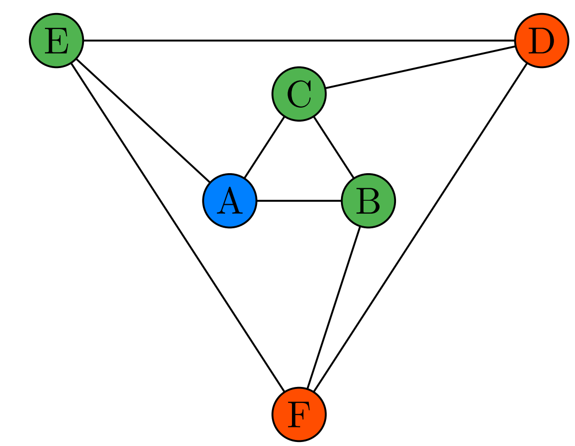
From Experts to Bandits

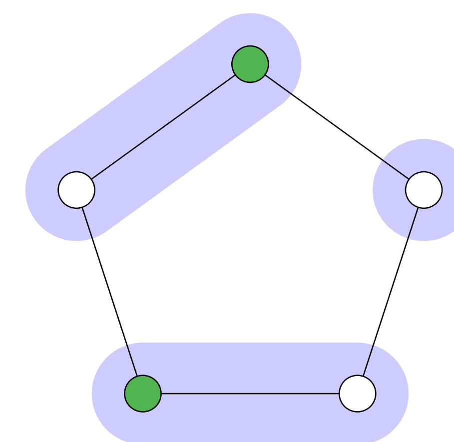
Mannor and Shamir 2011

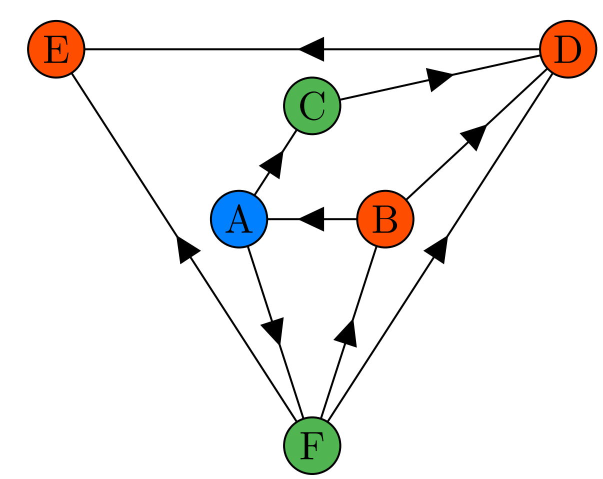


KNOWLEDGE OF OBSERVATION GRAPHS

- ELP (Mannor and Shamir 2011)
  - **EXP3** - with **"LP balanced exploration"**
  - undirected $O(\sqrt{(\alpha T)})$ ✅ – needs to know G_t
  - directed case $O(\sqrt{(cT)})$ – needs to know G_t
- EXP3-SET (Alon, Cesa-Bianchi, Gentile, Mansour, 2013)
  - undirected $O(\sqrt{(\alpha T)})$ ✅ does not need to know G_t ✅
- EXP3-DOM (Alon, Cesa-Bianchi, Gentile, Mansour, 2013)
  - directed $O(\sqrt{(\alpha T)})$ ✅ – need to know G_t
  - **calculates dominating set**







EXP3-IX: IMPLICIT EXPLORATION

Algorithm 1 EXP3-IX

Input: Set of actions $\mathcal{S} = [d]$,
parameters $\gamma_t \in (0, 1)$, $\eta_t > 0$ for $t \in [T]$.
for $t = 1$ **to** T **do**
$w_{t,i} \leftarrow (1/d) \exp\left(-\eta_t \widehat{L}_{t-1,i}\right)$ for $i \in [d]$
An adversary privately chooses losses $\ell_{t,i}$ for $i \in [d]$ and generates a graph G_t
$W_t \leftarrow \sum_{i=1}^{d} w_{t,i}$
$p_{t,i} \leftarrow w_{t,i}/W_t$
Choose $I_t \sim \boldsymbol{p}_t = (p_{t,1}, \ldots, p_{t,d})$
Observe graph G_t
Observe pairs $\{i, \ell_{t,i}\}$ for $(I_t \rightarrow i) \in G_t$
$o_{t,i} \leftarrow \sum_{(j \rightarrow i) \in G_t} p_{t,j}$ for $i \in [d]$
$\hat{\ell}_{t,i} \leftarrow \frac{\ell_{t,i}}{o_{t,i} + \gamma_t} \mathbb{1}_{\{(I_t \rightarrow i) \in G_t\}}$ for $i \in [d]$
end for

Benefits of the **implicit exploration**

- **no need to know the graph before**
- no need to estimate dominating set
- no need for doubling trick
- no need for aggregation

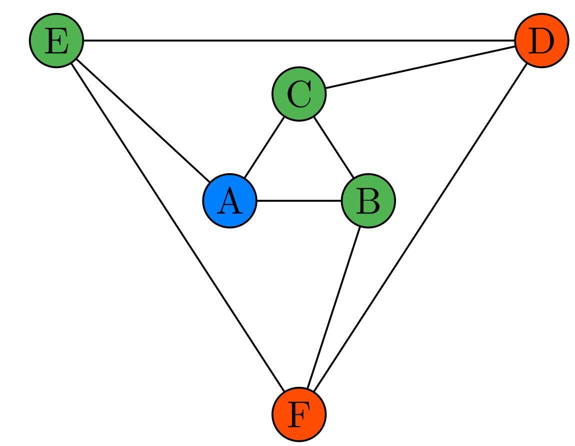
$$R_T = \widetilde{\mathcal{O}}\left(\sqrt{\alpha T \ln N}\right)$$

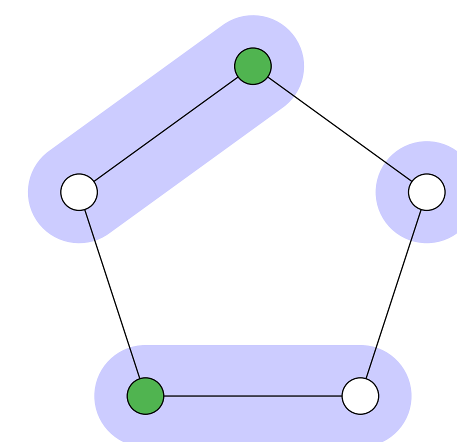
Optimistic bias for the loss estimates

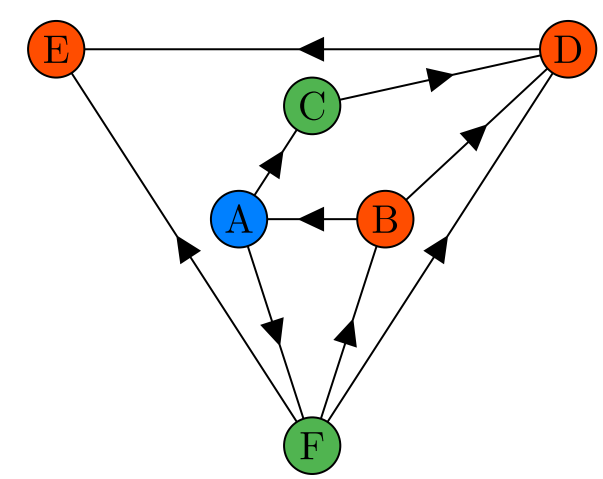
$$\mathbb{E}[\hat{\ell}_{t,i}] = \frac{\ell_{t,i}}{o_{t,i} + \gamma} o_{t,i} + 0(1 - o_{t,i}) = \ell_{t,i} - \ell_{t,i} \frac{\gamma}{o_{t,i} + \gamma} \leq \ell_{t,i}$$

FOLLOW UPS

- EXP3-IX (Kocák, Neu, MV, Munos, 2014)
  - directed $O(\sqrt{(\alpha T)})$ ✅ does not need to know G_t ✅
- EXP3.G (Alon, Cesa-Bianchi, Dekel, Koren, 2015)
  - directed $O(\sqrt{(\alpha T)})$ ✅ does not need to know G_t ✅
  - mixes uniform distribution
  - more general algorithm for settings **beyond bandits**
  - high-probability bound
- Neu 2015: high-probability bound for EXP3-IX
- TextBook: **Bandit Algorithms** T. Lattimore & Cs. Szepesvári

