

Algorithms on graphs and networks

Leonid Zhukov

Adapted by Dmitry Ignatov



Networks

- **Explicit: graphs & networks**
 - Power grids, phone lines
 - Internet, Web graph
 - Social networks (friends, business etc)
- **Implicit: transaction**
 - Email, Messenger, ICQ
 - Online auctions, E-store (Amazon)
 - Foreign trade
- **Constructed: affinity between data points**
 - Music (online radio)
 - Books, videos, ...



Flickr – social network

Photos: [Yours](#) • [Upload](#) • [Organize](#) • [Your Contacts](#) • [Explore](#)

flickr BETA

Hi leonid68!

◆ You have [4 new messages](#).
◆ [Choose your Flickr web address!](#)

Printing? Can it be true?
Well, it's true if you're in the U.S.... with more countries coming online soon! Get 10 free prints with your first order! [Click here to set yourself up for printing.](#)

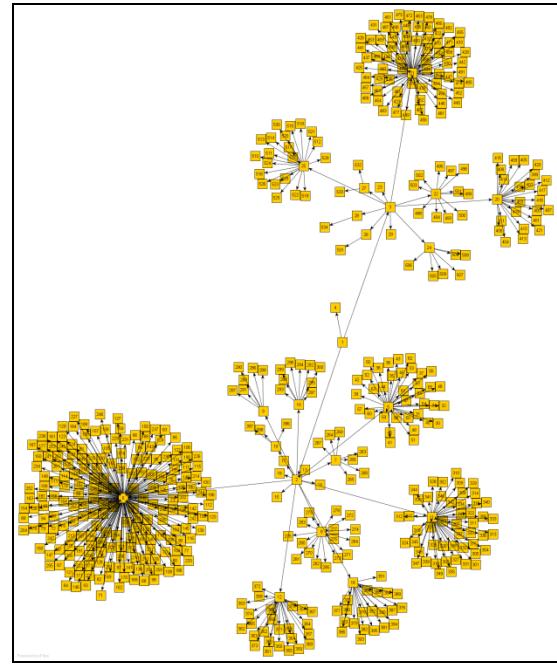
Flickr News
09 Feb 06 - Ladies and gentlemen, you'll notice a brand, spanking new link down in the footer. It's the Flickr Community Guidelines (applause, please).... [read more news](#)

[» Flickr Blog](#) Great photos & latest news, daily!

Do more with your photos! DISMIS X
Posters! Books! DVDs!
Now there's even [more you can do](#) with your photos:
◆ [PhotoShow](#) DVDs NEW
◆ [Qoop](#) Calendars NEW, Posters & Books
◆ [Zazzle](#): U.S. postage stamps with your photos
◆ [Englaze](#) back-up

Make your photos happy
And don't forget to set your

Photos: [Yours](#) • [Upload](#) • [Organize](#) • [Your Contacts](#) • [Explore](#)



flickr BETA

Your contacts

Friends (5)

[Valeria Temple](#) [skoof](#) [katerina 2006](#) [Steshka](#) [mithandor](#)

Search for people [SEARCH](#)
(Or, try the [advanced search](#).)

◆ [Who counts you as a contact?](#)

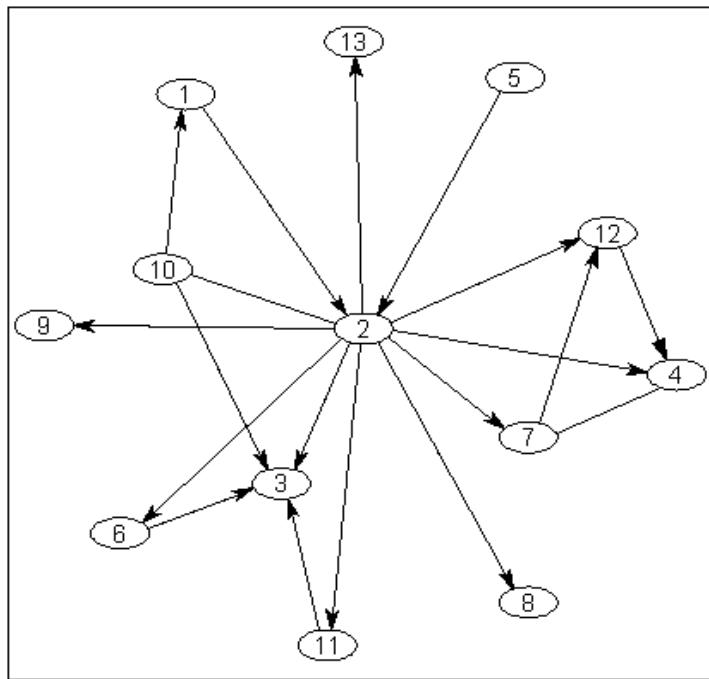


Flickr – a graph



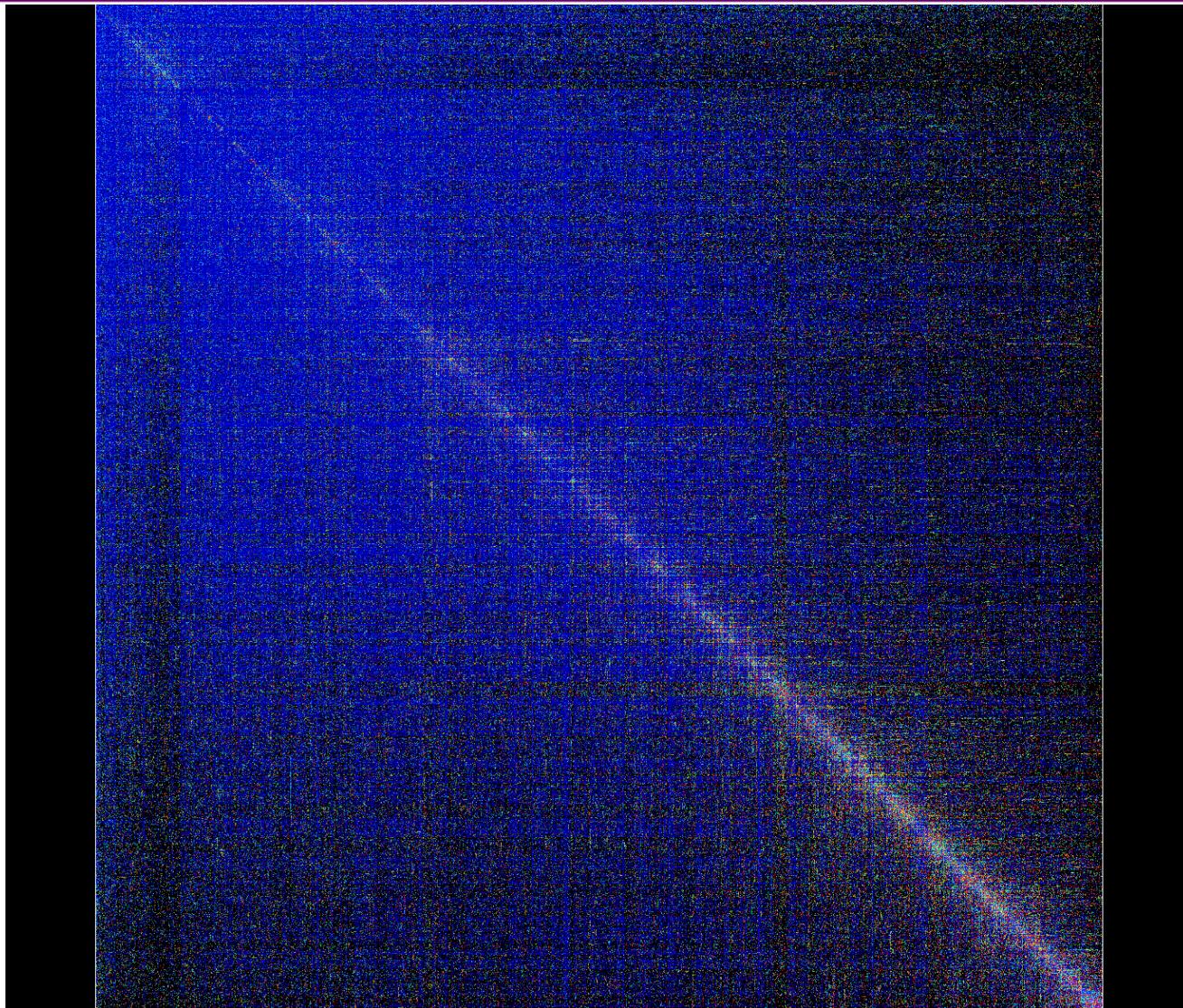


Adjacency matrix



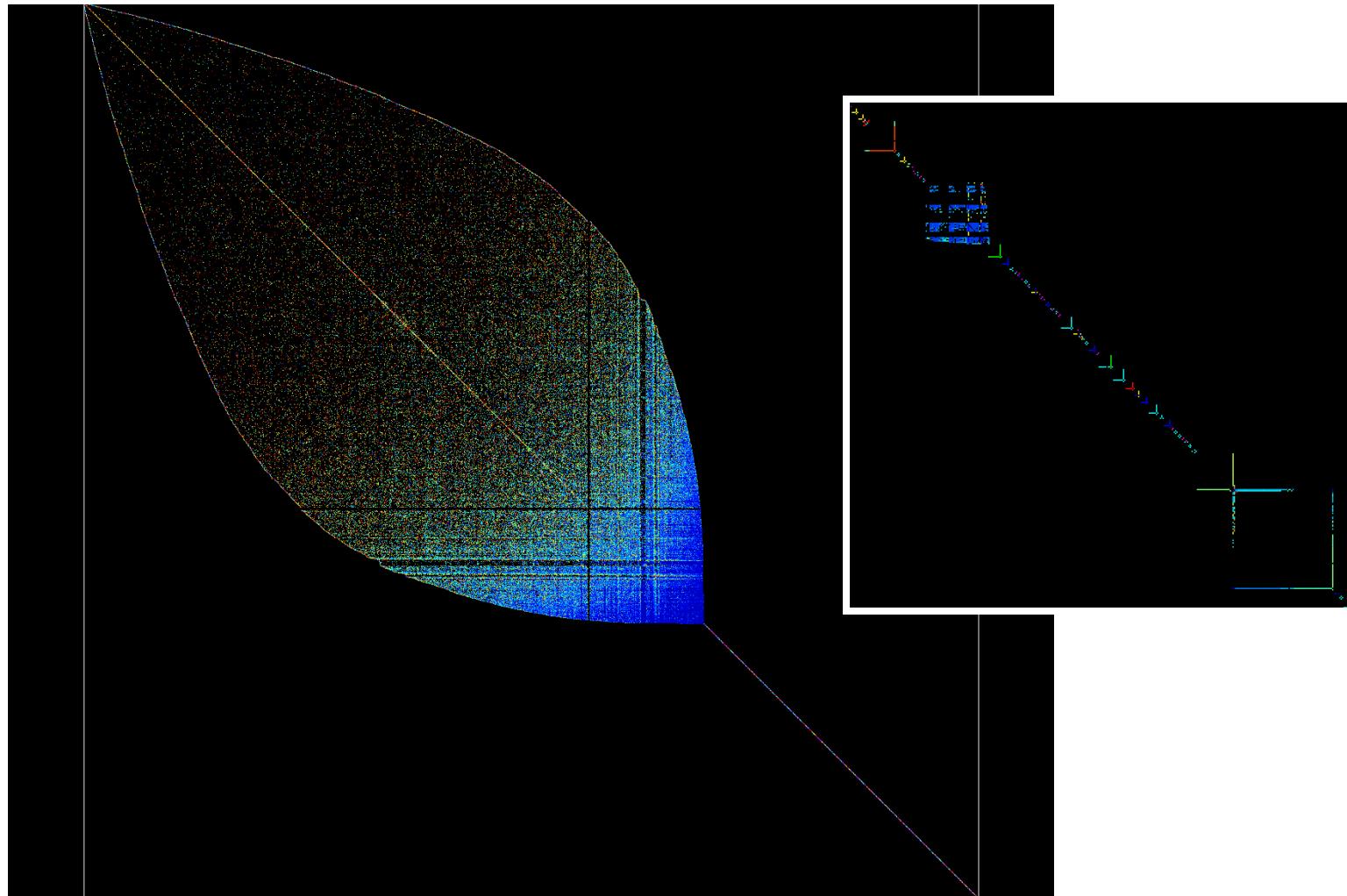


Adjacency matrix





Flickr: reverse Cuthill-McKee



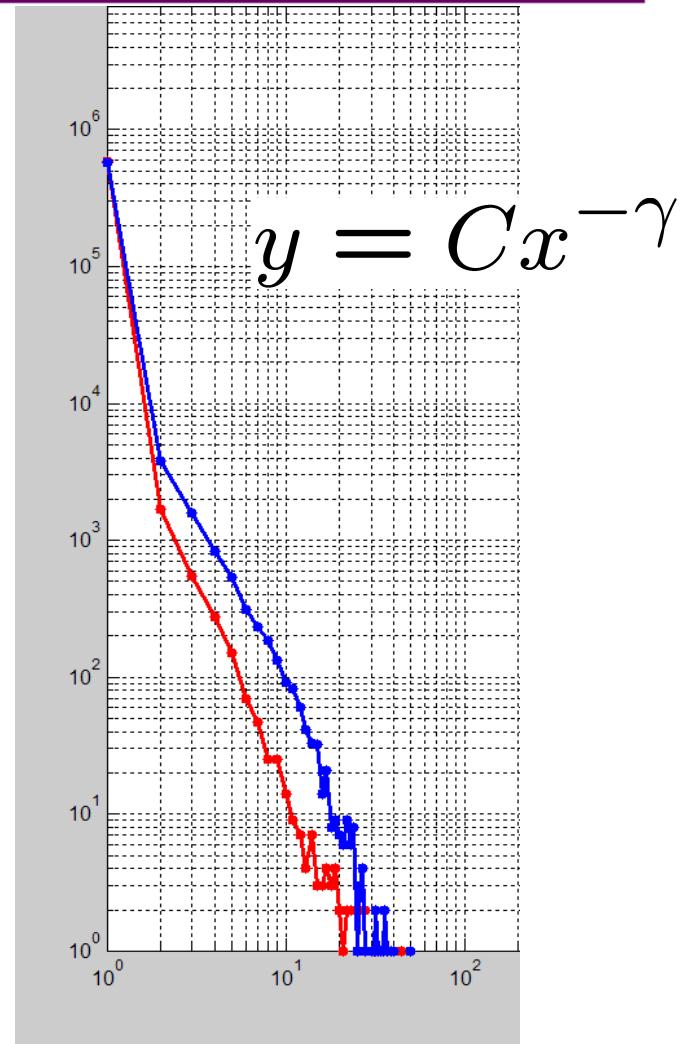
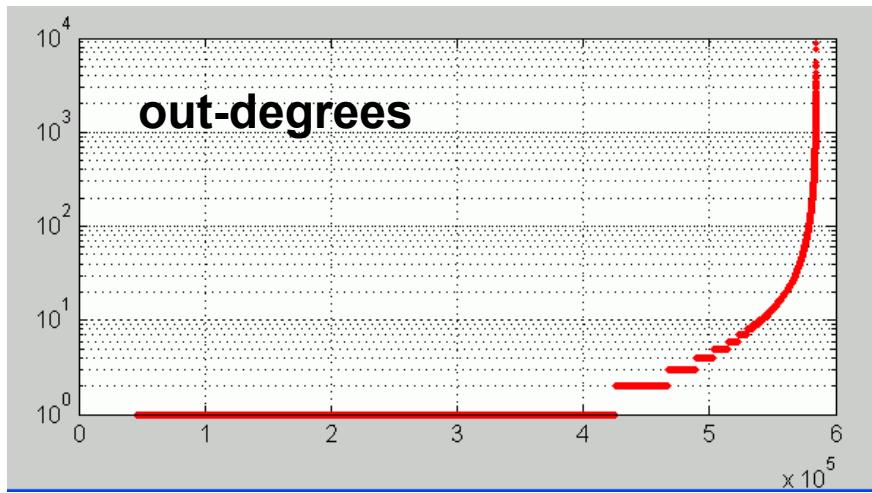
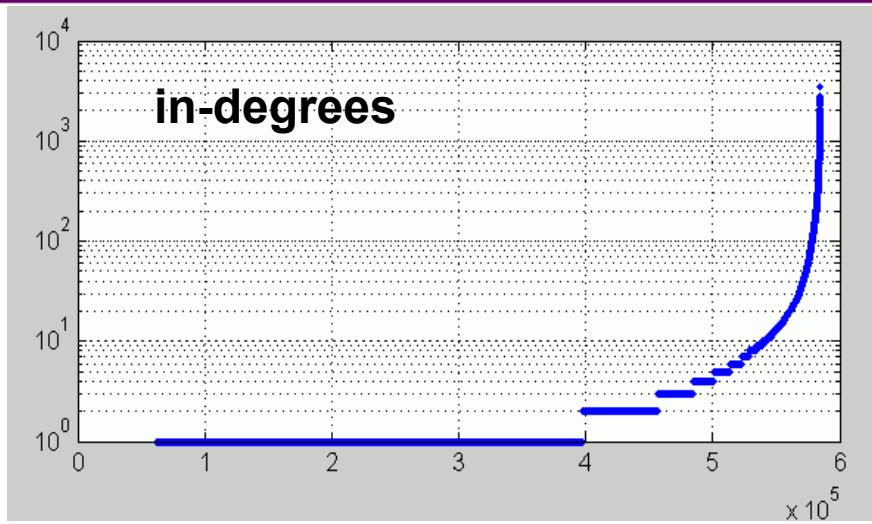


