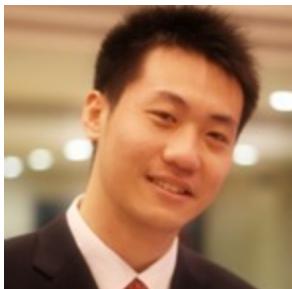




KDD 2017

Halifax, Nova Scotia - Canada  
August 13 - 17, 2017

# Tutorial: Data-Driven Approaches towards Malicious Behavior Modeling



Meng Jiang  
University of Notre Dame



Srijan Kumar  
Stanford University



Christos Faloutsos  
Carnegie Mellon University



V.S. Subrahmanian  
University of Maryland, College Park

Tutorial link: <http://bit.ly/kdd2017>

# Outline

## Introduction

Feature-based algorithms

Sockpuppets

Spectral-based algorithms

Vandals

Hoaxes

Density-based algorithms

Visualization: “spokes”, “blocks”, “staircases”

Camouflage

Theoretical guarantee

Ill-gotten Likes

Synchronized Behaviors

Advertising campaigns

Social spam

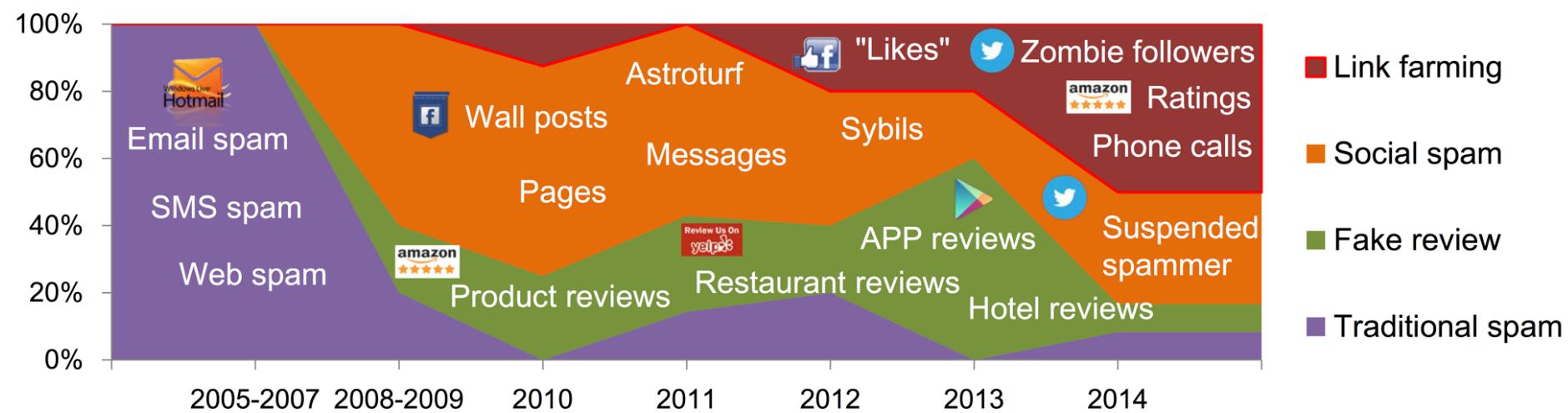
Conclusions and future directions

Tutorial link: <http://bit.ly/kdd2017>

# Suspiciousness and Density

- What is worth of inspections in our data?
- Suspiciousness: tending to cause or excite suspicion
  - Unexpected high density
  - ...
- Suspicious **density** of user behaviors in applications
  - Ill-gotten Likes: Facebook, etc.
  - Zombie followers: Twitter, Weibo, etc.
  - Social spam/fake reviews: Twitter, Weibo, Amazon, etc.
  - Advertising campaigns: Twitter, Weibo, etc.