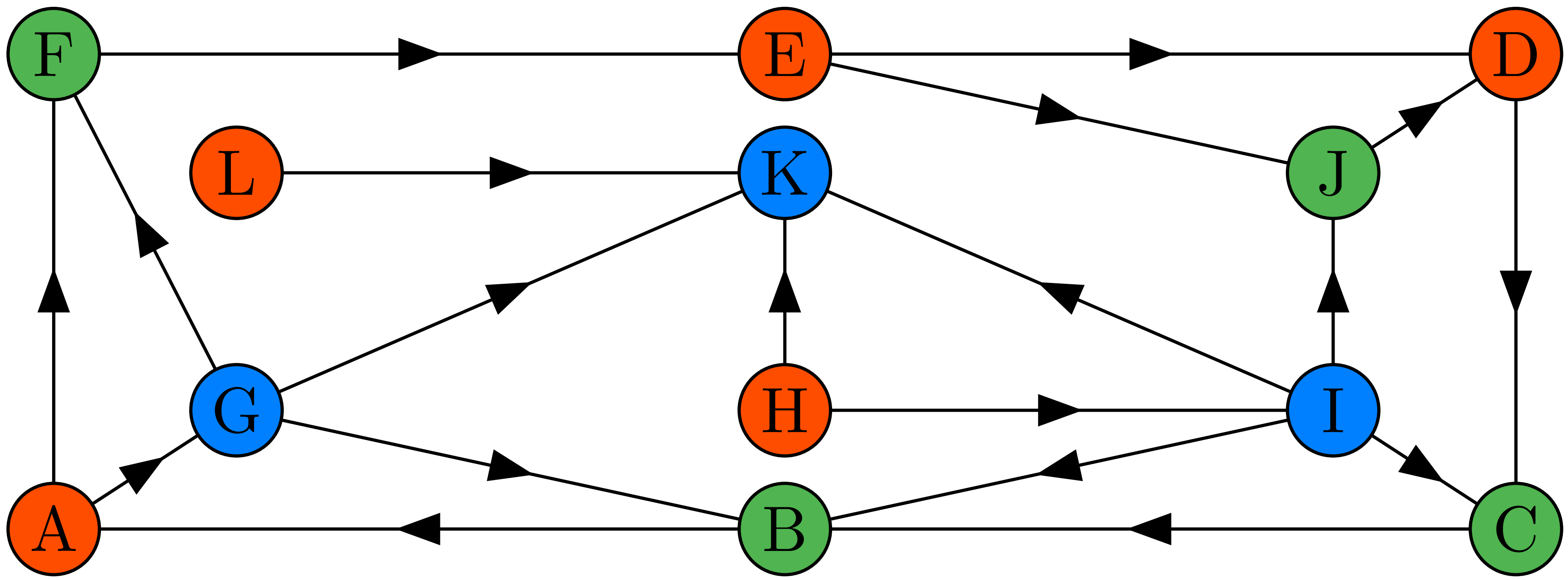






EXTENSION: COMPLEX GRAPH ACTIONS

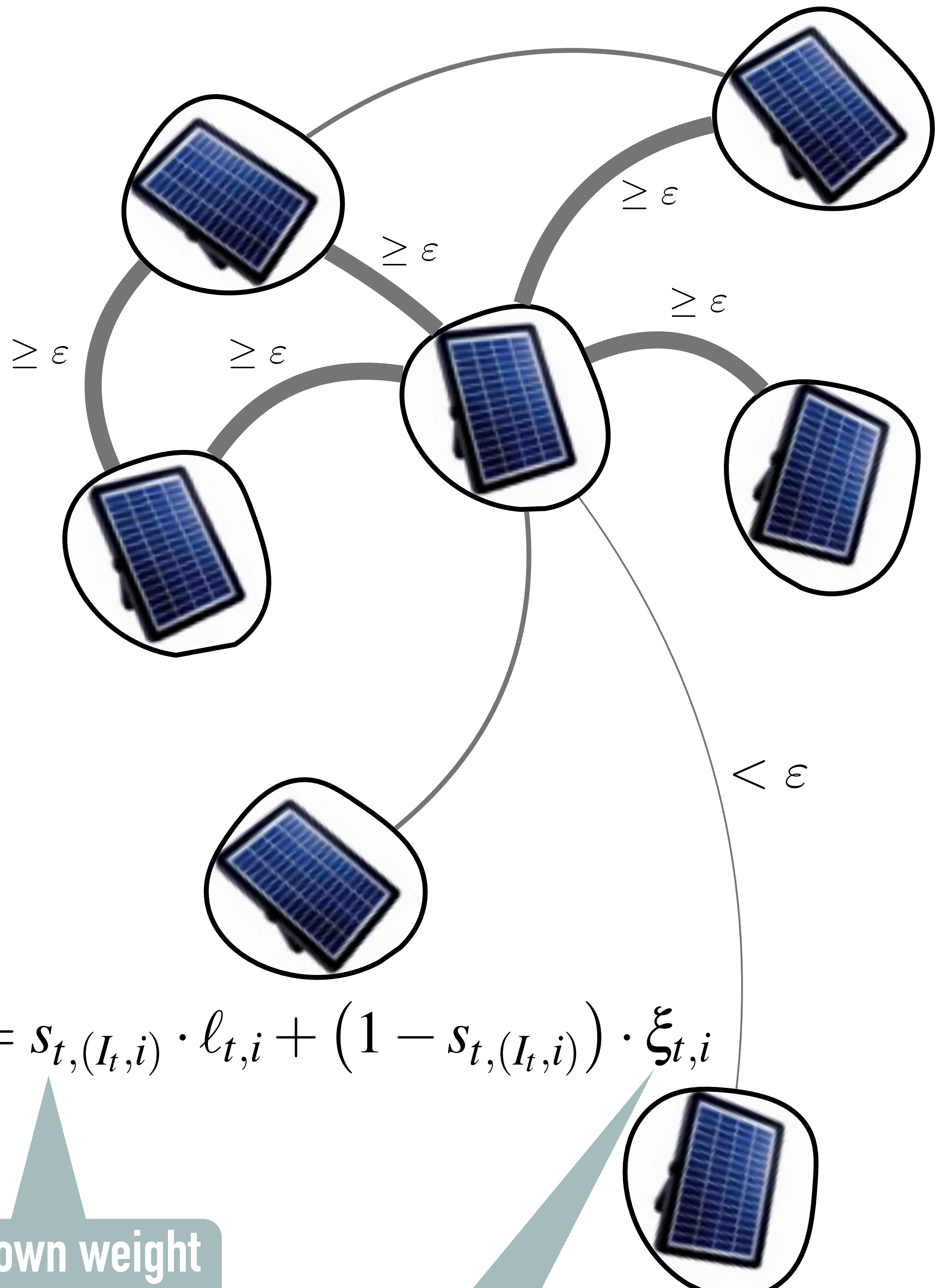
Example: online shortest path semi-bandits with observing traffic on the side streets



- Play action $\mathbf{V}_t \in S \subset \{0,1\}^N$, $\|\mathbf{v}\|_1 \leq m$ from all $\mathbf{v} \in S$
- Obtain losses $\mathbf{V}_t^\mathsf{T} \ell_t$
- Observe additional losses according to the graph

$$R_T = \widetilde{\mathcal{O}}\left(m^{3/2}\sqrt{\sum_{t=1}^{T} \alpha_t}\right) = \widetilde{\mathcal{O}}\left(m^{3/2}\sqrt{\overline{\alpha} T}\right)$$

EXTENSION: NOISY SIDE OBSERVATIONS



$$c_{t,i} = s_{t,(I_t,i)} \cdot \ell_{t,i} + \left(1 - s_{t,(I_t,i)}\right) \cdot \xi_{t,i}$$

known weight

zero-mean noise

Want: only reliable information!

1) If we know the perfect cutoff ε

- reliable: use as exact
- unreliable: rubbish

then we can improve over pure bandit setting!

2) Treating noisy observation induces bias

What can we hope for?

$$\tilde{\mathcal{O}}\left(\sqrt{1T}\right) \leq \tilde{\mathcal{O}}\left(\sqrt{\overline{\alpha}^\star T}\right) \leq \tilde{\mathcal{O}}\left(\sqrt{NT}\right)$$

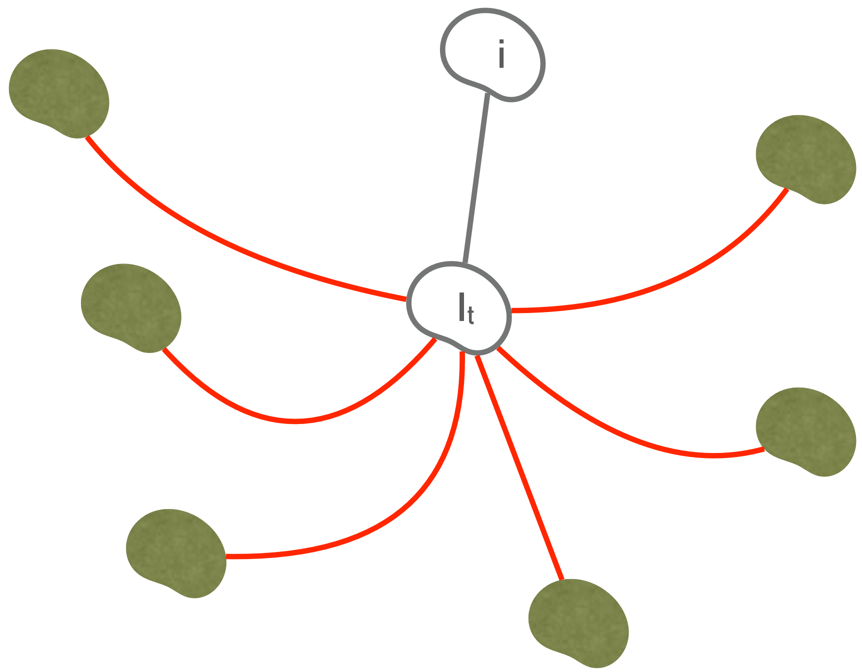
effective independence number

Can we learn without knowing either ε or α* ?