Flickr: some stats

- # nodes = 584,207
- # edges (nnz) = 3,555,115
- max in-degree = 3531
- max out-degree = 8976
- $\langle \text{in-degree} \rangle = \langle \text{out-degree} \rangle = 6$
- diameter = 18
- # strongly connected components = 152,324
- largest strongly connected components = 274,649 374 186 155 11
- # connected component = 43,189
- largest connected components = 404,893 378 112 108 103
- highest core number = 249 (size 668)

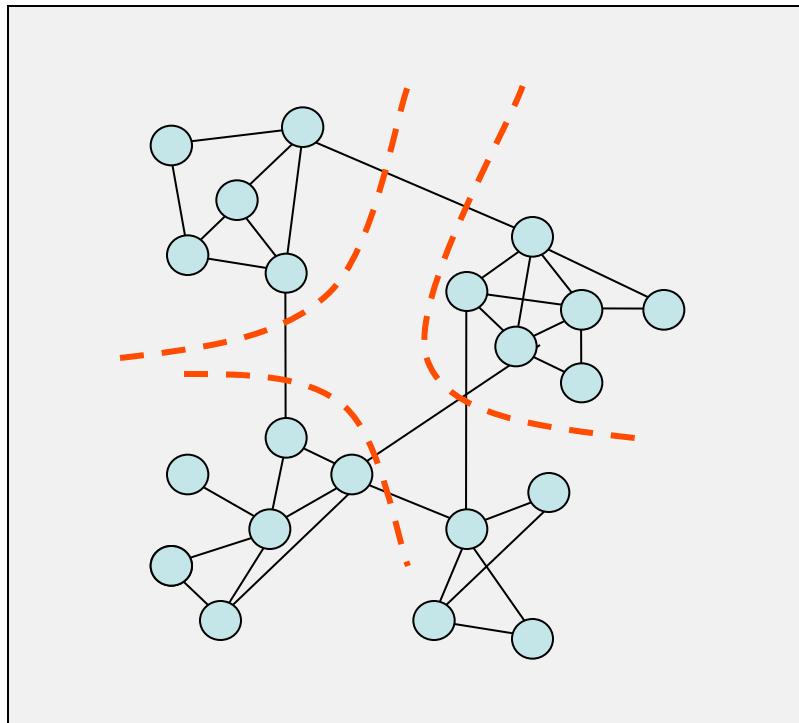


Flickr: scale free graph



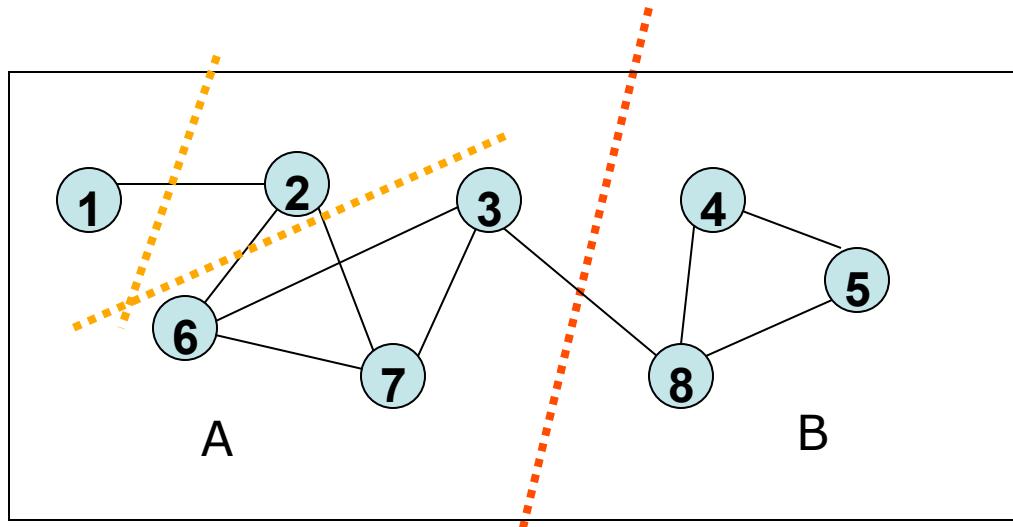


Graph partitioning





Graph partitioning

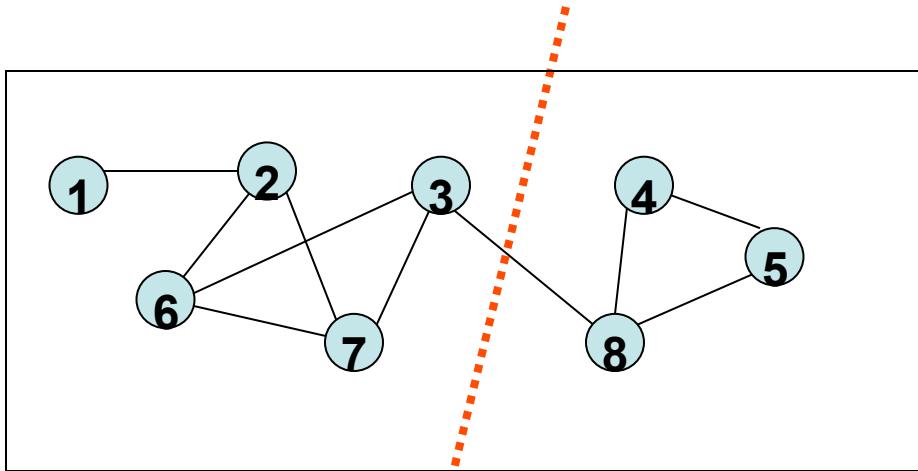


Graph separators: $cut(A, B) = \sum_{v_a \in A, v_b \in B} w(v_a, v_b)$

Normalized cut: $NCut(A, B) = \frac{cut(A, B)}{assoc(A, V)} + \frac{cut(A, B)}{assoc(B, V)}$



Spectral graph partitioning



- assign each node indicator $p = \pm 1$ $p = (-1, -1, -1, +1, -1, +1, -1, +1)$
- smallest cut: $cut = \frac{1}{4} \sum_{i>j} (p_i - p_j)^2 w_{ij}$
- combinatorial optimization, NP hard, relax:

$$p_i = \{-1, 1\}^N \Rightarrow x_i \in [-1, 1], x_i \in R^1$$



Spectral: solution

- Quadratic optimization:

$$E = \frac{1}{4} \sum_{i>j} (x_i - x_j)^2 w_{ij} = \frac{\mathbf{x}^T \mathbf{L} \mathbf{x}}{4}$$

$$\sum_i x_i^2 = N, \quad (\mathbf{x}^T \mathbf{e}) = 0$$

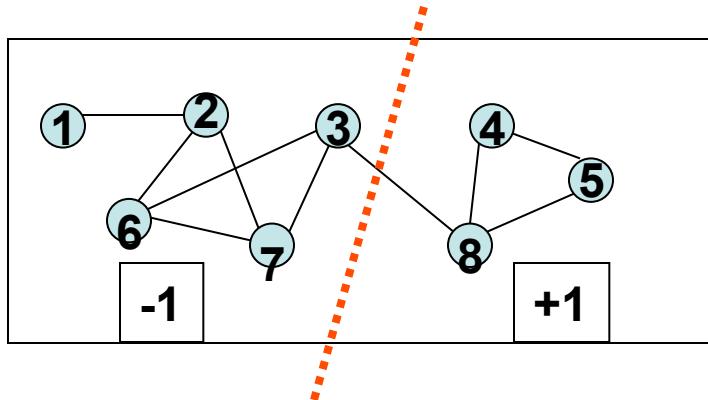
- Eigenvalue problem:

$$\mathbf{L} \mathbf{x} = \lambda \mathbf{x}, \quad \lambda_2, \quad x_2$$

$$p_i = +1, \quad \text{if } x_i > 0; \quad p_1 = -1, \quad \text{if } x_i < 0$$



Example: normalized cuts



$$L = D - A$$

$$Lx = \lambda Dx$$

$$\lambda_2, x_2$$

$$L = \begin{matrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 3 & 0 & 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & 3 & 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 0 & 2 & -1 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 2 & 0 & 0 & -1 \\ 0 & -1 & -1 & 0 & 0 & 3 & -1 & 0 \\ 0 & -1 & -1 & 0 & 0 & -1 & 3 & 0 \\ 0 & 0 & -1 & -1 & -1 & 0 & 0 & 3 \end{matrix}$$

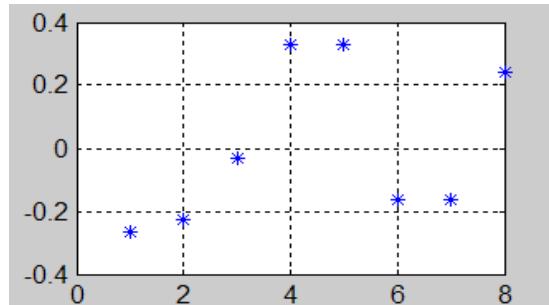
$$x = \begin{matrix} -0.2641 \\ -0.2285 \\ -0.0341 \\ 0.3294 \\ 0.3294 \\ -0.1646 \\ -0.1646 \\ 0.2406 \end{matrix} \Rightarrow \begin{matrix} -1 \\ -1 \\ -1 \\ 1 \\ 1 \\ -1 \\ -1 \\ 1 \end{matrix}$$



Spectral: ordering

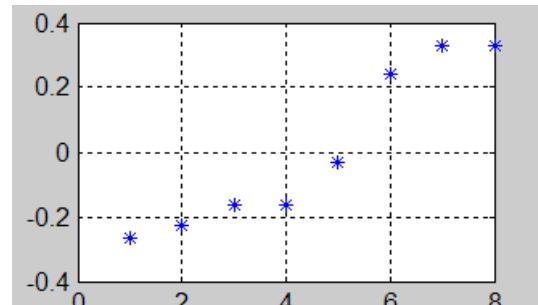
Eigenvector

-0.2641
-0.2285
-0.0341
0.3294
0.3294
-0.1646
-0.1646
0.2406

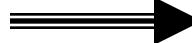
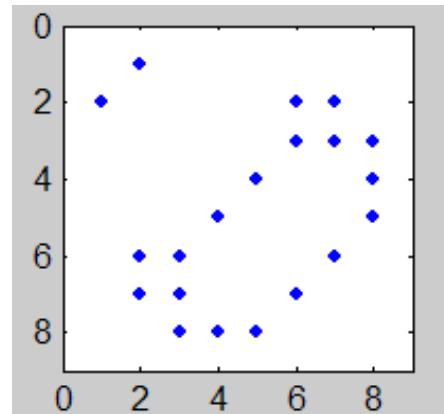


Eigenvector – sorted

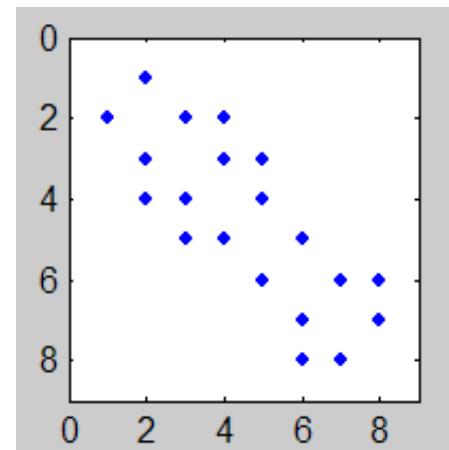
0.3294
0.3294
0.2406
-0.0341
-0.1646
-0.1646
-0.2285
-0.2641



Adjacency matrix



Adjacency matrix – re-ordered

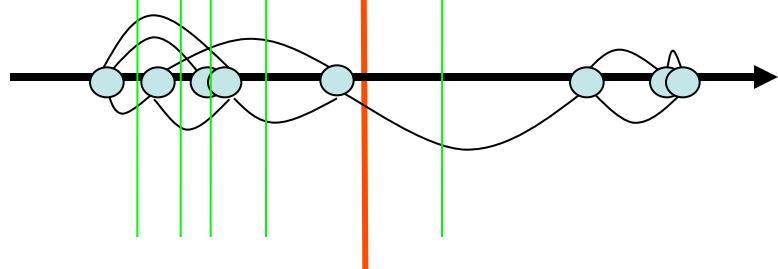
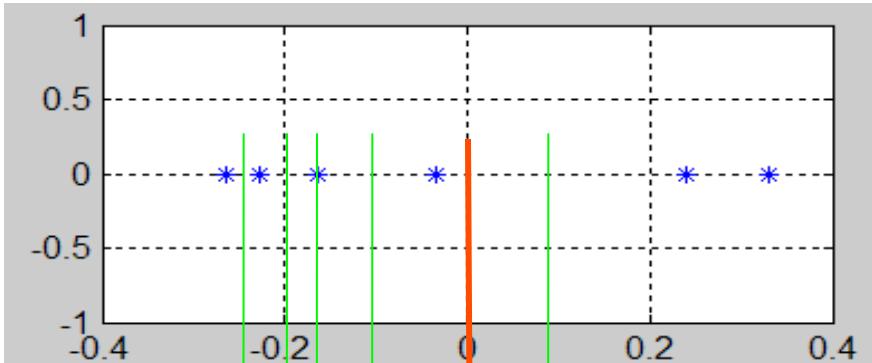


perm = [1 2 3 4 5 6 7 8]