# ... Social Spam, Social Link Farming

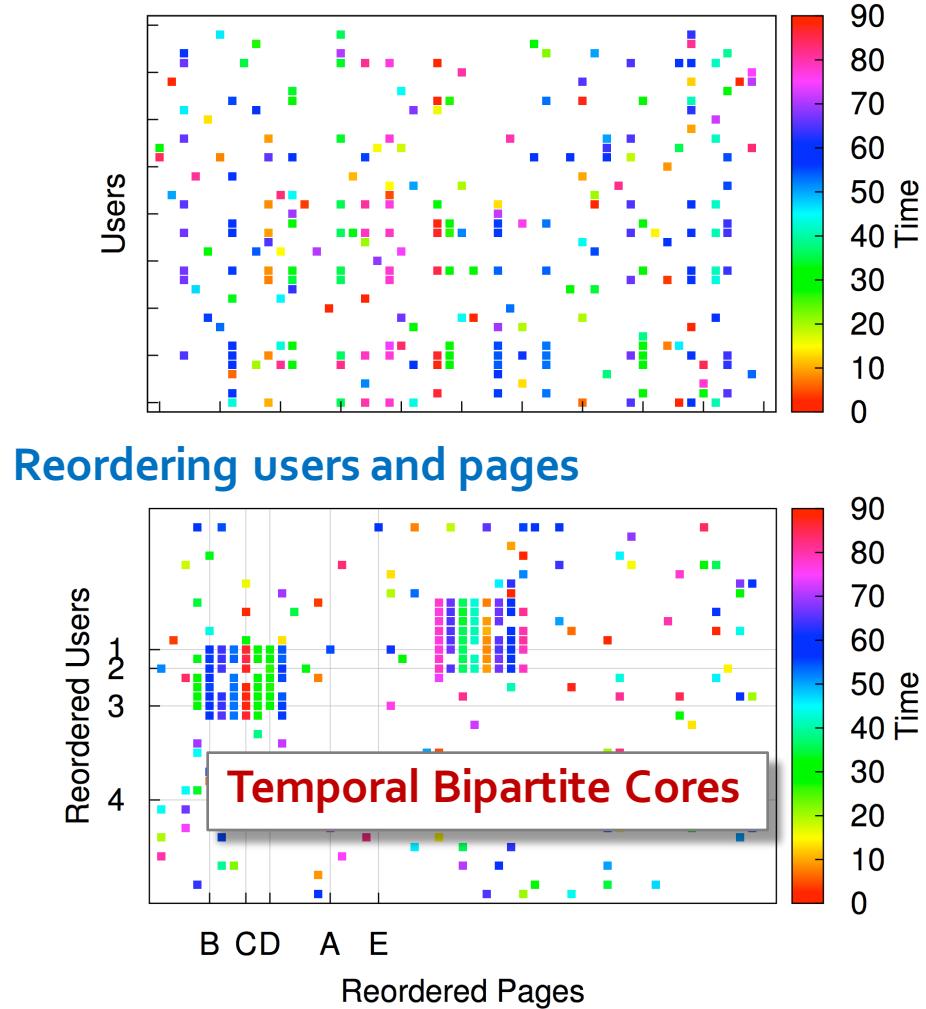
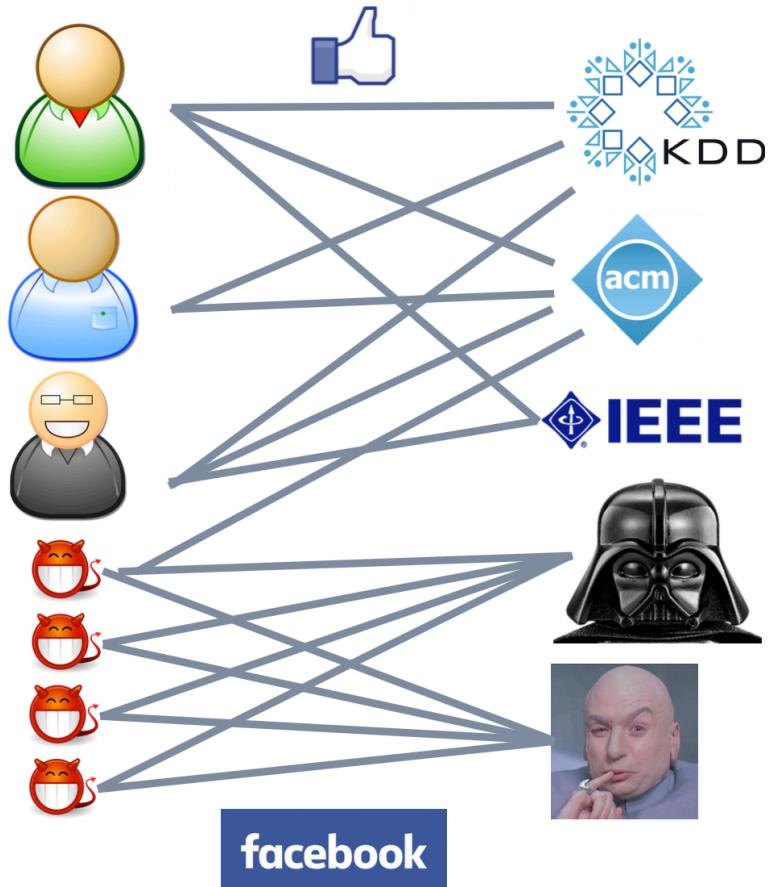