NEW DIRECTIONS: UNKNOWN GRAPHS!

- **Learning on the graph while learning the graph?**
  - **most of algorithms require (some) knowledge of the graph**
  - not always available to the learner
- Question: Can we learn faster without knowing the graphs?
  - example: social network provider has little incentive to reveal the graphs to advertisers
- Answer: **Cohen, Hazan, and Koren:** Online learning with **feedback** graphs without the graphs (ICML 2016)
  - **NO!** (in general we cannot, but possible in the stochastic case)
- NEXT in the talk: **examples where we can do something!**
  - **Erdös-Rényi side observation graphs**
  - **Influence Maximisation**

EXTENSION: ERDÖS-RÉNYI GRAPHS



- N-2 samples from Bernoulli(r_t) ... R(k)
- N-2 samples from p_{ti} ... P(k)
- O'(k) = P(k) + (1-P(k))R(k)
- G_{ti} = min{k : O'(k) = 1} U {N-1}
- E[G_{ti}] ≈ 1/(p_{ti} + (1-p_{ti})r_t)

$$\widehat{\ell}_{t,i} = G_{t,i} O_{t,i} \ell_{t,i}$$

is loss of i observed?

true loss

$$\widehat{\ell}^{\star}_{t,i} = \frac{O_{t,i}\ell_{t,i}}{p_{t,i} + (1 - p_{t,i})r_t}$$

probability of picking i

probability of side observation

If r_t ≥ (log T)/(2N-2) then

$$\mathcal{O}\left(\sqrt{\log N \sum_{t=1}^{T} \frac{1}{r_t}}\right)$$

Lower bound (Alon et al. 2013) $\Omega(\sqrt{T/r})$

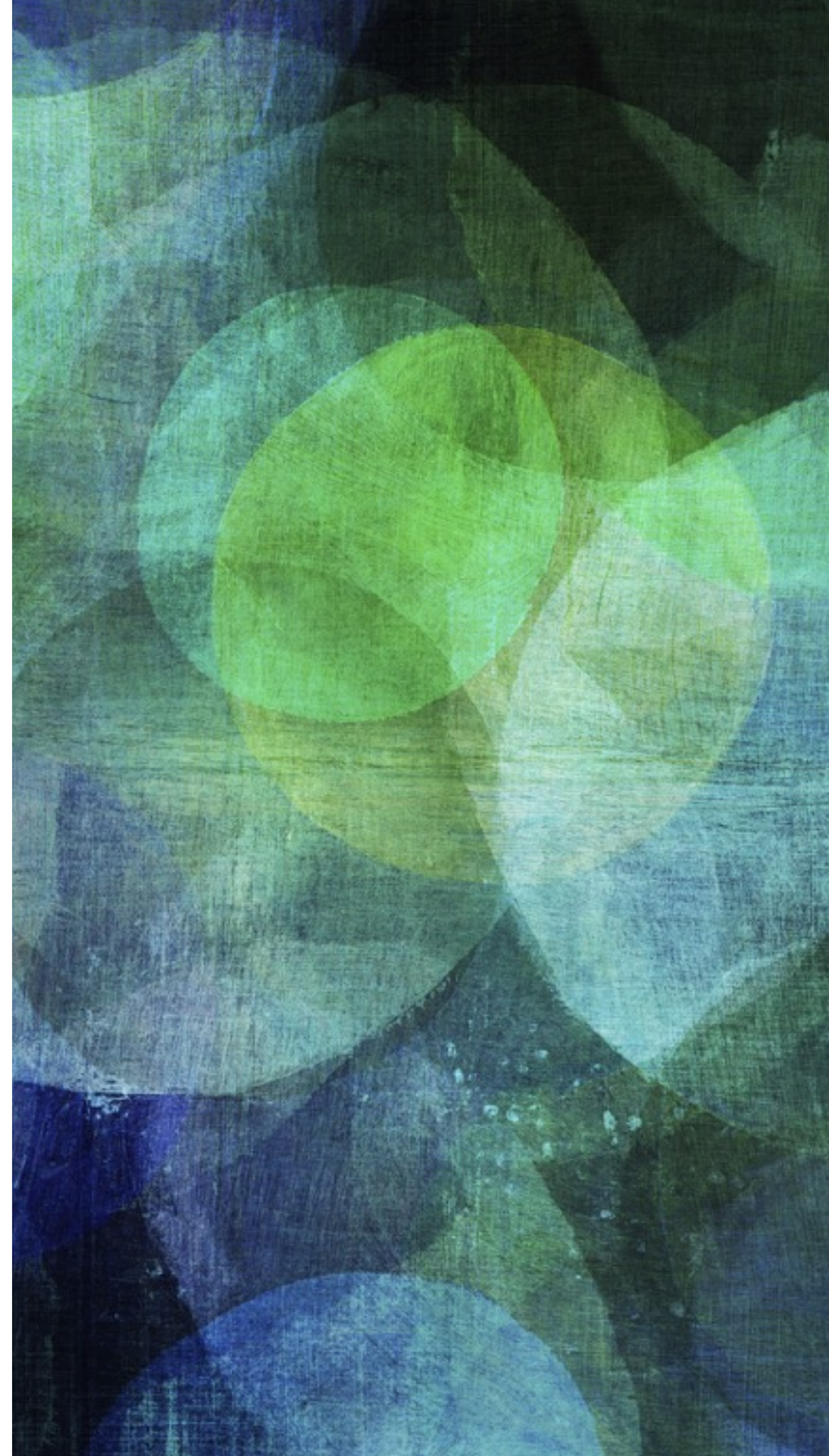
Get rid of r_t ≥ (log T)/(2N-2)?

Carpentier, MV: **Revealing Graph Bandits for Maximising Local Influence**, AISTATS 2016