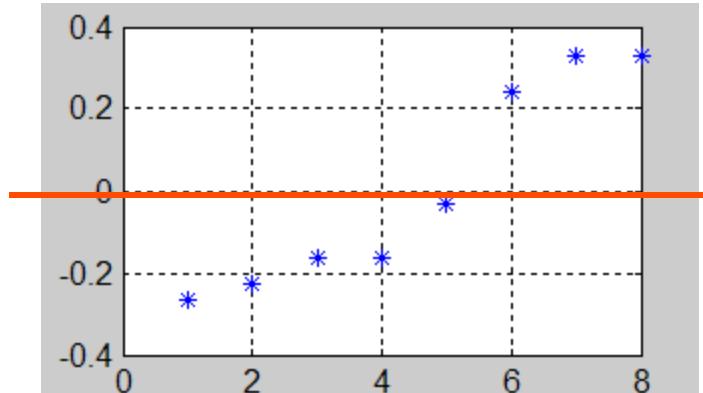
Spectral: graph cut

node layout

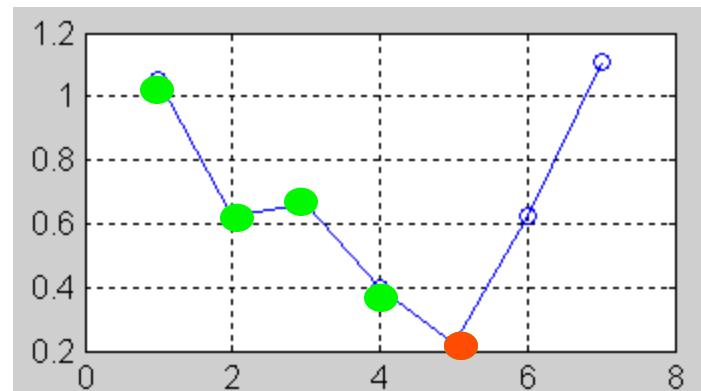


$$cut = \frac{1}{4} \sum_{i>j} (p_i - p_j)^2 w_{ij}$$

Eigenvector – sorted

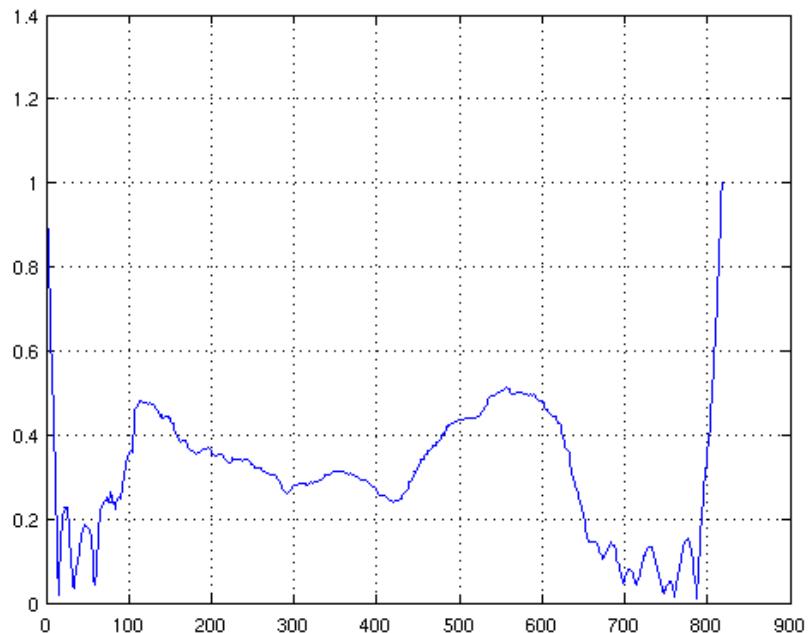
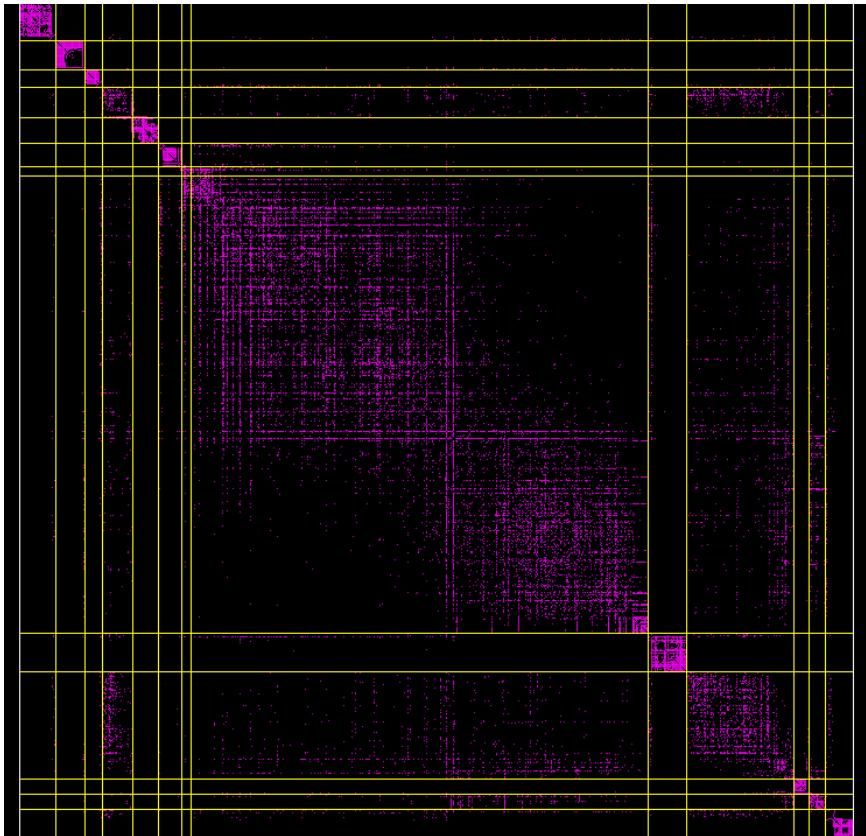


Cut values



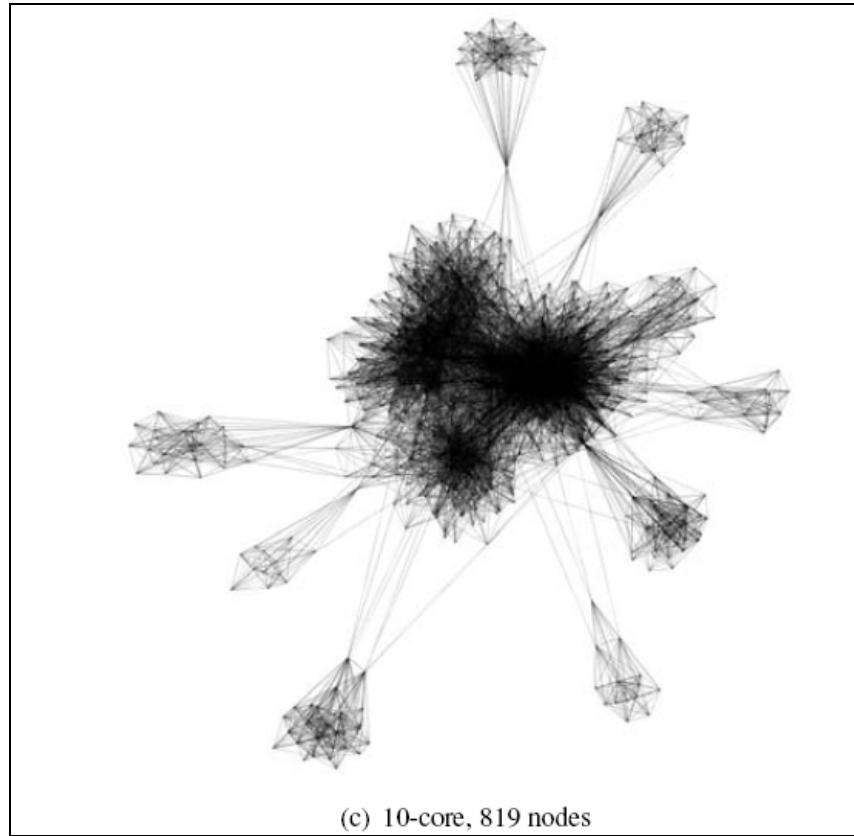
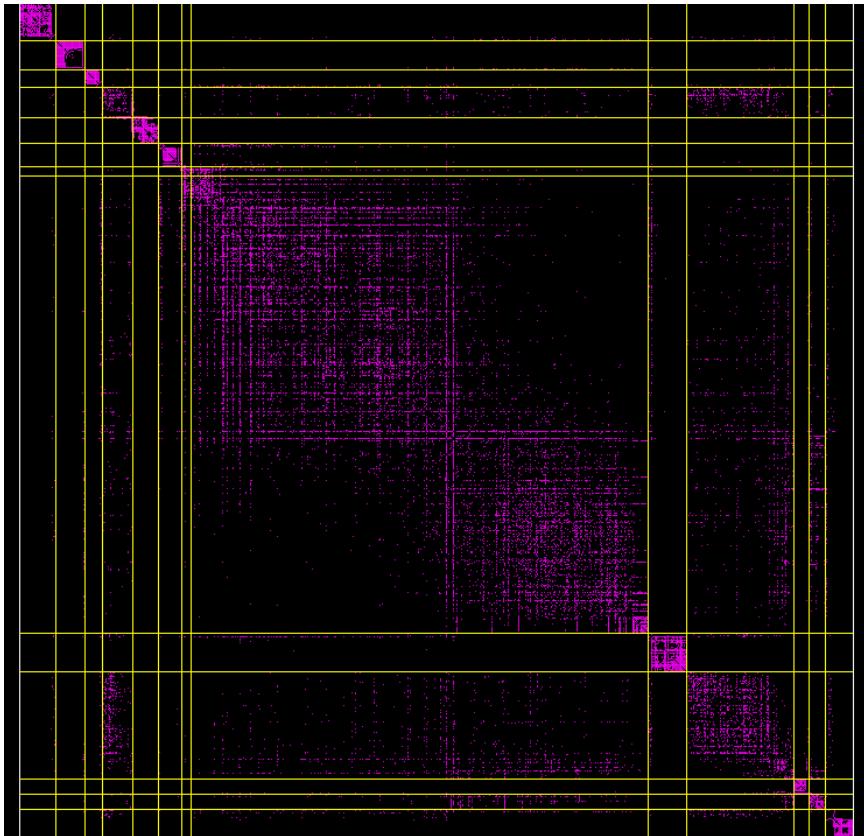


Clustering



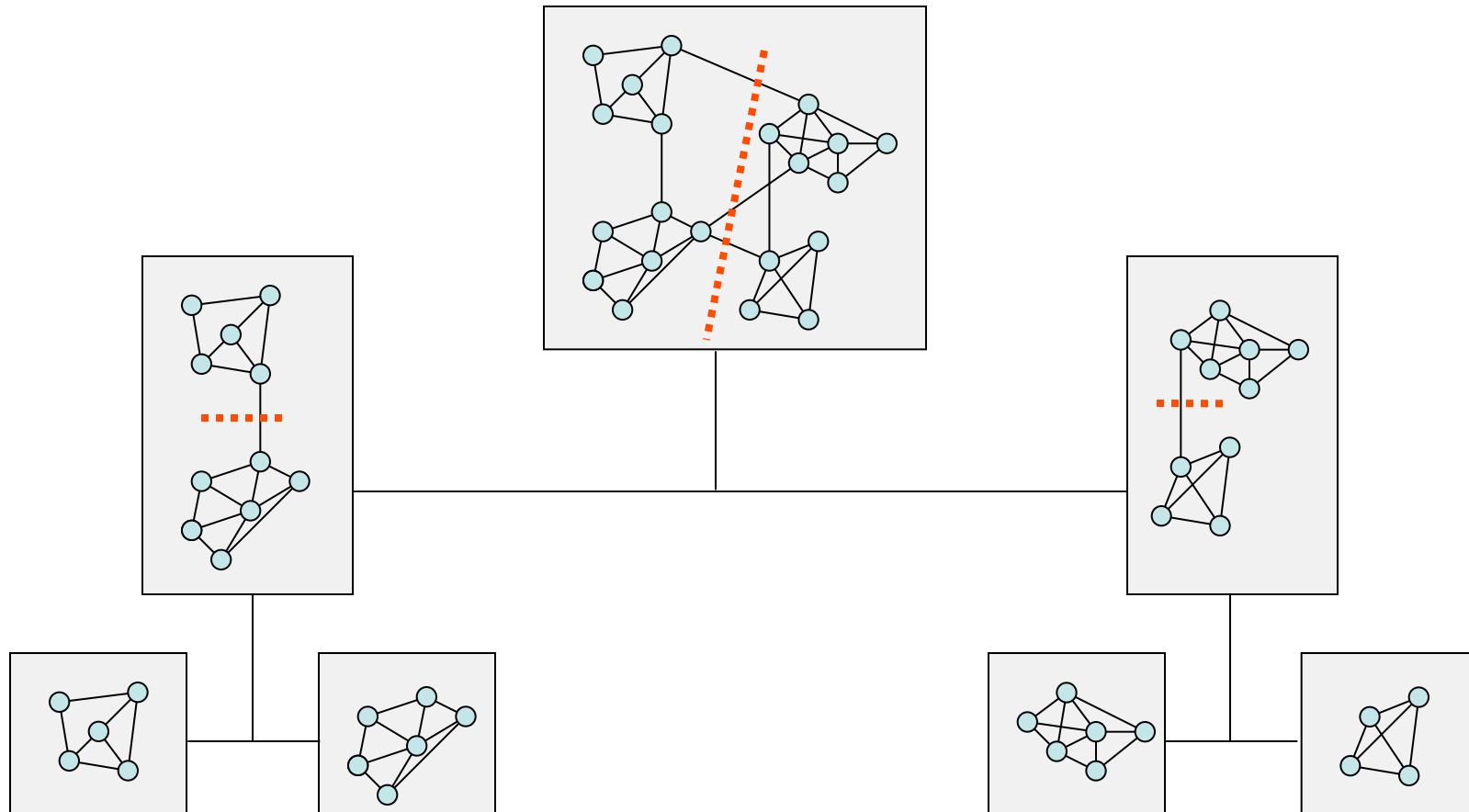


Clustering



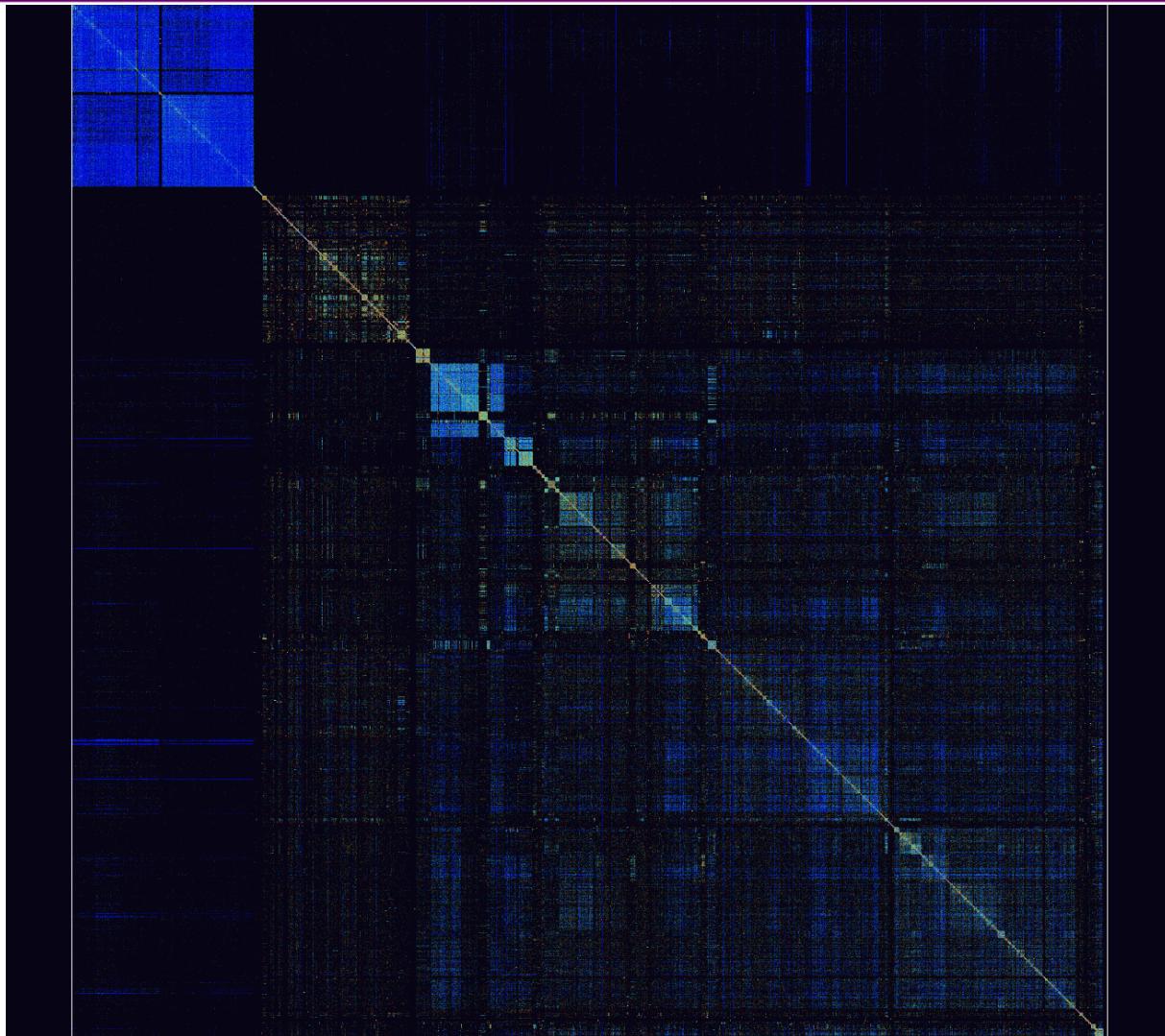


Recursive partitioning tree



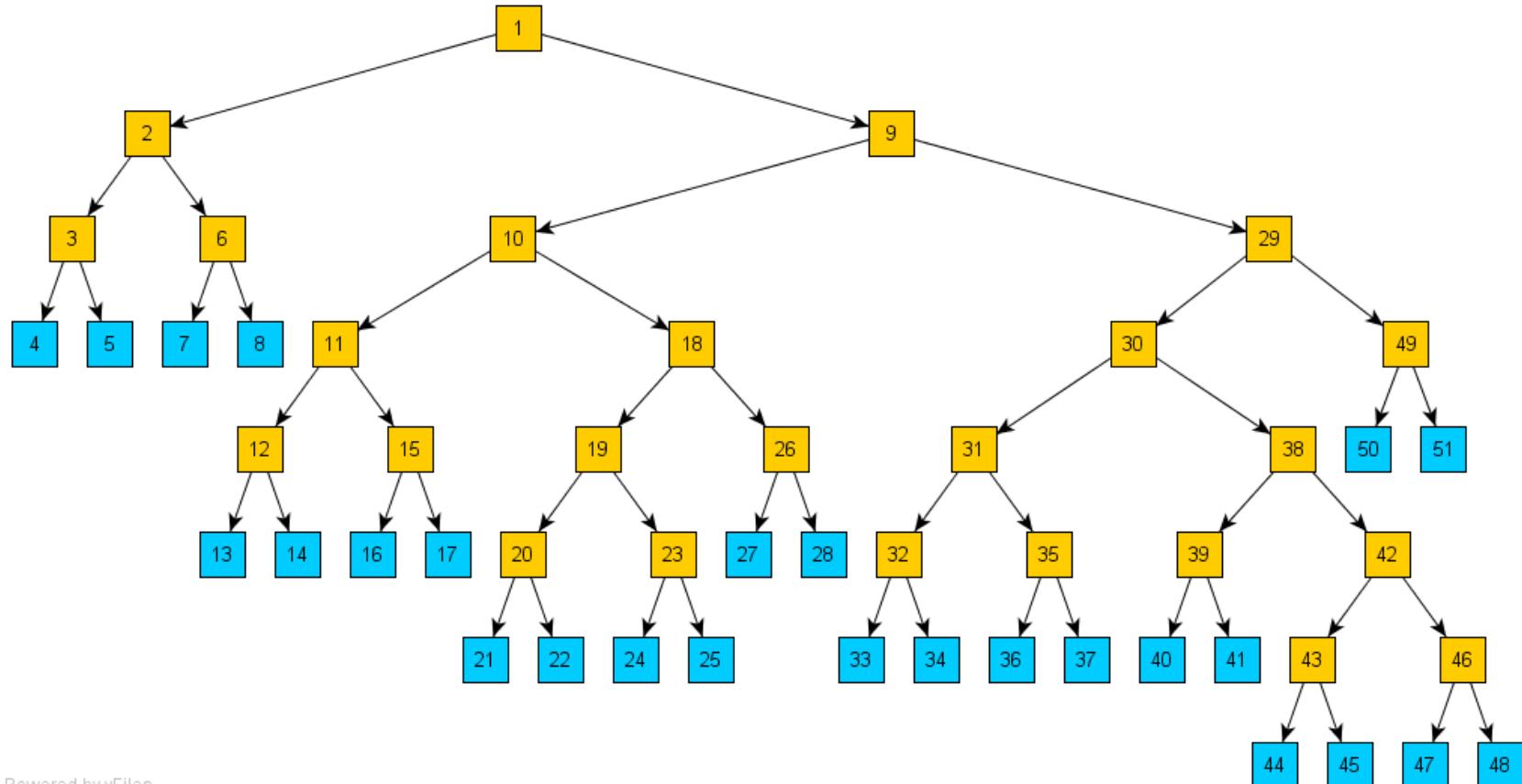


Flickr: “10-core” spectral ordering





Partition tree



Powered by yFiles



Sponsored search

Web | Images | Video | Directory | Local | News | Shopping

YAHOO! SEARCH epson

My Web BETA Subscriptions (New) Shortcuts Advanced Search Preferences

Search Results Results 1 - 10 of about 28,800,000 for **epson** - 0.02 sec. (About this page)

Also try: [epson printers](#), [epson driver](#), [epson p-2000](#), [epson scanners](#) More...

SPONSOR RESULTS

- **Epson** Go with an industry favorite in business printers-Hewlett-Packard. Reliable, high-quality laser, color and all-in-one printers. Compare HP printers head-to-head with **Epson**.
www.hp.com
- **Epson Compatible Ink Cartridges - \$6.95** High-quality **Epson** compatible inkjet cartridges from Inkjetcartridge.com. Toll-free sales and support. Yahoo five-star award-winning service.
www.inkjetcartridge.com
- **75% off Epson Ink - Free Shipping** Up to 75% off **Epson** ink and toner, plus a one-year quality guarantee and free shipping. Save an extra 5% with coupon code over15. Free promotional items with every purchase.
www.123inkjets.com

Y [Epson - Ink Jet Printers](#) - [Scanners](#) - [Projectors](#)
Yahoo! Shortcut - [About](#)

1. **Epson** creates and sells digital imaging products including printers, digital cameras, scanners, and projectors. Also offers ink jet cartridges, paper, and a variety of other supplies and accessories.
Category: [B2B Imaging Equipment](#)
www.epson.com - 8k - [Cached](#) - [More from this site](#) - [Save](#) - [Block](#)

Epson Batteries and Chargers - eBatts
eBatts sells **Epson** batteries and chargers. Batteries and chargers for...
www.ebatts.com

Save on Epson Inks
Buy 2 **Epson** printer ink cartridges, get 1 cartridge free. Free 2 day...
www.clickinks.com

Epson Inkjet Cartridges - Save
Save up to 80% on **Epson** inkjet cartridges. We offer a 100% money-back...
www.printpal.com

Epson Multimedia Projectors
Shop with us to save money on multimedia projection equipment.
www.projectorsforsale.com

Epson Compatible Ink

View Bids Tool

1. **Epson**
Go with an industry favorite in business printers-Hewlett-Packard. Reliable, high-quality laser, color and all-in-one printers. Compare HP printers head-to-head with Epson.
www.hp.com
(Advertiser's Max Bid: \$1.80)
2. **Epson Compatible Ink Cartridges - \$6.95**
High-quality Epson compatible inkjet cartridges from Inkjetcartridge.com. Toll-free sales and support. Yahoo five-star award-winning service.
www.inkjetcartridge.com
(Advertiser's Max Bid: \$1.79)
3. **75% off Epson Ink - Free Shipping**
Up to 75% off Epson ink and toner, plus a one-year quality guarantee and free shipping. Save an extra 5% with coupon code over15. Free promotional items with every purchase.
www.123inkjets.com
(Advertiser's Max Bid: \$1.79)

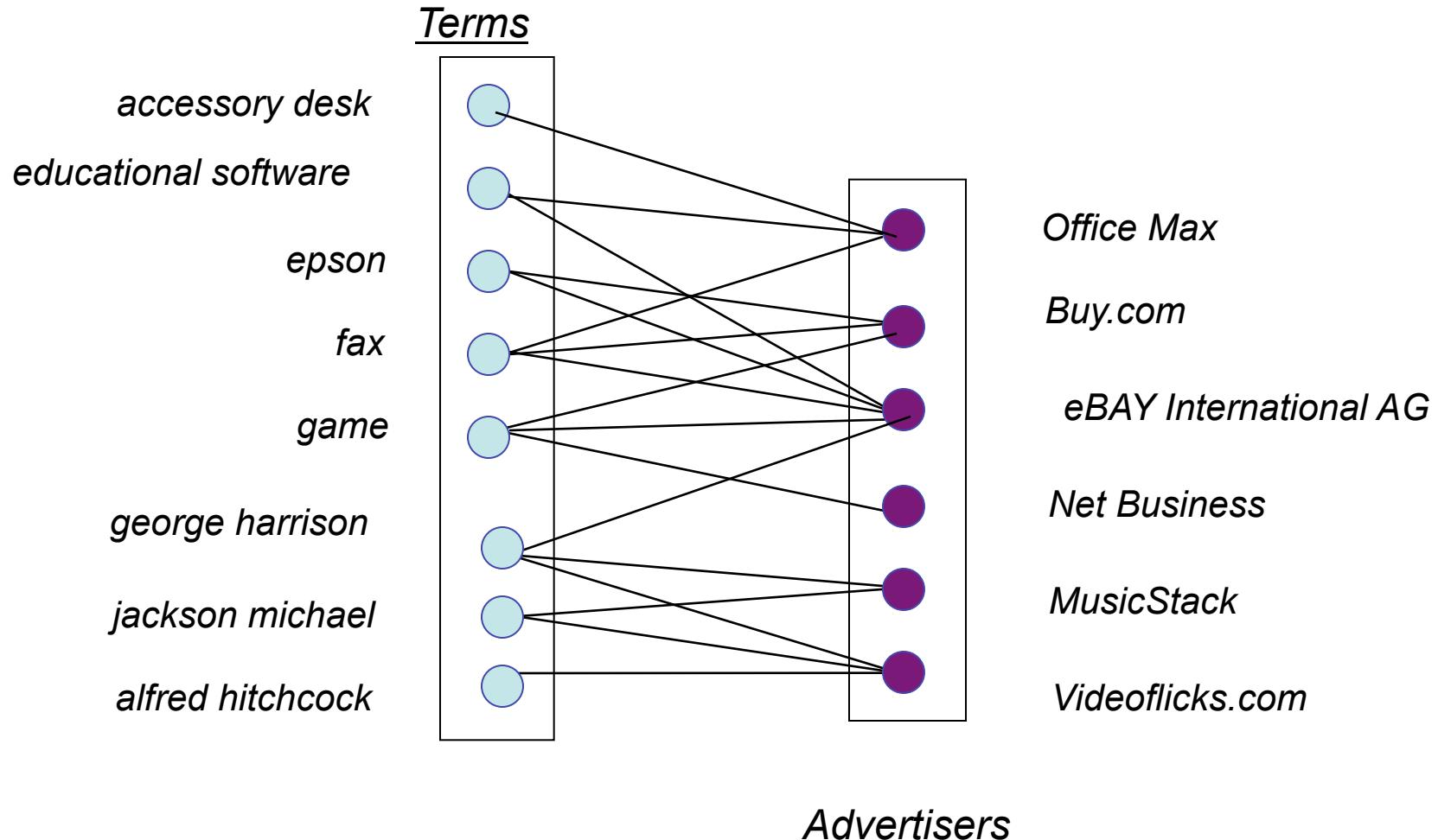


Sponsored search data

<i>bidded term</i>	<i>\$ bid</i>	<i>advertiser</i>
accessory desk	.83	Office Max
alfred hitchcock	.01	Videoflicks.com
educational software	.13	Buy.com
educational software	.4	eBAY International AG
educational software	.17	OfficeMax
epson	.28	Buy.com
epson	.4	eBAY International AG
fax	.13	Buy.com
fax	.4	eBAY International AG
fax	.38	OfficeMax
game	.02	Net Business
game	.25	Buy.com
george harrison	.15	eBAY International AG
george harrison	.05	MusicStack
george harrison	.01	Videoflicks.com
jackson michael	.05	MusicStack
jackson michael	.01	Videoflicks.com