Wen, Kveton, MV: **Influence Maximization with Semi-Bandit Feedback,** NIPS 2017

INFLUENCE MAXIMISATION

looking for the influential nodes **while** exploring the graph

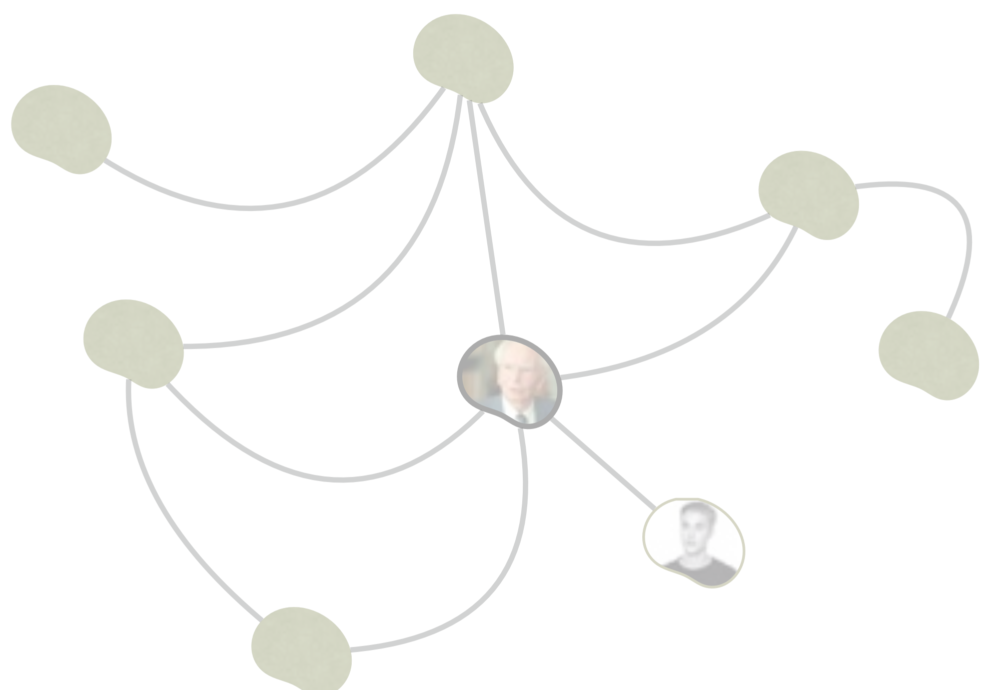


REVEALING BANDITS FOR LOCAL INFLUENCE

Unknown $\mathbf{M} = (p_{i,j})_{i,j}$ symmetric matrix of influences

In each time step $t = 1, \ldots, T$

- learners picks a node k_t
- set $S_{k_t,t}$ of influenced nodes is *revealed*



Select influential people = Find the strategy maximising

$$L_T = \sum_{t=1}^{T} |S_{k_t,t}|$$

The number of expected influences of node **k** is by definition

$$r_k = \mathbb{E}\left[|S_{k,t}|\right] = \sum_{j \leq N} p_{k,j}$$

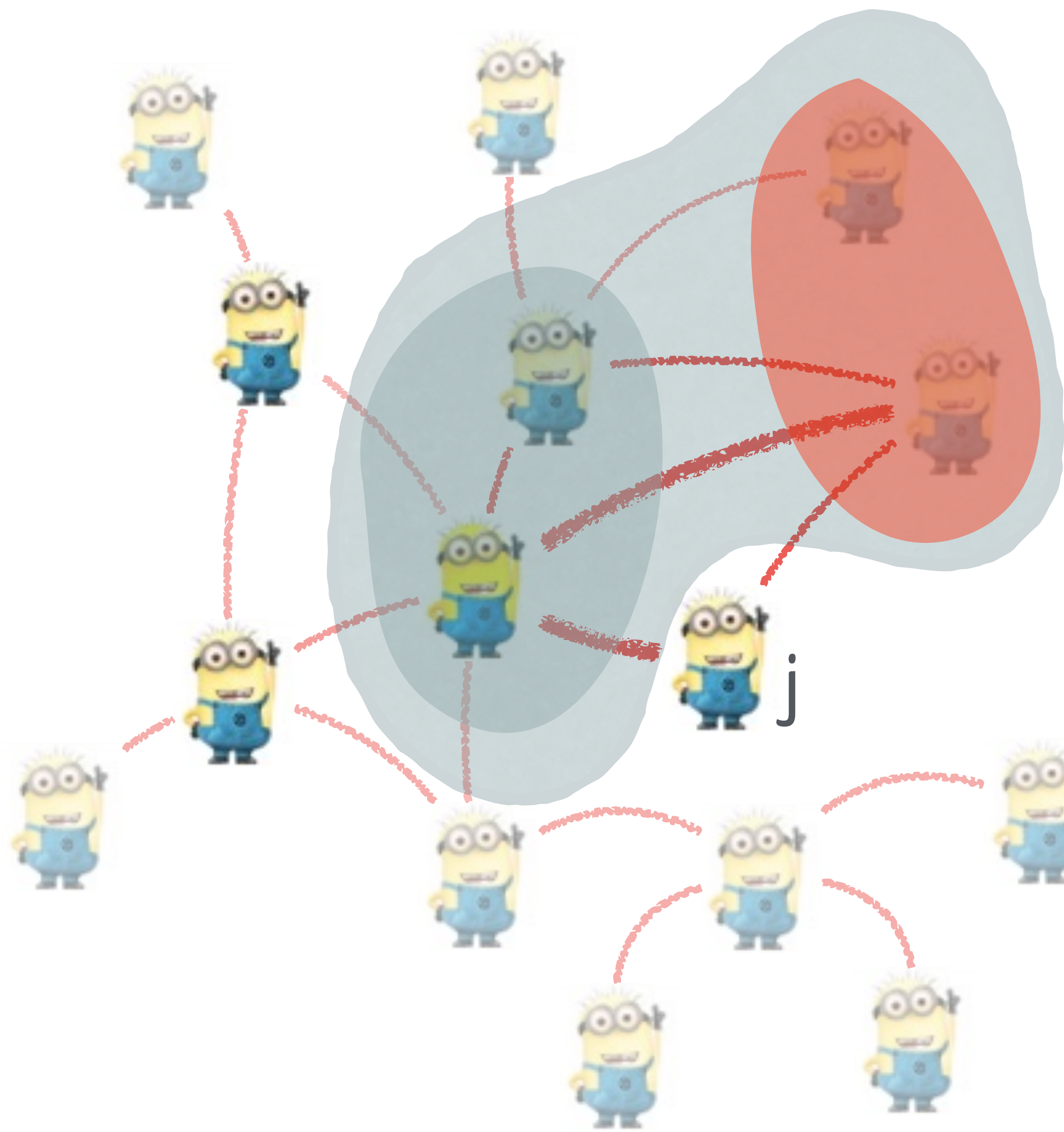
Oracle strategy always selects the best

$$k^* = \arg\max_k \mathbb{E}\left[\sum_{t=1}^{T} |S_{k,t}|\right] = \arg\max_k T r_k$$

Expected regret of any adaptive, non-oracle strategy unaware of **M**

$$\mathbb{E}\left[R_T\right] = \mathbb{E}\left[L_T^*\right] - \mathbb{E}\left[L_T\right]$$

GLOBAL DIFFUSION PROCESS OF A MARKOV CHAIN



- Sets of progressive diffusion
  - modeling diffusion steps
- Random stopping time
  - but bounded
- Topological ordering

$$\mathcal{S}^0 \triangleq \mathcal{S}_t$$

$$\mathcal{S}^{\tau+1} \triangleq \left\{ u_2 \in \mathcal{V}_{\mathcal{S}_t,v} : u_2 \notin \cup_{\tau'=0}^{\tau} \mathcal{S}^{\tau'} \text{ and } \exists e = (u_1, u_2) \in \mathcal{E}_{\mathcal{S}_t,v} \text{ s.t. } u_1 \in \mathcal{S}^{\tau} \text{ and } \mathbf{w}(e) = 1 \right\}$$

EMPIRICAL RESULTS

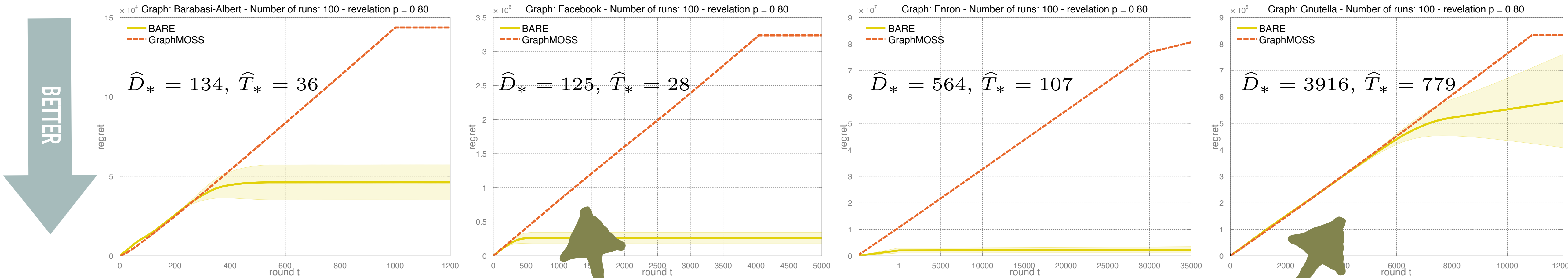
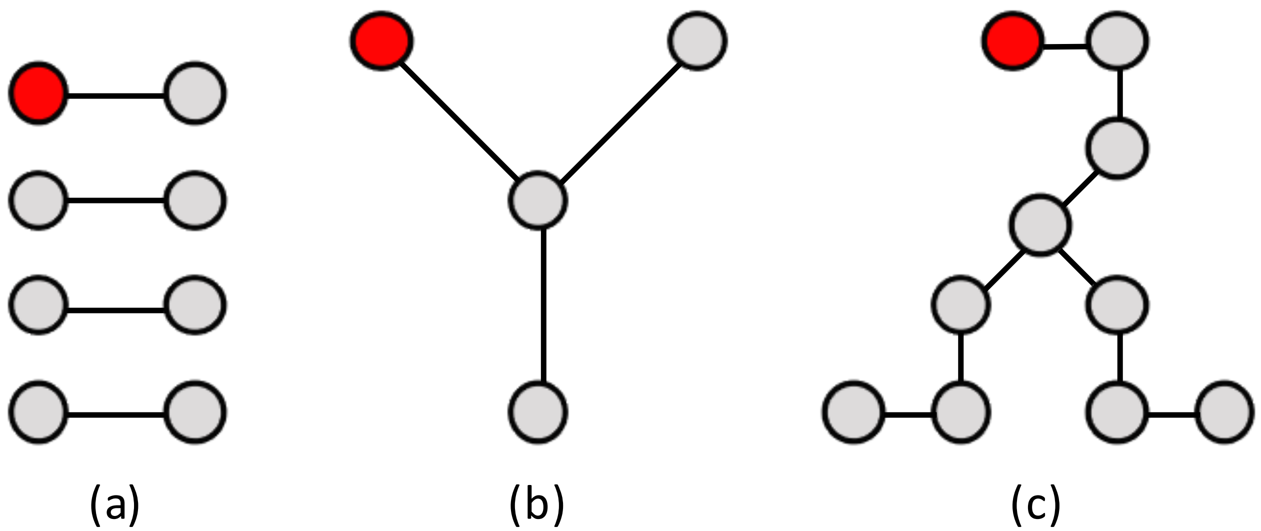


Figure 1: *Left*: Barabási-Albert. *Middle left*: Facebook. *Middle right*: Enron. *Right*: Gnutella.