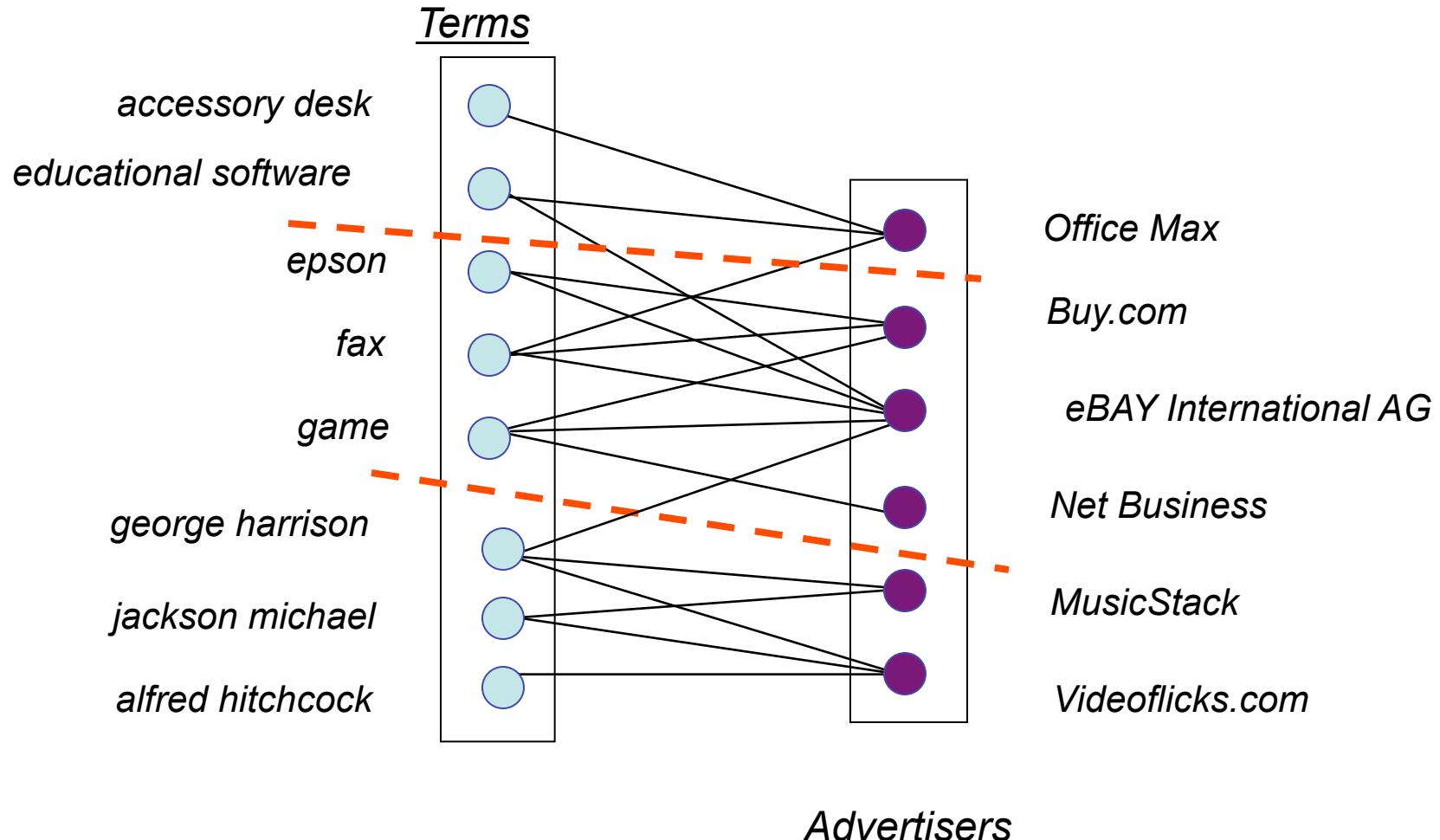


Sponsored search bid graph





Graph partitioning





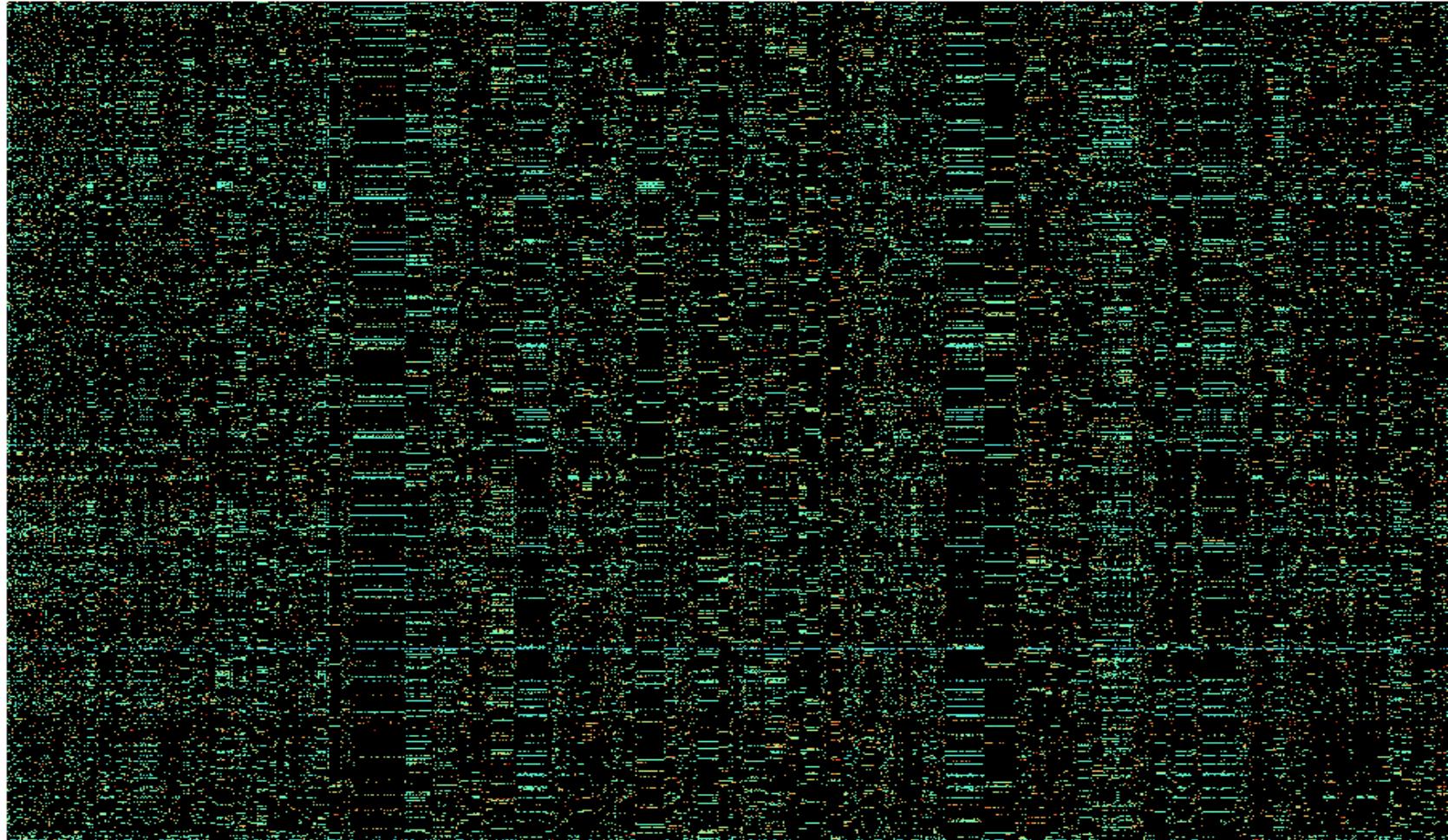
Sponsored search bid matrix

	1	2	3	4	5	6	7	8
Office Max								
Net Business								
Buy.com								
eBAY								
MusicStack								
Videoflicks.com								

- | | |
|-------------------------|---------------------|
| 1. accessory desk | 5. game |
| 2. educational software | 6. george harrison |
| 3. epson | 7. jackson michael |
| 4. fax | 8. alfred hitchcock |

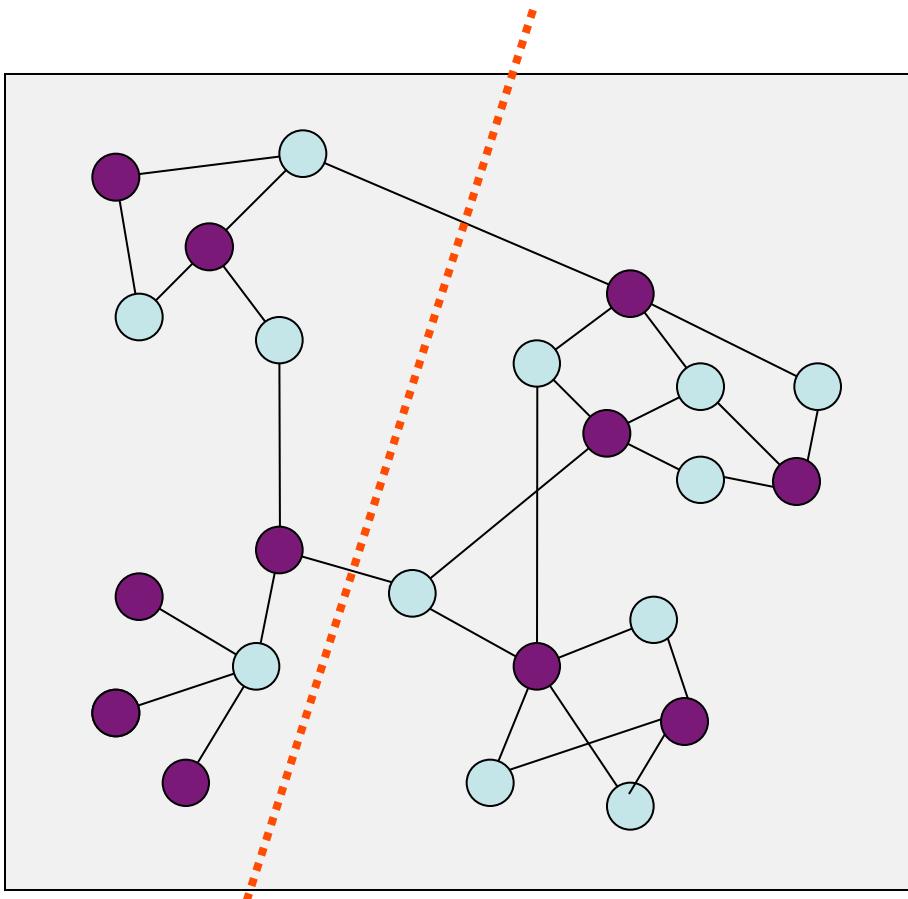


Overture bid graph : real life





Bi-partite graph partitioning



● Terms

● Advertisers

$$\hat{\mathbf{A}} = \begin{pmatrix} 0 & \mathbf{A} \\ \mathbf{A}^T & 0 \end{pmatrix}$$



Bi-partite formulation

- Eigensystem:

$$\begin{pmatrix} D_1 & -A \\ -A^T & D_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

- SVD:

$$A_n = D_1^{-1/2} A D_2^{-1/2}$$

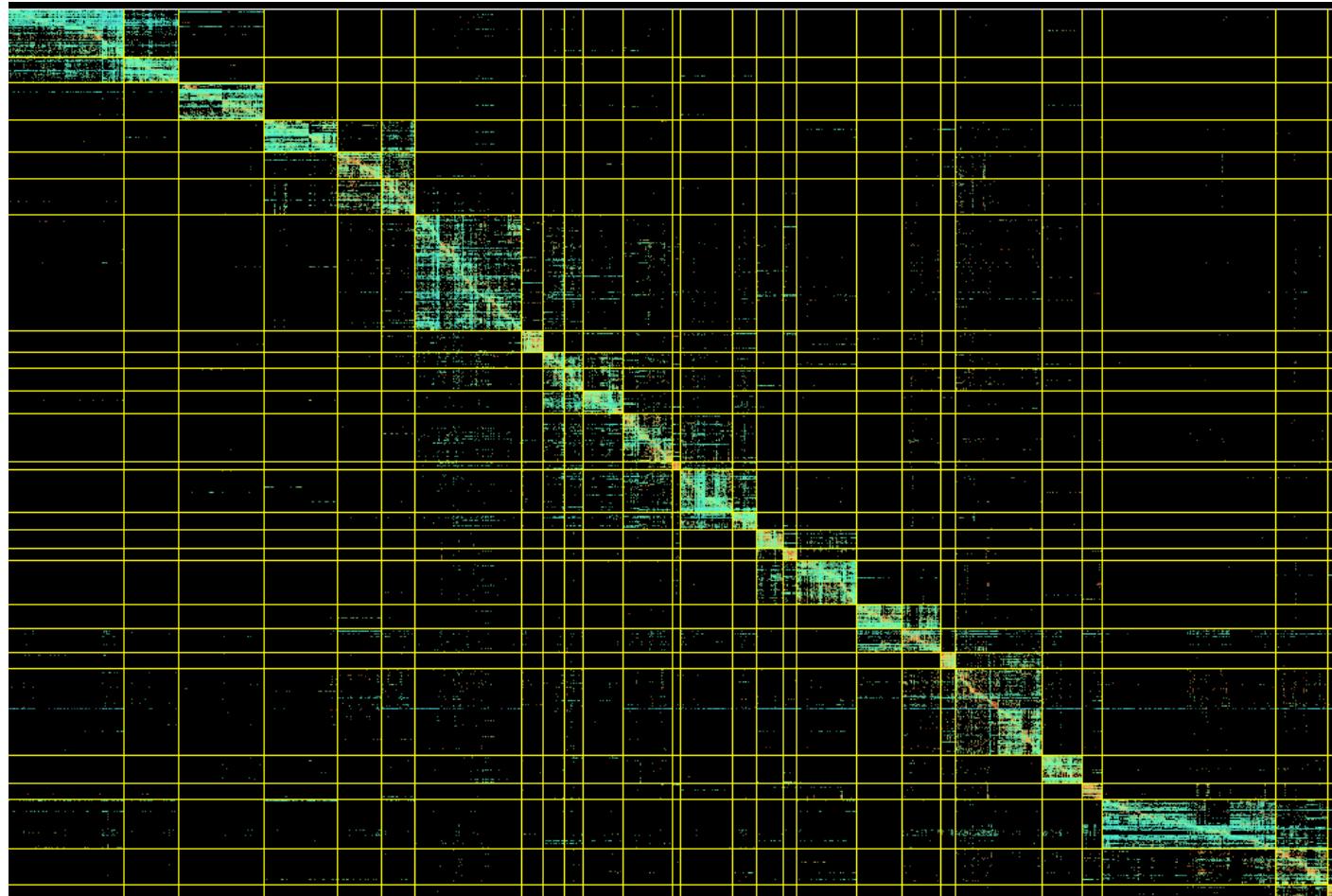
$$A_n = u \sigma v^T, \quad \sigma = 1 - \lambda$$

- Solution:

$$x = D^{-1/2} u, \quad y = D^{-1/2} v$$

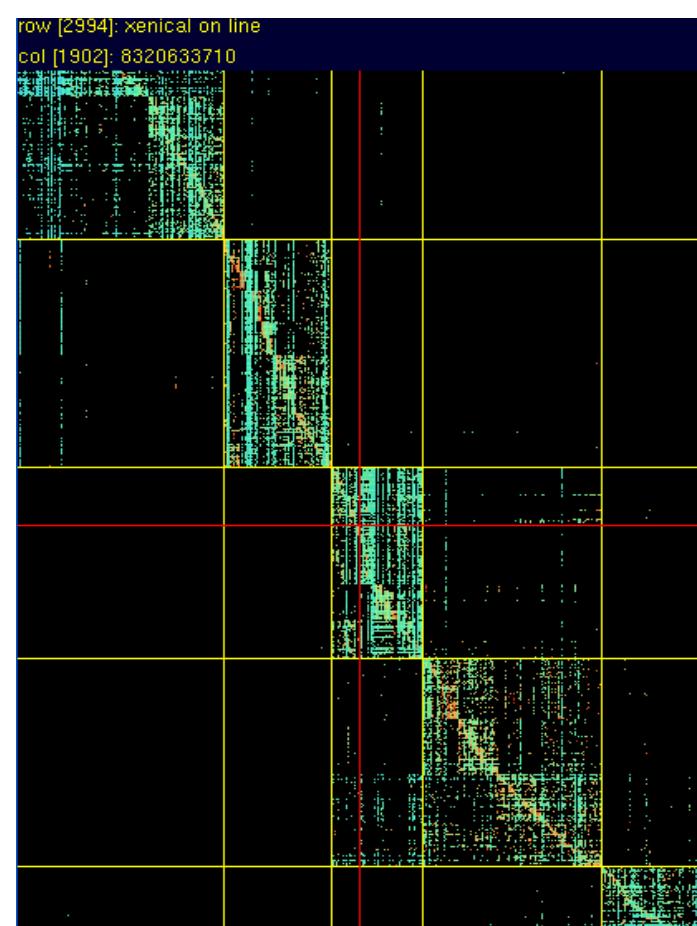
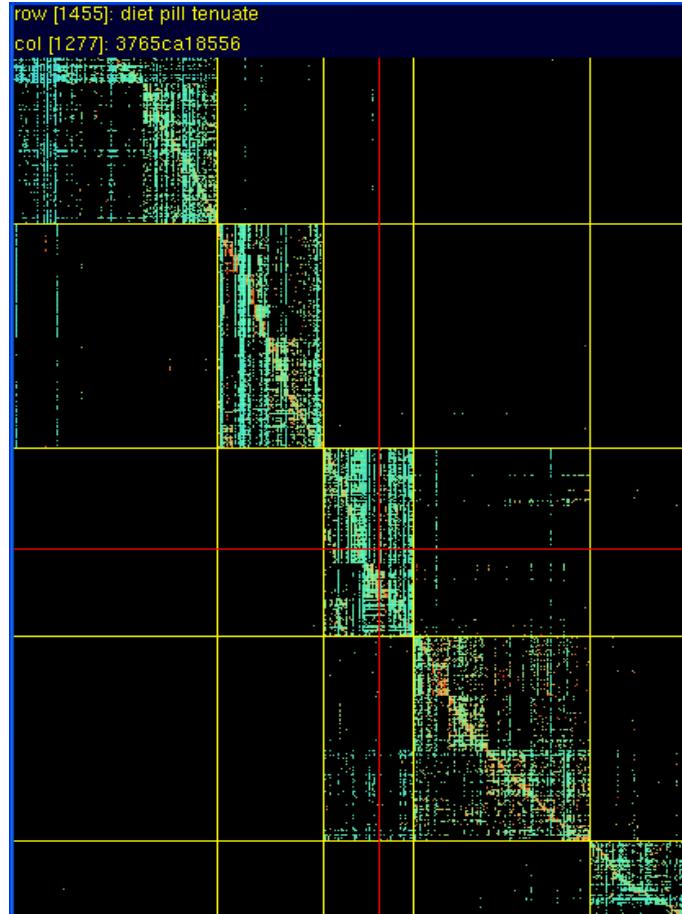


Back to data





Clusters





Sample clusters

'A.S.C. Incorporated'
'Atlantic Telecom'
'AudioLink'
'Cost Plus Electronics'
'Headset Express'
'Hello Direct'
'PDA Mountain'
'PK TECH INC'

.....

'accessory phone'
'cordless headset'
'cordless headset phone'
'free hands'
'headset'
'headset microphone'
'headset phone'
'headset plantronics'
'headset wireless'
'isdn'
'plantronics'

.....

'Berent Associates Center for
Shyness and Social Therapy'
'Dr. Puff'
'New York Psychotherapy
Collective'
'Ruby Shoes'
'Self-employed psychologist'
'www.kabbalah.com'

.....

'health mental'
'help self'
'improvement self'
'parenting'



Yahoo! Music – LAUNCHcast radio

player

The screenshot shows the Yahoo! Music LAUNCHcast plus player. At the top, it says "MUSIC LAUNCHcast plus" and "music that listens to you". Below that, there are tabs for "SONG INFO", "STATION DIRECTORY", and "HELP & OPTIONS". The "STATION" dropdown is set to "My Station". On the right, a user "Hi leonid68m" is logged in. In the center, a song is playing: "Pancho Piloncho" by Larry Harlow from the album "Larry Harlow's Latin Jive" (2003). There are rating options: "Can't get enough" (0 stars), "Like it" (1 star), and "Love it" (2 stars). Below the song info, there's a "BUY SONG" button and a note about buying the album on Yahoo! Shopping. At the bottom, there are playback controls (rewind, play/pause, forward) and volume controls. The audio quality is set to "Med". The status bar shows "Playing 02:42 / 10:22".

user profile

My Station > Edit >		Artists	1 - 20 (of 46 artists)
ARTISTS	ARTIST	YOUR RATING Why rate?	
	Gipsy Kings	0 ★★★★☆	
	Tito Nieves	0 ★★★★☆	
	Ricardo Montaner	0 ★☆☆☆☆	
	Eddie Palmieri	0 ★★★☆☆	
	Tito Puente	0 ★★★★☆	
	Sade	0 ★★★★☆	
	Santana	0 ★★★★☆	
	Vanessa-Mae	0 ★★★★☆	
	Ruben Blades	0 ★☆☆☆☆	
	Mary J. Blige	0 ★★★☆☆	



Launch cast data

user ID artist ID ratings

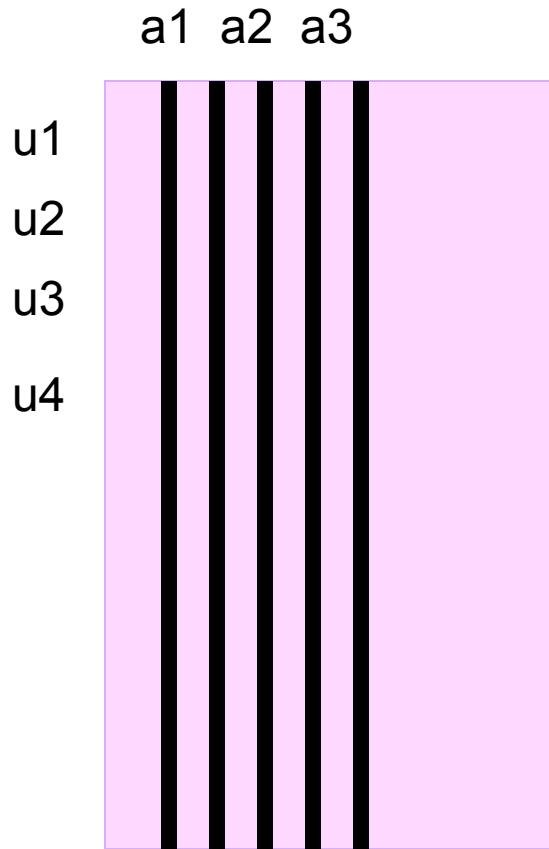
190389	1034047	100
190389	1034060	20
190389	1034129	50
190389	1034276	100
190389	1034388	80
190389	1034801	100
190389	1034831	100
190389	1034882	0
190389	1034883	40
190389	1035010	20
190389	1035473	60
190389	1035840	0
190389	1035926	30
190389	1036157	30
190389	1036454	60
190389	1036736	30
190389	1036737	100
190389	1036740	70
190389	1036746	80
190389	1036762	100

artist ID artist

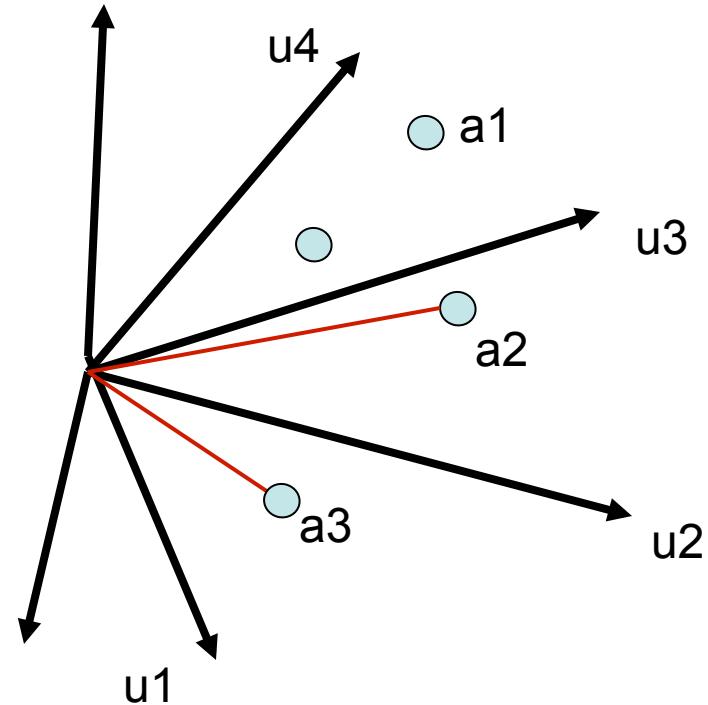
1037729	Coral
1037730	Rick Braun
1037731	Britney Spears
1037732	Gary Motley
1037733	Candido Fabre
1037734	K'Stalia Y Los Salchichas
1037735	Donny Hoffa
1037736	Rafael Mendez
1037737	Lithops
1037738	The Paris Ensemble
1037739	The Sunset Orchestra
1037740	The Mariachis Of Chiapas
1037741	Jenny Simpson
1037742	Doc West & The Yard Dogs
1037743	Geoff Bartley
1037744	Twilight Circus Dub Sound...
1037745	Tekneek



Ratings: data representation



ratings: users - artists



Vector space model:
 $w_{ij} = \cos(a_i, a_j) =$
 $(a_i \cdot a_j) / (\|a_i\| \|a_j\|)$

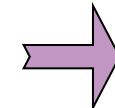


Similarity graph

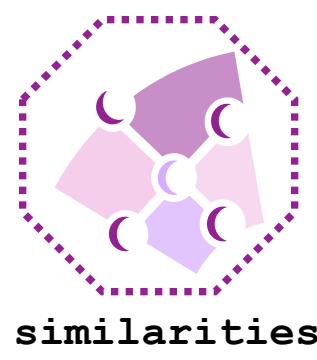
artists

users	0	3	1	0	4	5
0	1	0	0	0	0	5
1	4	0	0	0	0	0
2	2	0	4	0	0	0
3	0	5	0	0	0	5
4	1	0	3	0	0	4

Launch Ratings

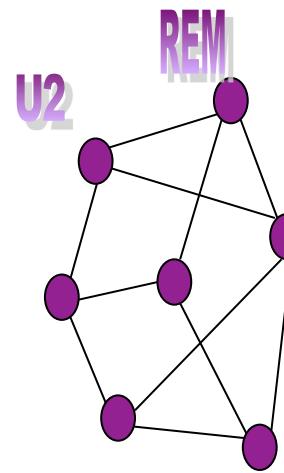
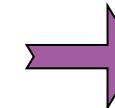


Compute



similarities

artists



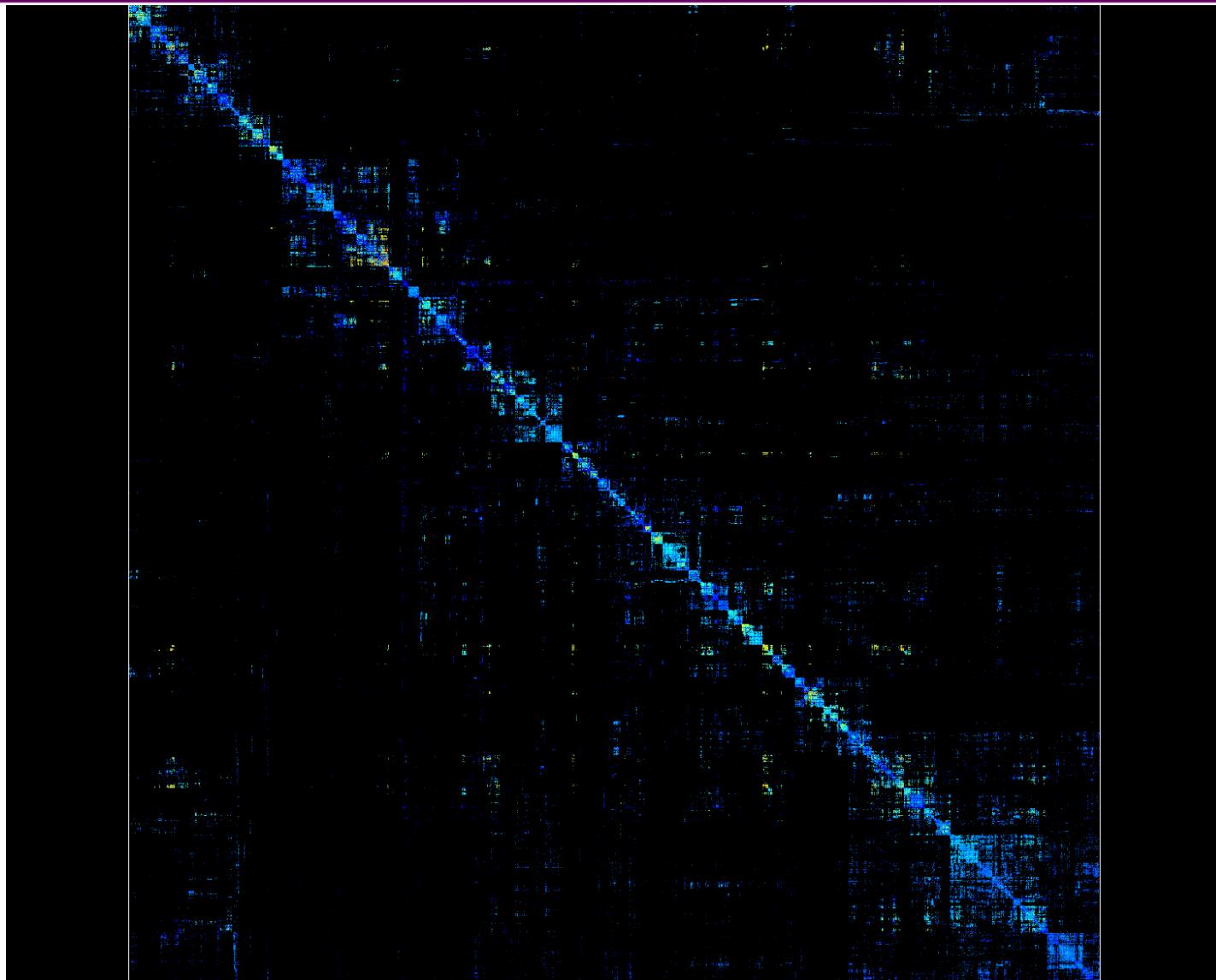
artists

1	0.3	0.8	0.0	0.5
0.3	1	0.4	0.0	0.2
0.8	0.4	1	0.3	0.0
0.0	0.0	0.3	1	0.0
0.5	0.2	0.0	0.0	1



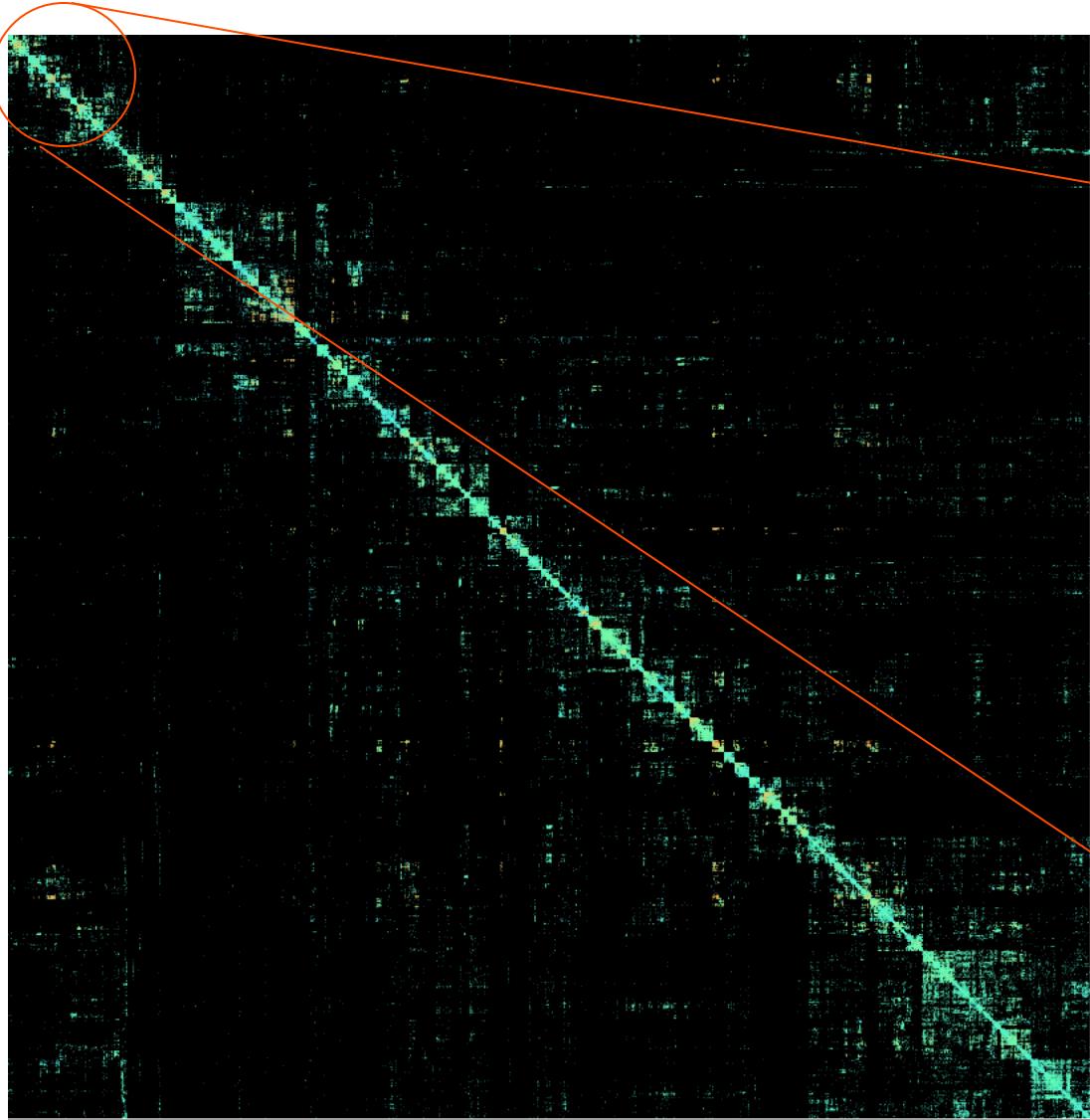


Yahoo! Music, artist-artist



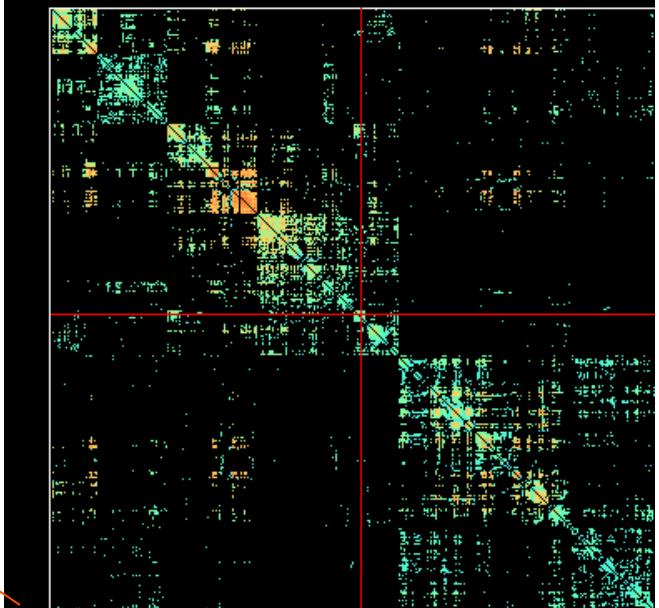


Artist clusters



row [8014]: Primal Fear

col [4509]: Jag Panzer





Sample clusters

cluster # 6

Rank	Label
1	Mozart
2	Royal Philharmonic Orchestra
3	Yo-Yo Ma
4	Scottish Chamber Orchestra
5	New York Philharmonic
6	Gilbert Johnson
7	New York Pro Music Antiqua
8	New Philharmonia Orchestra
9	Izzy
10	London Symphony Orchestra
11	Philadelphia Orchestra
12	Roberto Michelucci
13	Columbia Symphony Orchestra
14	San Francisco Symphony
15	London Philharmonic Orchestra
16	Berlin Philharmonic Orchestra