Enron and Facebook **vs.** Gnutella (decentralised)

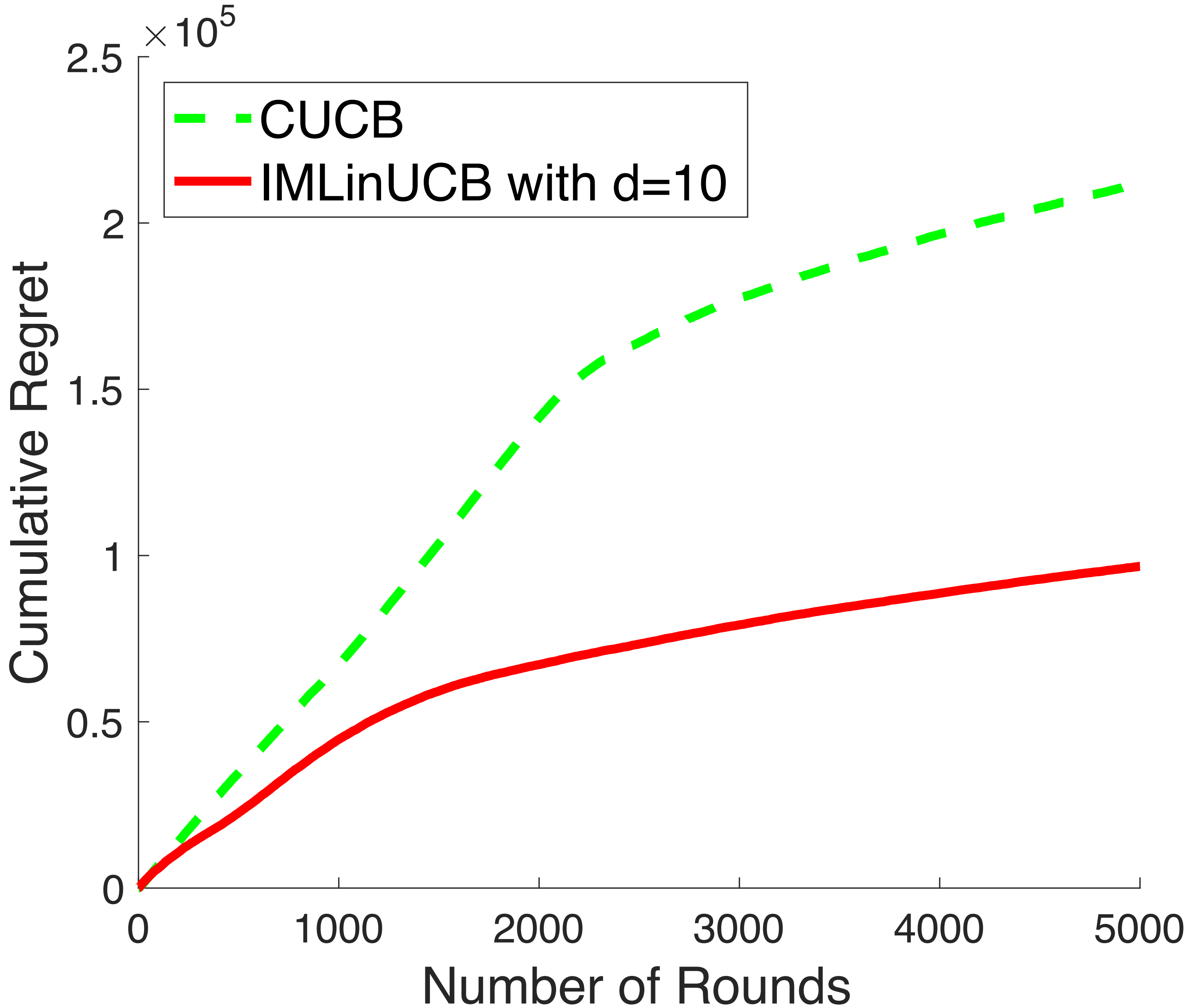
GLOBAL WORST-CASE BOUNDS



topology	$C_{\mathcal{G}}$ (worst-case C_*)	$R^{\alpha\gamma}(n)$ for general $\mathbf{X}$	$R^{\alpha\gamma}(n)$ for $\mathbf{X} = \mathbf{I}$
bar graph	$\mathcal{O}(\sqrt{K})$	$\widetilde{\mathcal{O}}\left(dK\sqrt{n}/(\alpha\gamma)\right)$	$\widetilde{\mathcal{O}}\left(L\sqrt{Kn}/(\alpha\gamma)\right)$
star graph	$\mathcal{O}(L\sqrt{K})$	$\widetilde{\mathcal{O}}\left(dL^{\frac{3}{2}}\sqrt{Kn}/(\alpha\gamma)\right)$	$\widetilde{\mathcal{O}}\left(L^2\sqrt{Kn}/(\alpha\gamma)\right)$
ray graph	$\mathcal{O}(L^{\frac{5}{4}}\sqrt{K})$	$\widetilde{\mathcal{O}}\left(dL^{\frac{7}{4}}\sqrt{Kn}/(\alpha\gamma)\right)$	$\widetilde{\mathcal{O}}\left(L^{\frac{9}{4}}\sqrt{Kn}/(\alpha\gamma)\right)$
tree graph	$\mathcal{O}(L^{\frac{3}{2}})$	$\widetilde{\mathcal{O}}\left(dL^2\sqrt{n}/(\alpha\gamma)\right)$	$\widetilde{\mathcal{O}}\left(L^{\frac{5}{2}}\sqrt{n}/(\alpha\gamma)\right)$
grid graph	$\mathcal{O}(L^{\frac{3}{2}})$	$\widetilde{\mathcal{O}}\left(dL^2\sqrt{n}/(\alpha\gamma)\right)$	$\widetilde{\mathcal{O}}\left(L^{\frac{5}{2}}\sqrt{n}/(\alpha\gamma)\right)$
complete graph	$\mathcal{O}(L^2)$	$\widetilde{\mathcal{O}}\left(dL^3\sqrt{n}/(\alpha\gamma)\right)$	$\widetilde{\mathcal{O}}\left(L^4\sqrt{n}/(\alpha\gamma)\right)$

Table 1: $C_{\mathcal{G}}$ and *worst-case* regret bounds for different graph topologies