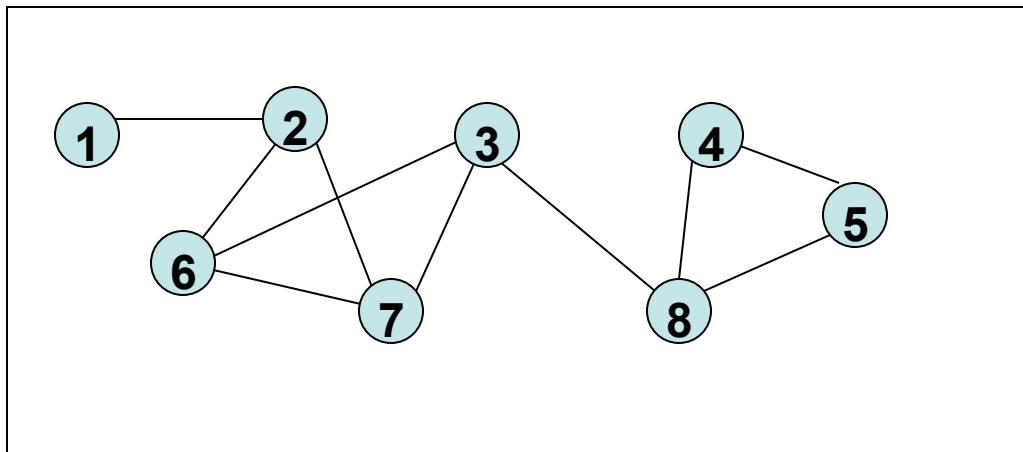
cluster # 79

Rank	Label
1	Enrique Iglesias
2	Avril Lavigne
3	Backstreet Boys
4	Kylie Minogue
5	Good Charlotte
6	Simple Plan
7	T.A.T.U.
8	Hilary Duff
9	Paula Abdul
10	No Doubt
11	Britney Spears
12	Westlife
13	O-Town
14	LFO (Lyte Funky Ones)
15	98 Degrees
16	Shakira



Ordering: graph embedding in 1D

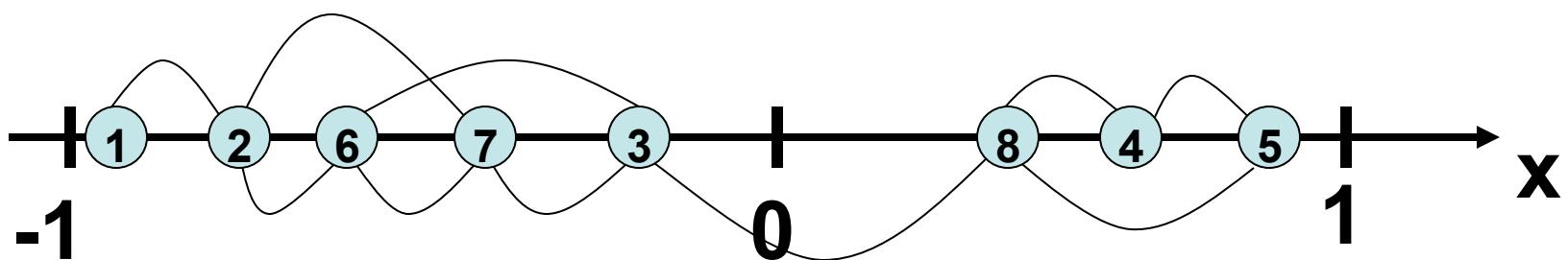
- every node – coordinate $x(i)$



$$E = \frac{1}{2} \sum_{i,j} w_{ij} (x_i - x_j)^2$$

$$\sum_i x_i^2 = N$$

$$\sum_i x_i = 0$$



create ordering/permutation



Graph embedding : 2D

- Quadratic optimization problem

$$E(x) = \frac{1}{2} \sum_{i,j} w_{ij} (x_i - x_j)^2 + \lambda \sum_i (x_i^2 - N)$$

- Replace $x \rightarrow r = (x, y)$ $r_i^2 = x_i^2 + y_i^2$
 $r_i r_j = x_i x_j + y_i y_j$
- Quadratic optimization problem

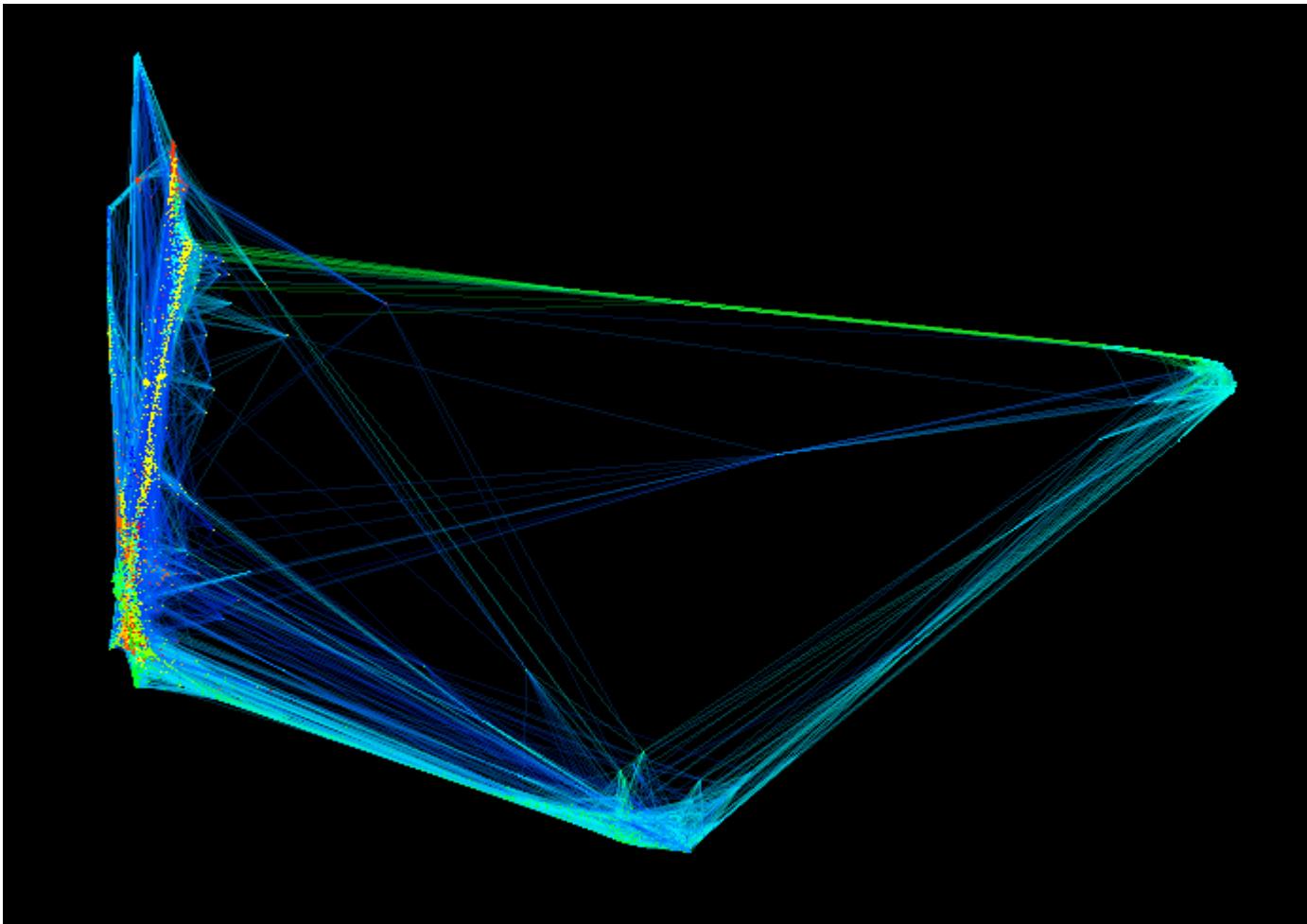
$$E(r) = \frac{1}{2} \sum_{i,j} w_{ij} (r_i - r_j)^2 + \lambda \sum_i (r_i^2 - N)$$

$$E(r) \sim E(x) + E(y)$$

- Embedding: $x = \text{ev}(2)$, $y = \text{ev}(3)$

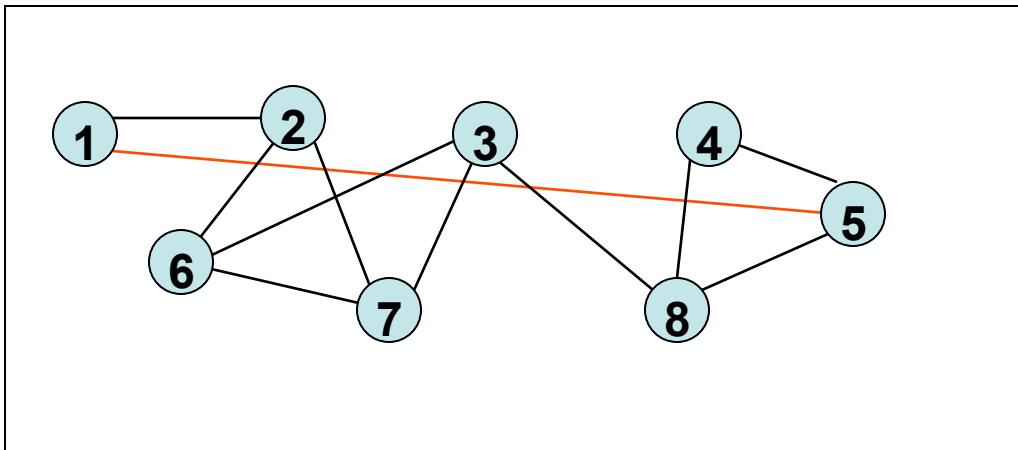


2D spectral embedding

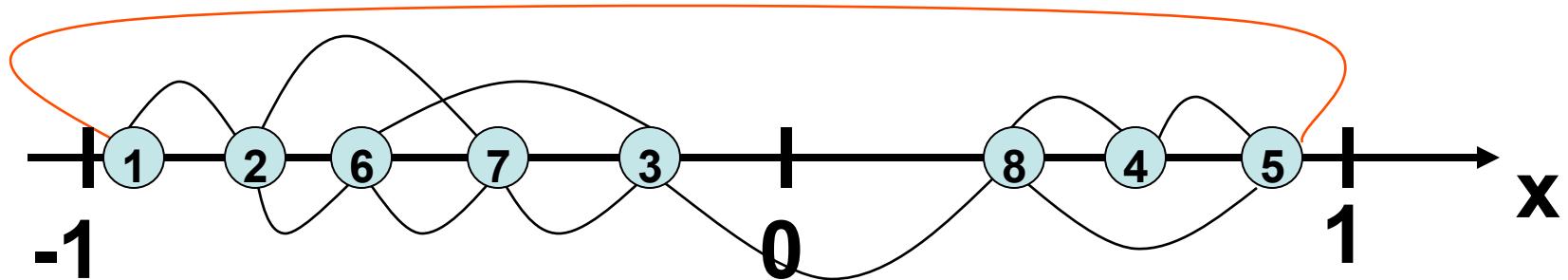




intuition



$$E = \frac{1}{2} \sum_{i,j} w_{ij} (x_i - x_j)^2$$





Embedding 2D - better

- Change constraints: average \rightarrow “per node”

$$\frac{1}{N} \left(\sum_i x_i^2 \right) = 1 \quad \rightarrow \quad r_i^2 = x_i^2 + y_i^2 = 1$$

- Optimization:

$$E(r) = \frac{1}{2} \sum_{i,j} w_{ij} (r_i - r_j)^2 + \lambda \sum_i (r_i^2 - N)$$



$$E(r) = \frac{1}{2} \sum_{i,j} w_{ij} (r_i - r_j)^2 + \sum_i \lambda_i (r_i^2 - 1)$$



Semi Definite Programming (SDP)

- Standard form primal SDP

$$\min(C \cdot X), \quad A_i \cdot X = b_i, \quad X \geq 0$$

- Minimum bisection SDP relaxation

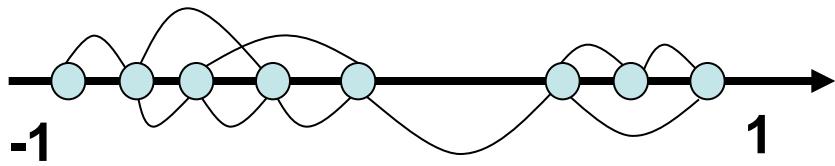
$$\min(L \cdot X), \quad \text{diag}(X) = e, \quad e^T X e = 0, \quad X \geq 0$$

- Factor:

$$X = RR^T$$

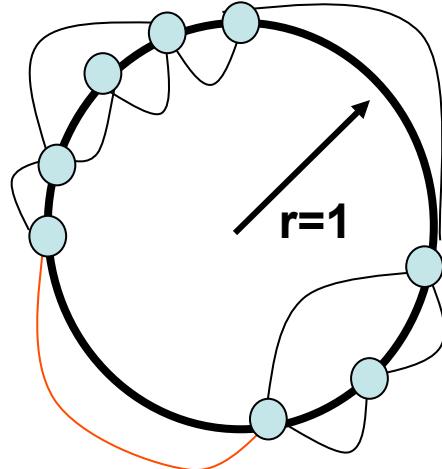


SDP embedding



- Spectral:

$$\min \left(\frac{\mathbf{x}^T \mathbf{L} \mathbf{x}}{4} \right), \quad x \in \mathbf{R}^1;$$
$$\sum x_i^2 = N, \quad (\mathbf{x}^T \mathbf{e}) = 0$$

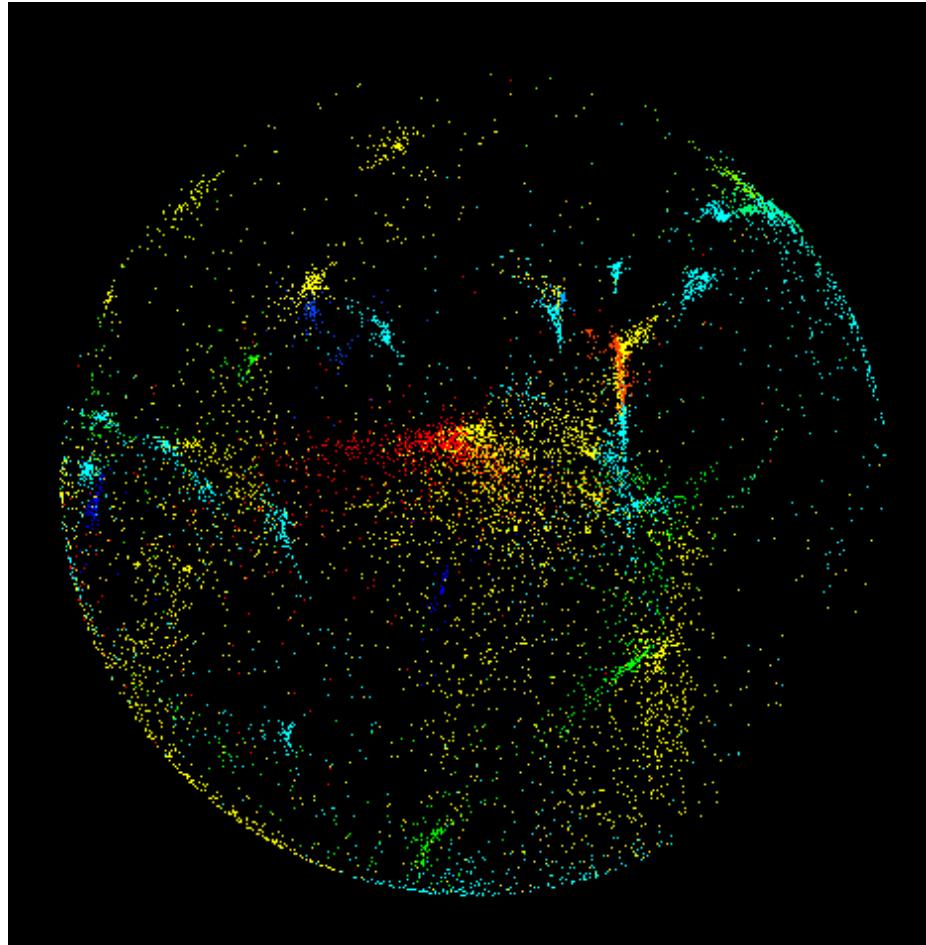


- SDP:

$$\min \left(Tr \frac{\mathbf{R}^T \mathbf{L} \mathbf{R}}{4} \right), \quad R_i \in \mathbf{R}^N;$$
$$\|R_i\| = 1, \quad (\mathbf{R}^T \mathbf{R}) \geq 0$$

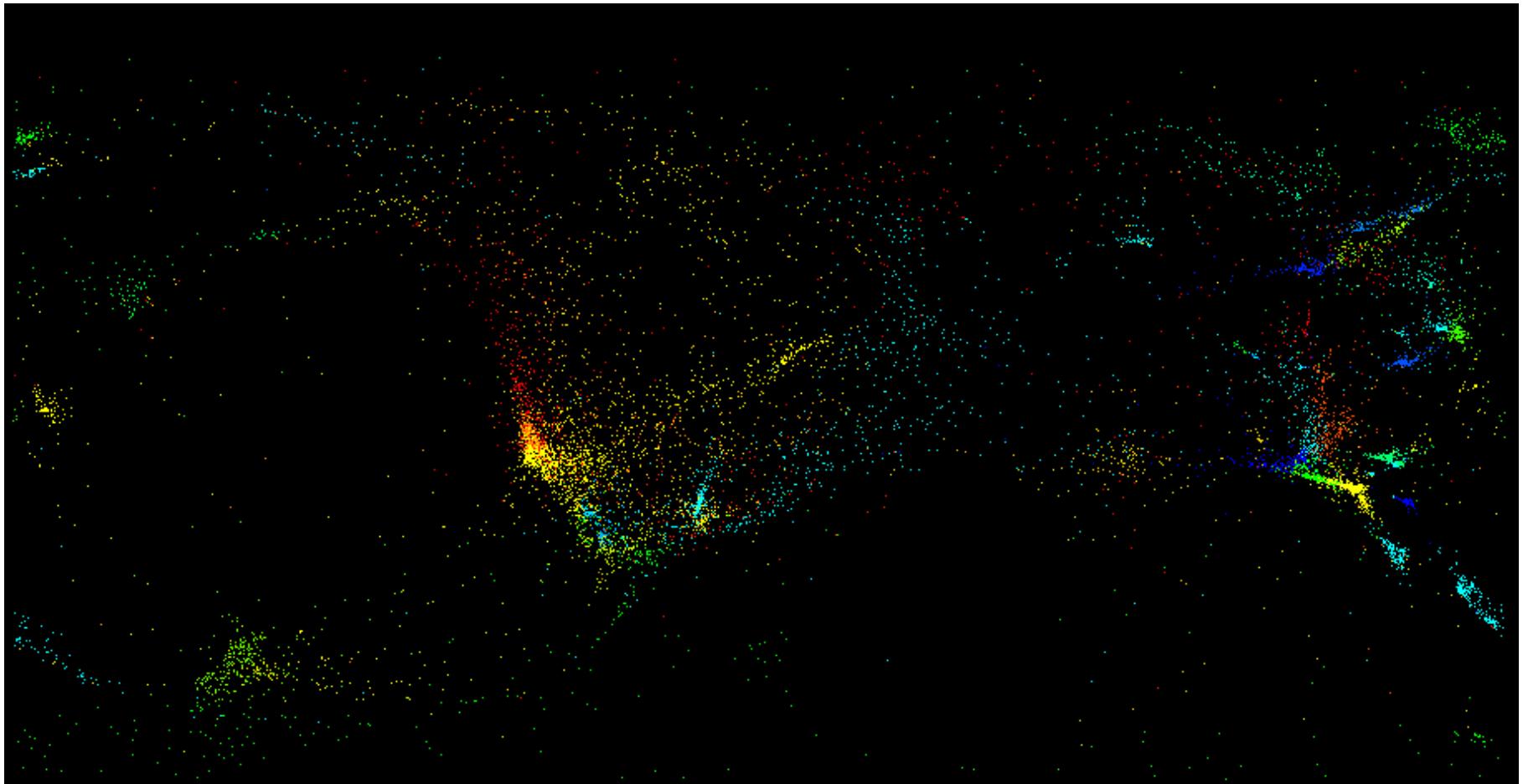


SDP spherical embedding



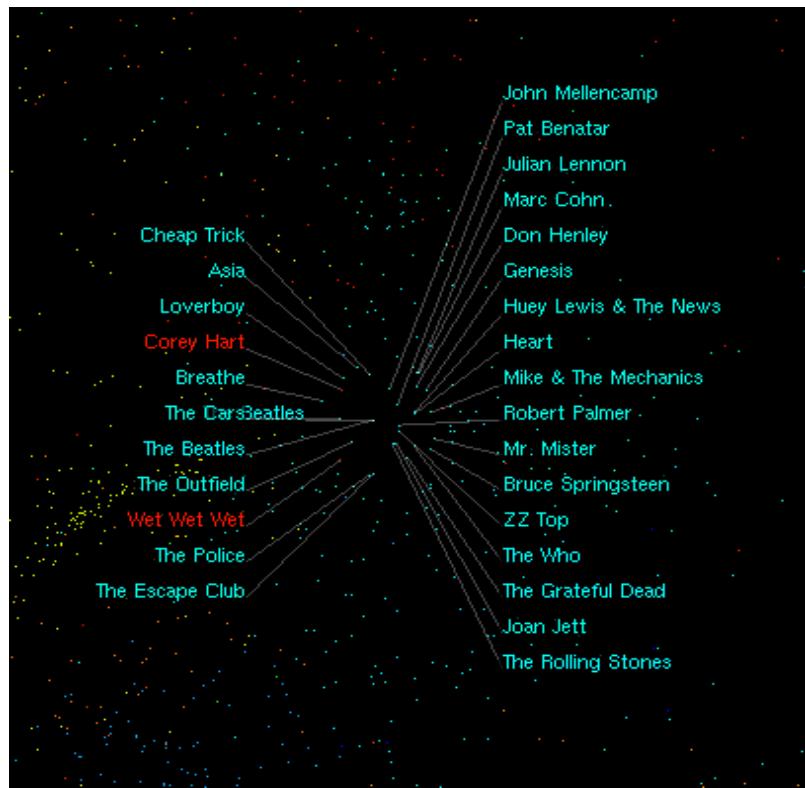
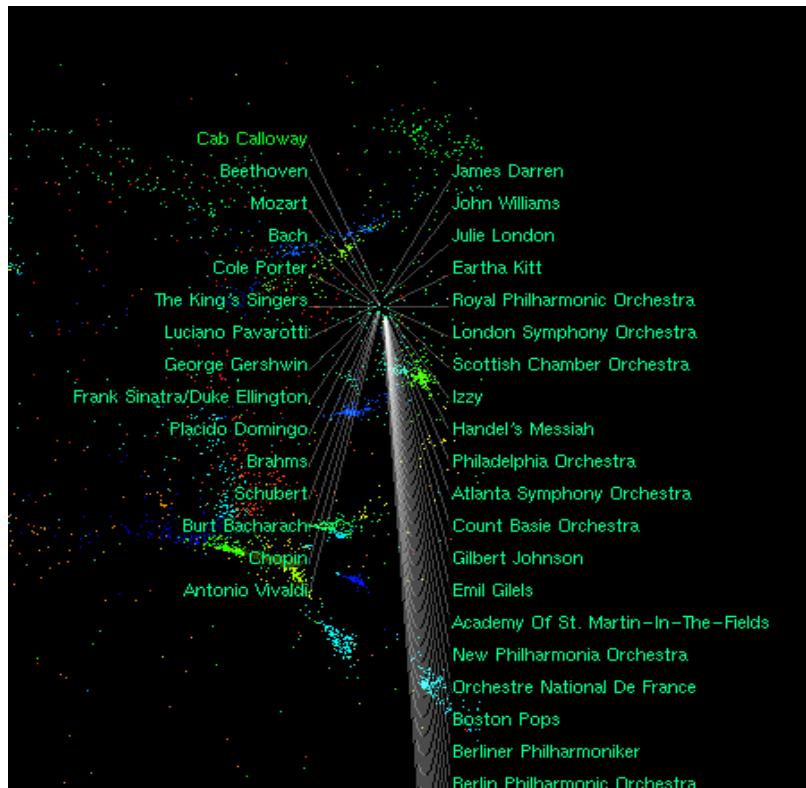


SDP sphere unrolled





Clusters

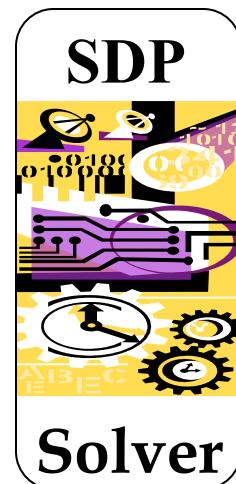
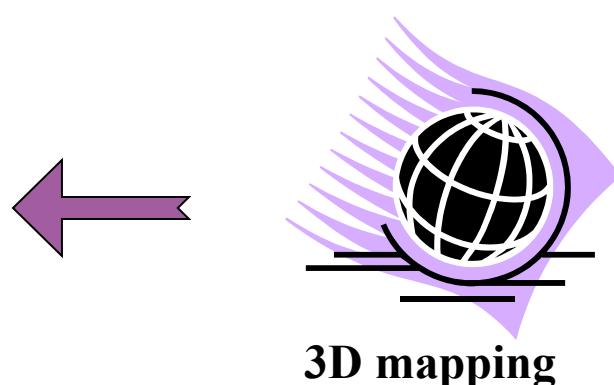
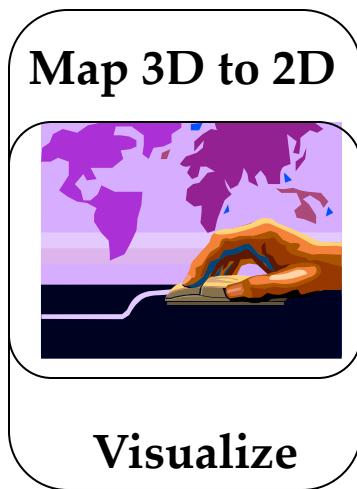
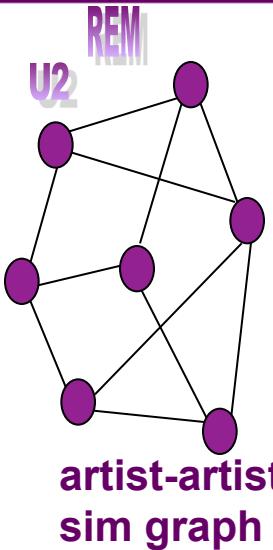
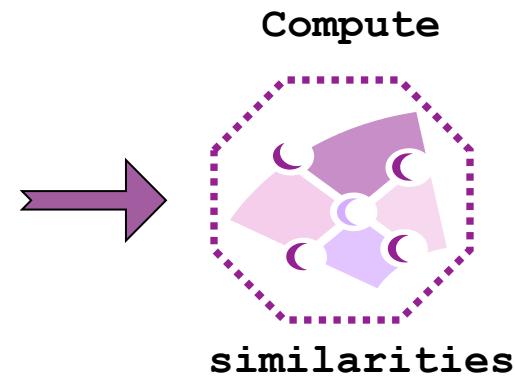


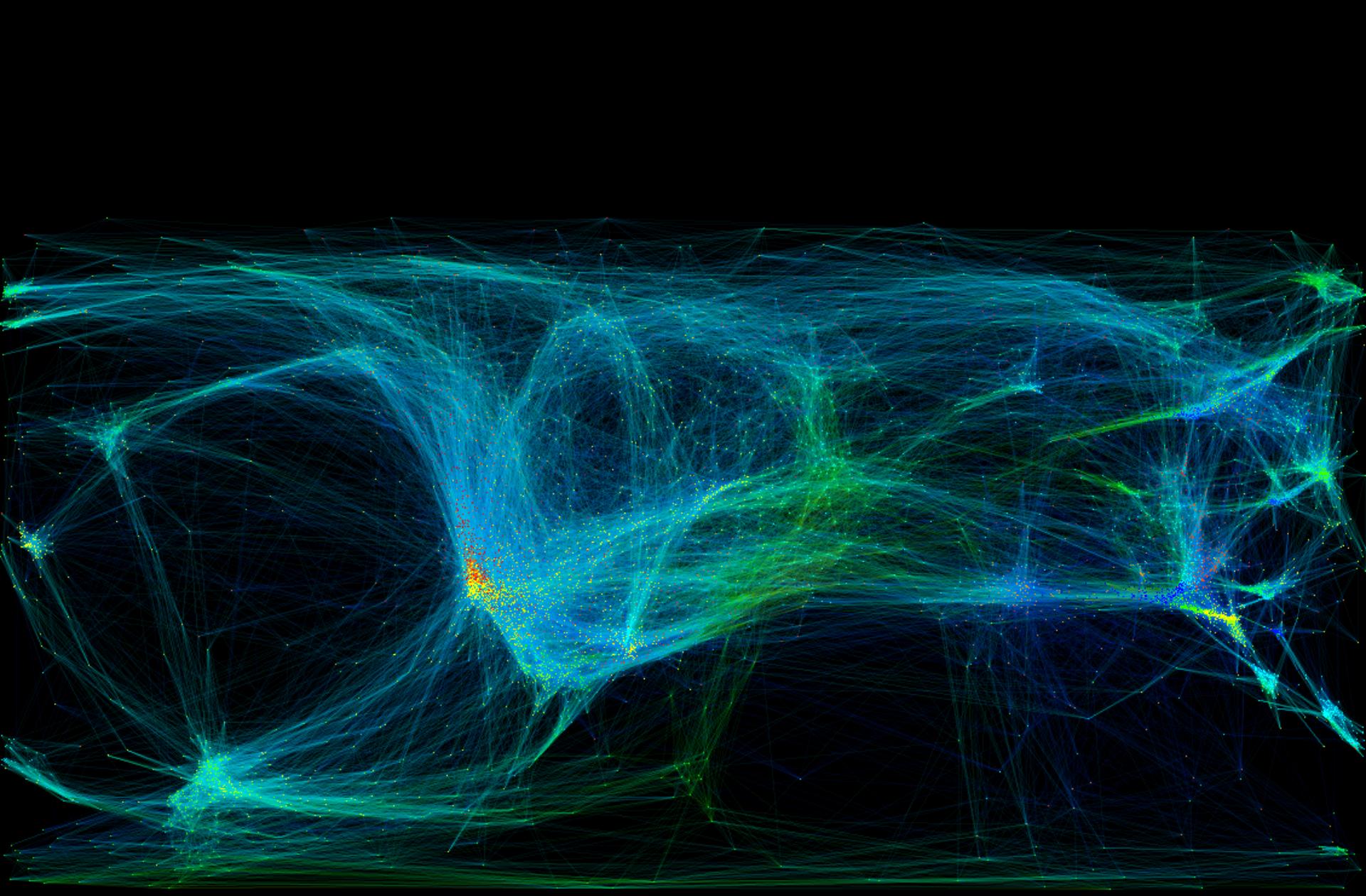


Visualization pipeline

users	artists	0	3	1	0	4	5
	0	1	0	0	0	5	
	1	0	0	0	0	5	
	0	4	0	0	0	0	
	2	0	2	0	4	0	
	0	5	0	0	0	5	
	1	0	3	0	0	4	

Launch Ratings





The World of Music: SDP layout of high-dimensional data

YAHOO!