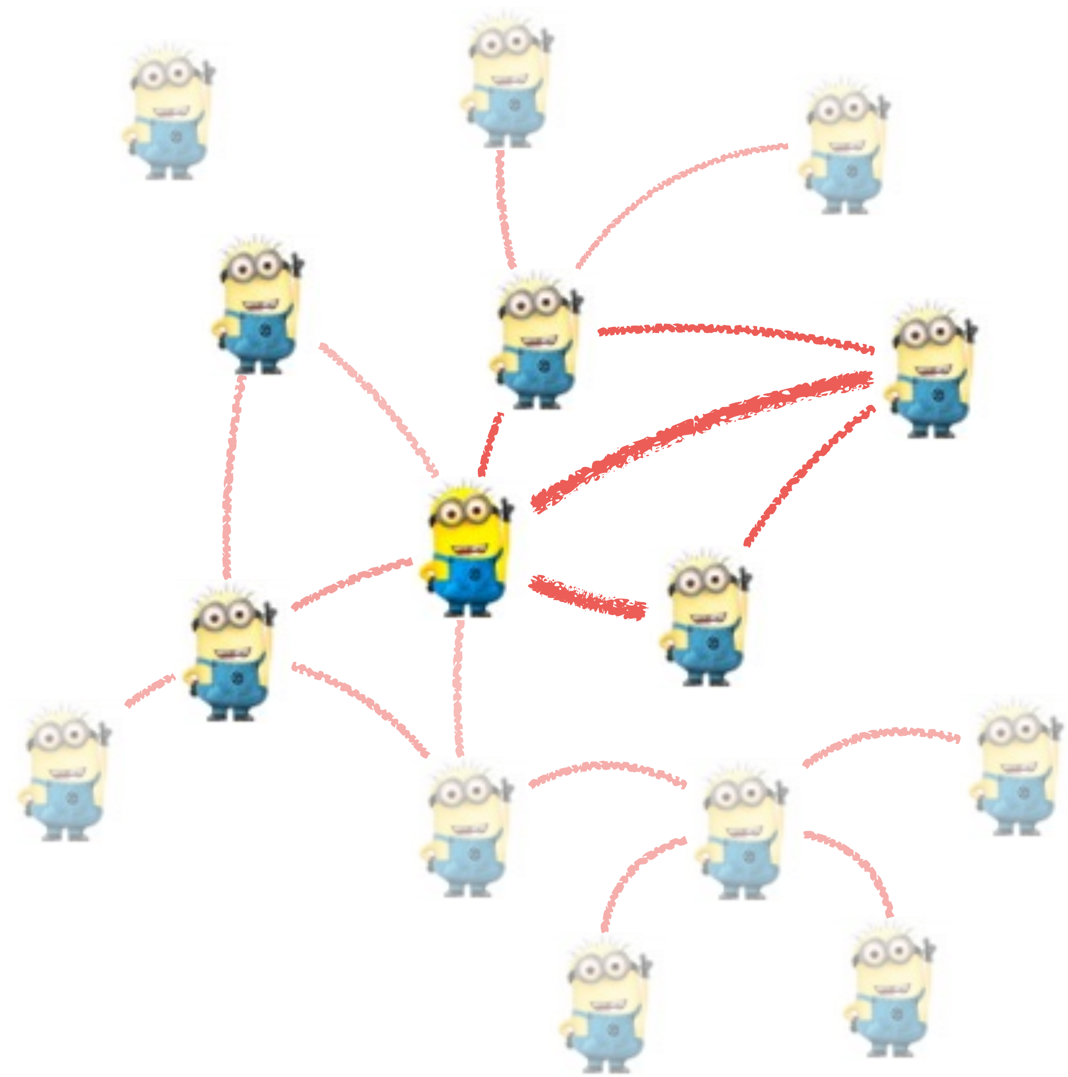
FACEBOOK EXPERIMENT



- real Facebook (a small subgraph)
- weights from U(0,0.1)
- **nodetovec** with d=10
  - imperfect
- K = 10
- CUCB with no linear generalisation

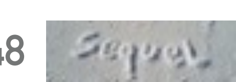
NEXT STEPS

- **Active learning on graphs: online influence maximization**
  - learning the graph **while** acting on it optimally
  - global cascading model with edge level feedback
  - **difficulty of the problem** and scaling with it

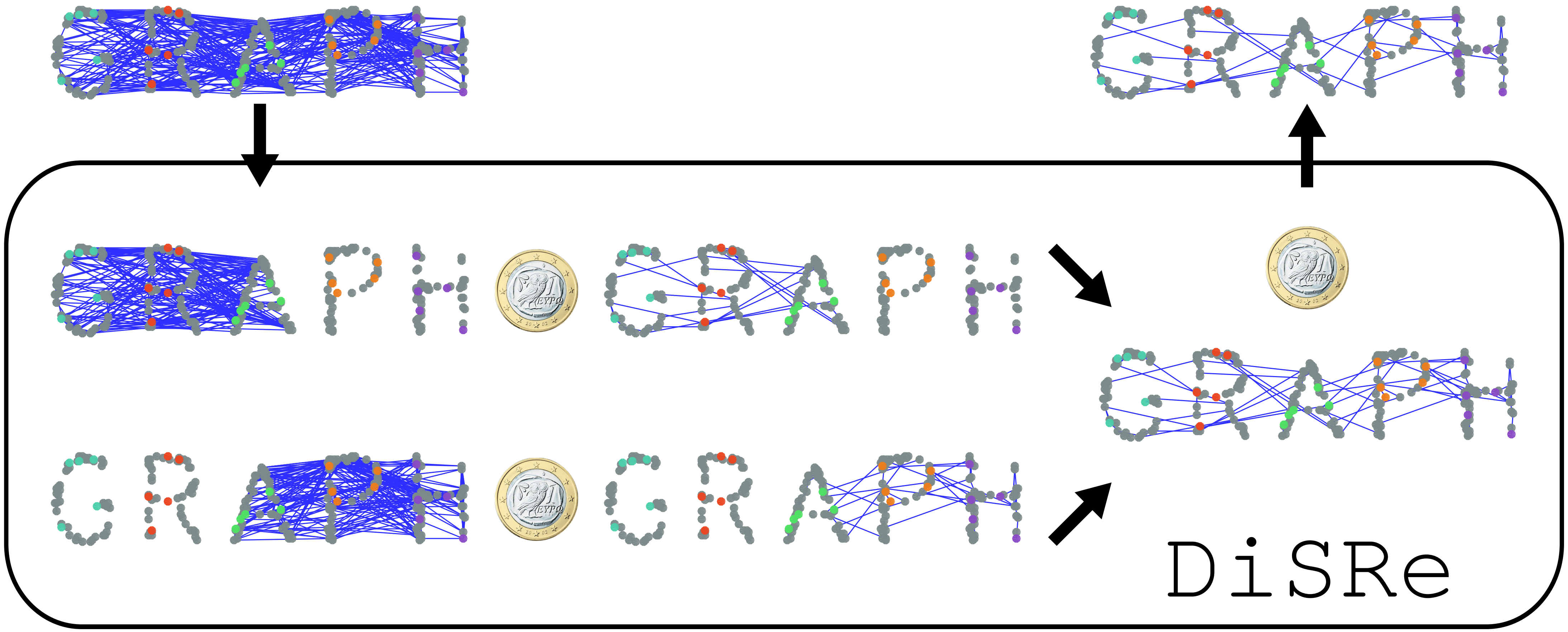


- **What is next?**
  - node-level feedback
  - dynamic/evolving graphs
  - realistic accessibility constraints

Survey: http://researchers.lille.inria.fr/~valko/hp/publications/valko2016bandits.pdf



SCALE UP!!!



MV, Kveton, Huang, Ting: **Online Semi-Supervised Learning on Quantized Graphs** UAI 2010

Kveton, MV, Rahimi, Huang: **Semi-Supervised Learning with Max-Margin Graph Cuts** AISTATS 2010

Calandriello, Lazaric, MV: **Distributed sequential sampling for kernel matrix approximation** AISTATS 2017

Calandriello, Lazaric, MV: **Second-order kernel online convex optimization with adaptive sketching**, ICML 2017

Calandriello, Lazaric, MV: **Efficient second-order online kernel learning with adaptive embedding**, NIPS 2017

Calandriello, Koutis, Lazaric, MV: **Improved large-scale graph learning through ridge spectral sparsification**, ICML 2018

code: http://researchers.lille.inria.fr/~valko/hp/publications/squeak.py

SCALING UP GRAPH LEARNING

- Large graphs do not fit in a single machine memory
- multiple passes slow, distribution has communication costs
- removing edges impacts structure/accuracy
- Make the graph sparse, while preserving its structure for learning

$$(1-\varepsilon)\mathbf{L}_{\mathcal{G}} \preceq \mathbf{L}_{\mathcal{H}} \preceq (1+\varepsilon)\mathbf{L}_{\mathcal{G}}$$

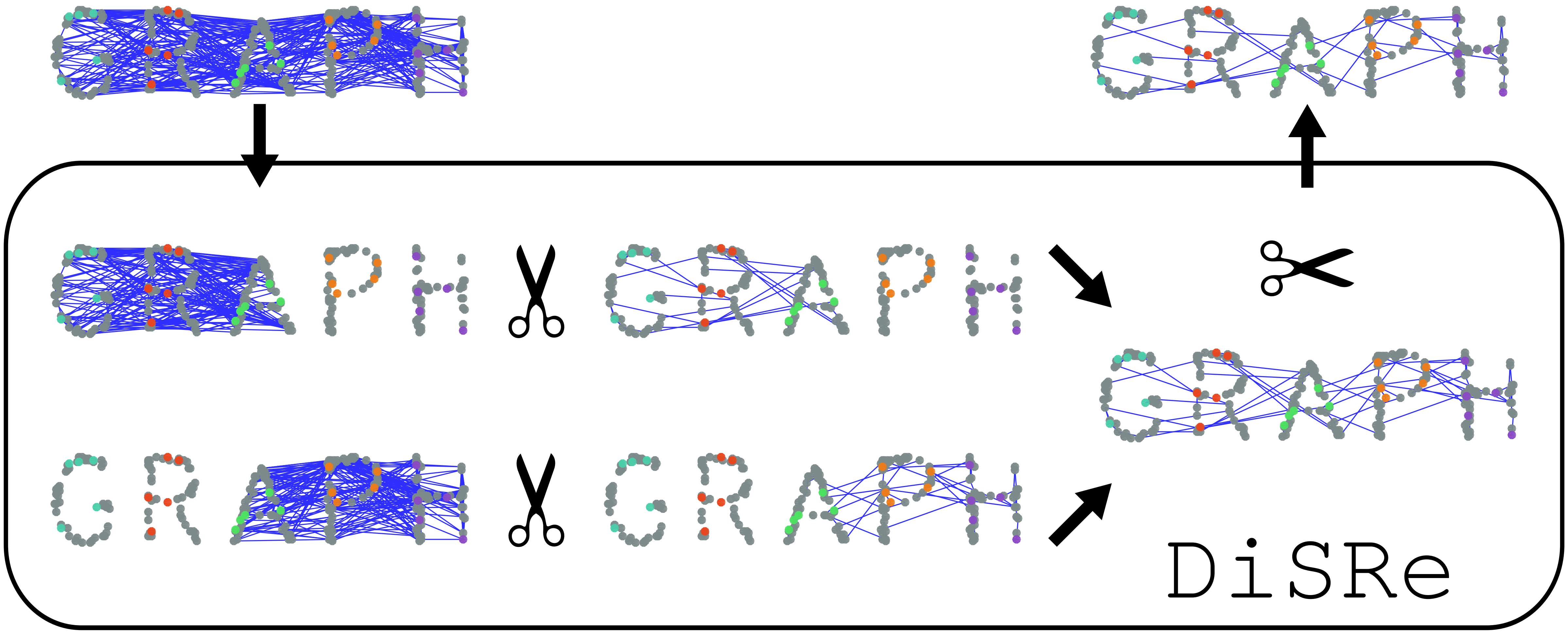
$$(1-\varepsilon)\mathbf{L}_{\mathcal{G}} - \varepsilon\gamma\mathbf{I} \preceq \mathbf{L}_{\mathcal{H}} \preceq (1+\varepsilon)\mathbf{L}_{\mathcal{G}} + \varepsilon\gamma\mathbf{I}$$

Mixed multiplicative / additive error

large (i.e. $\geq \gamma$) directions reconstructed accurately

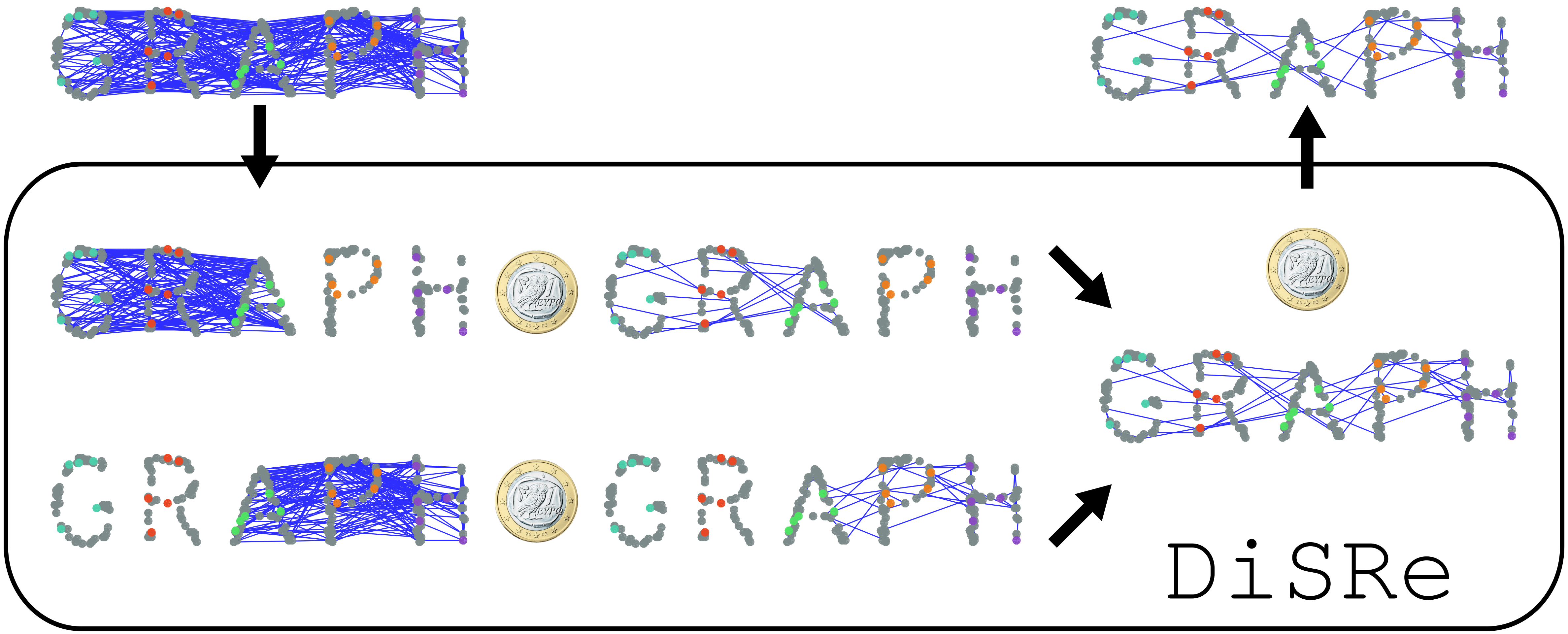
small (i.e. $\leq \gamma$) directions uniformly approximated ($\gamma\mathbf{I}$)

HOW DOES IT WORK?



arbitrarily split in subgraphs that fit in a single machine

DISRE GUARANTEES



Theorem

Given an arbitrary graph $\mathcal{G}$ w.h.p. DiSRe satisfies

(1) each sub-graphs is an (ε, γ)-sparsifier

(2) with at most $\mathcal{O}(d_{\text{eff}}(\gamma)\log(n))$ edges.

DISRE EXPERIMENTS

Dataset: Amazon co-purchase graph [Yang and Leskovec 2015]
↳ natural, artificially sparse (true graph known only to Amazon)
↳ we compute 4-step random walk to recover removed co-purchases [Gleich and Mahoney 2015]

Target: eigenvector $\mathbf{v}$ associated with $\lambda_2(\mathbf{L}_{\mathcal{G}})$ [Sadhanala et al. 2016]

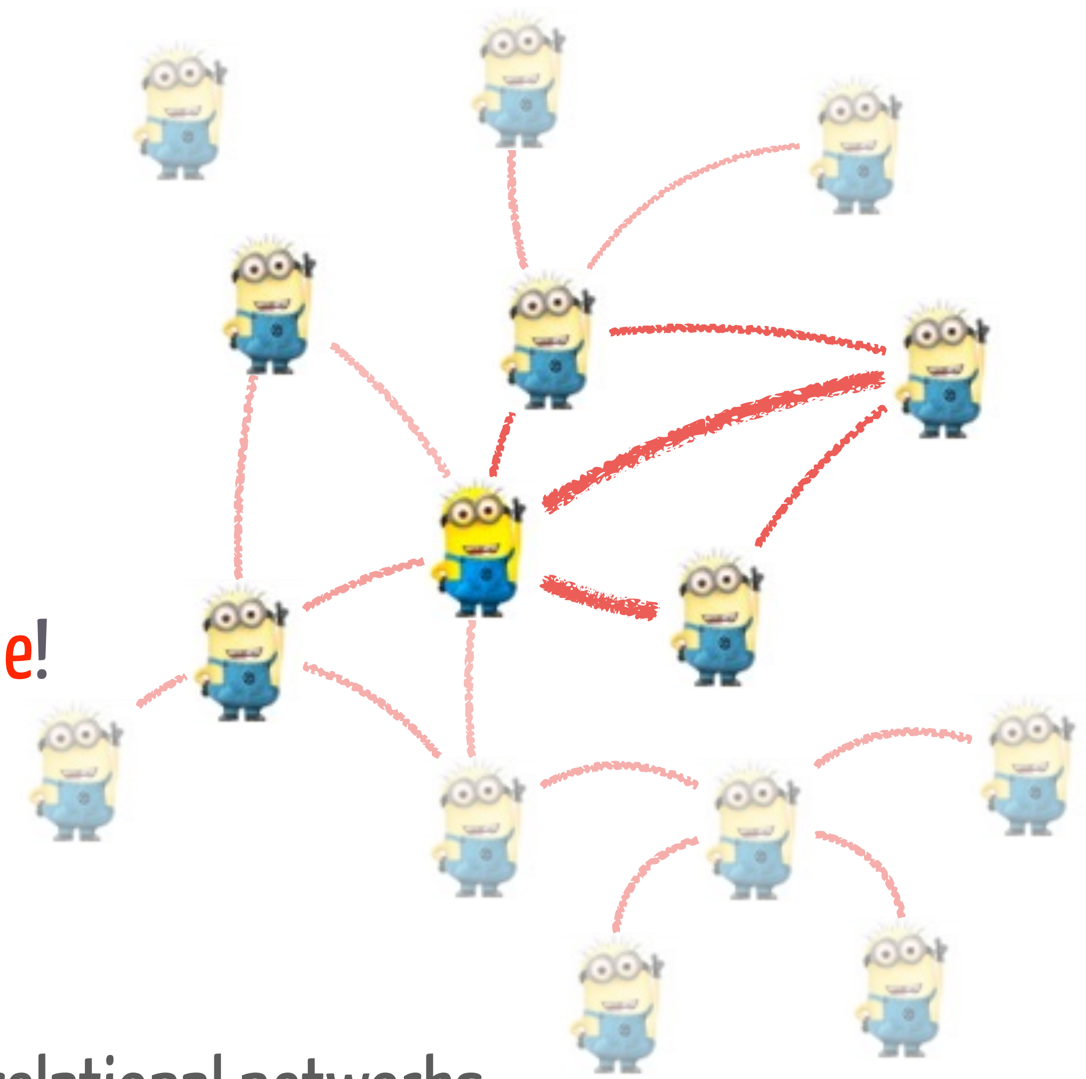
$n = 334,863$ nodes, $m = 98,465,352$ edges (294 avg. degree)

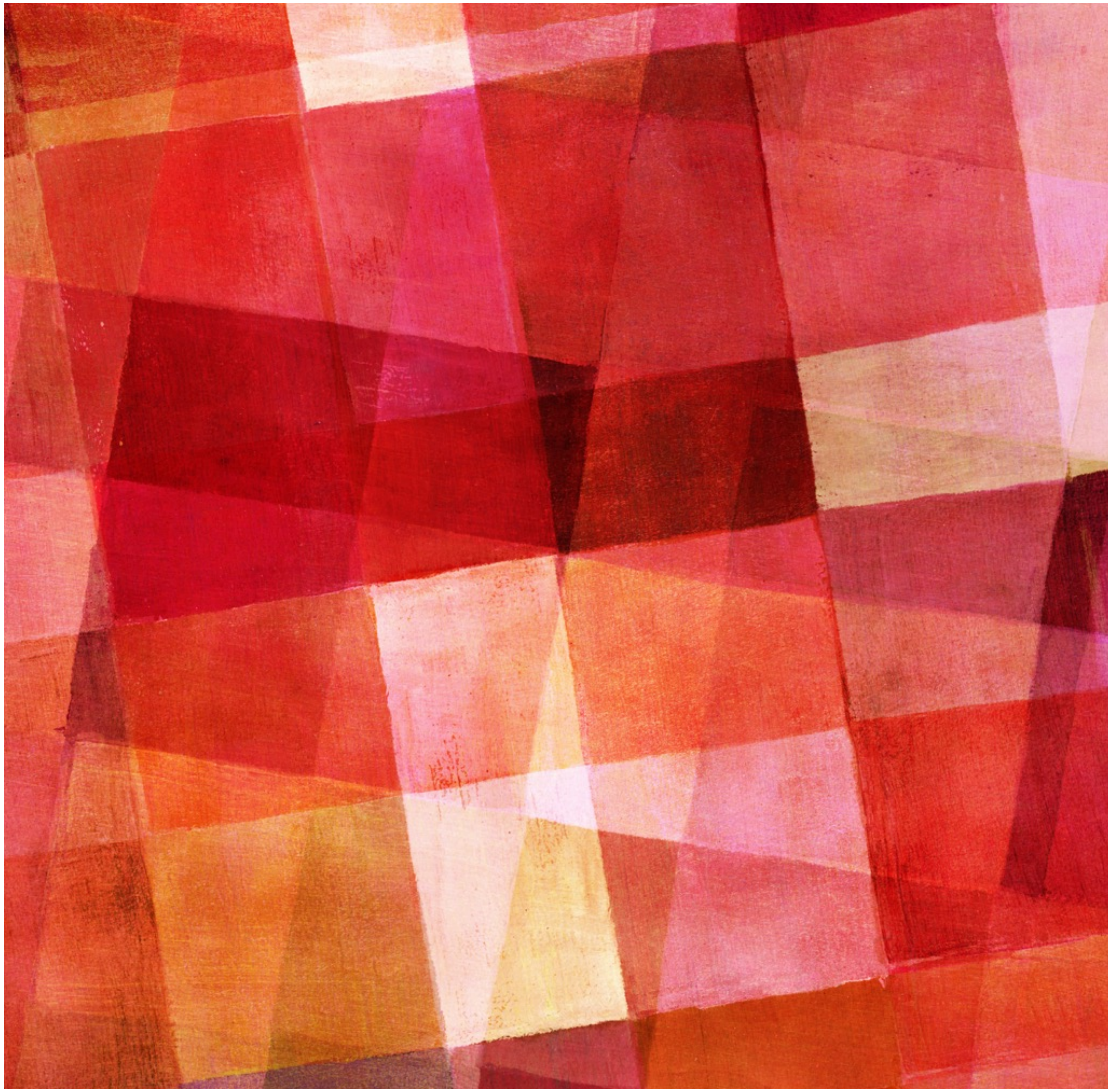
Alg.	*Parameters*	$\lvert\mathcal{E}\rvert$ *(x10^6)*	$\lVert\widetilde{\mathbf{f}} - \mathbf{v}\rVert_2^2$ $(\sigma = 10^{-3})$	$\lVert\widetilde{\mathbf{f}} - \mathbf{v}\rVert_2^2$ $(\sigma = 10^{-2})$
EXACT		98.5	0.067 ± 0.0004	0.756 ± 0.006
kN	$k = 60$	15.7	0.172 ± 0.0004	0.822 ± 0.002
DISRE	$\gamma = 0$	22.8	0.068 ± 0.0004	$\mathbf{0.756} \pm 0.005$
DISRE	$\gamma = 10^2$	11.8	$\mathbf{0.068} \pm 0.0002$	0.772 ± 0.004

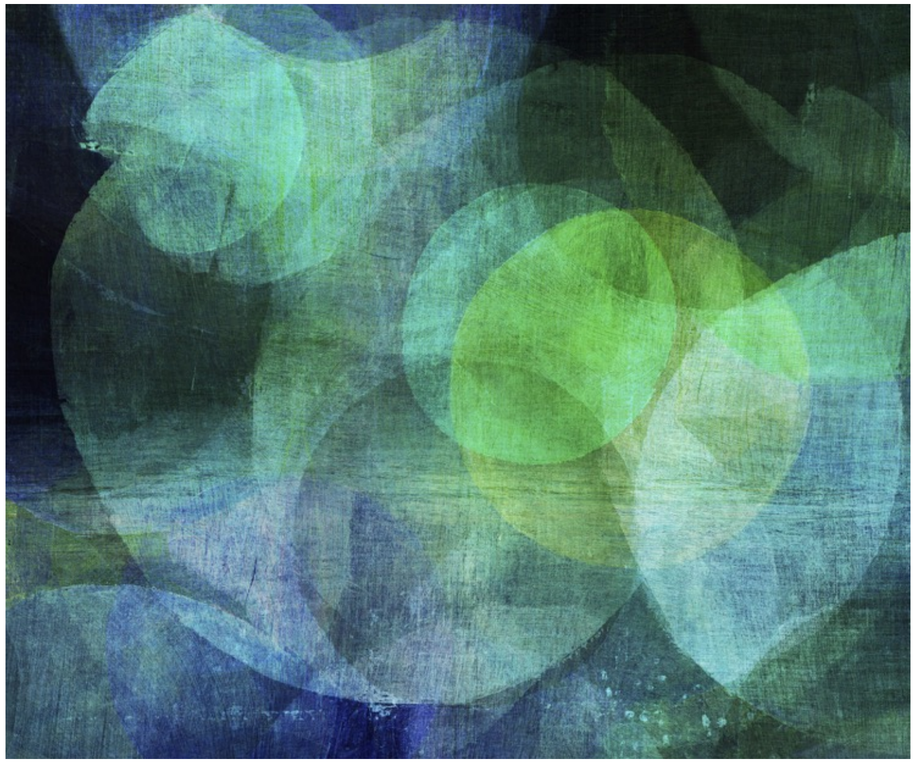
Time: Loading $\mathcal{G}$ from disk 90sec, DISRE 120sec($k = 4 \times 32$ CPU), computing $\widetilde{\mathbf{f}}$ 120sec, computing $\widehat{\mathbf{f}}$ 720sec

CONCLUSION AND NEXT STEPS

- **Graphs** give way to reason and act with **relationships**
  - step above over only considering entities
  - high-level cognition - this is how children learn!
  - online learning and online decision-making
  - **optimal allocation of resources (samples, time)**
  - trees and Monte-Carlo tree search
  - tools to scale up the learning with near-linear time!
- **What is next?**
  - find **the** way to low-level representation
  - graph-networks that operate on the graphs and relational networks
  - intrinsic exploration over graph and other structures







Michal Valko, SequeL, Inria Lille - Nord Europe, michal.valko@inria.fr
http://researchers.lille.inria.fr/~valko/hp/