Meng Jiang, Peng Cui, and Christos Faloutsos. "Suspicious behavior detection: current trends and future directions." **IEEE Intelligent Systems**, 2016. (Survey paper)

# 1. Ill-gotten Facebook Likes

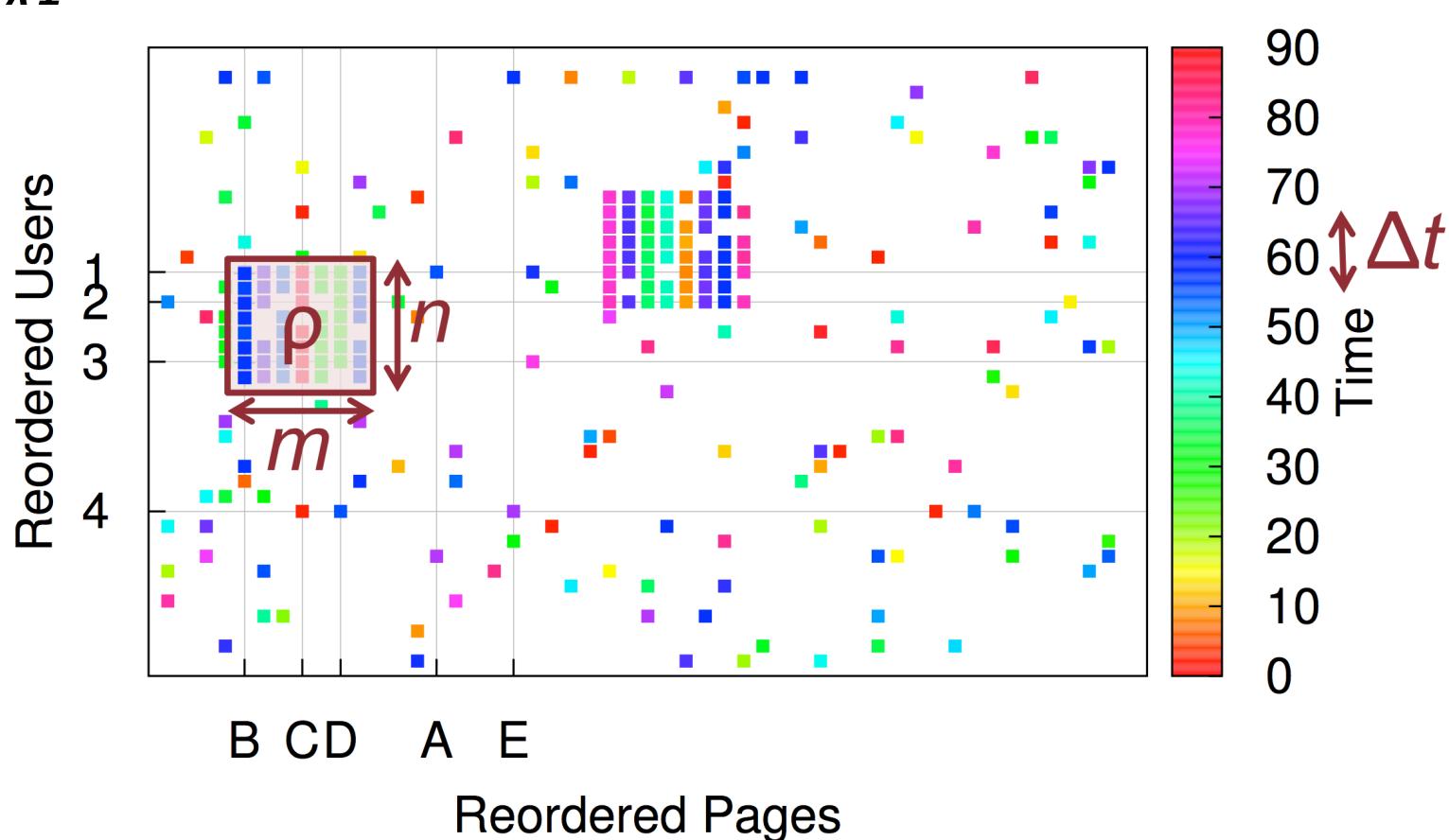
25,000 Facebook Likes	50,000 Facebook Likes	100,000 Facebook Likes	200,000 Facebook Likes
\$265	\$525	\$1,000	\$1,750
Lifetime Replacement Warranty	Lifetime Replacement Warranty	Lifetime Replacement Warranty	Lifetime Replacement Warranty
Dedicated 24/7 Customer Service			
100% Risk Free, Try Us Today			
Order starts within 24 - 48 hours			
Order completed within 22 days	Order completed within 35 days	Order completed within 35 days	Order completed within 35 days

# Density in Temporal Bipartite Graph

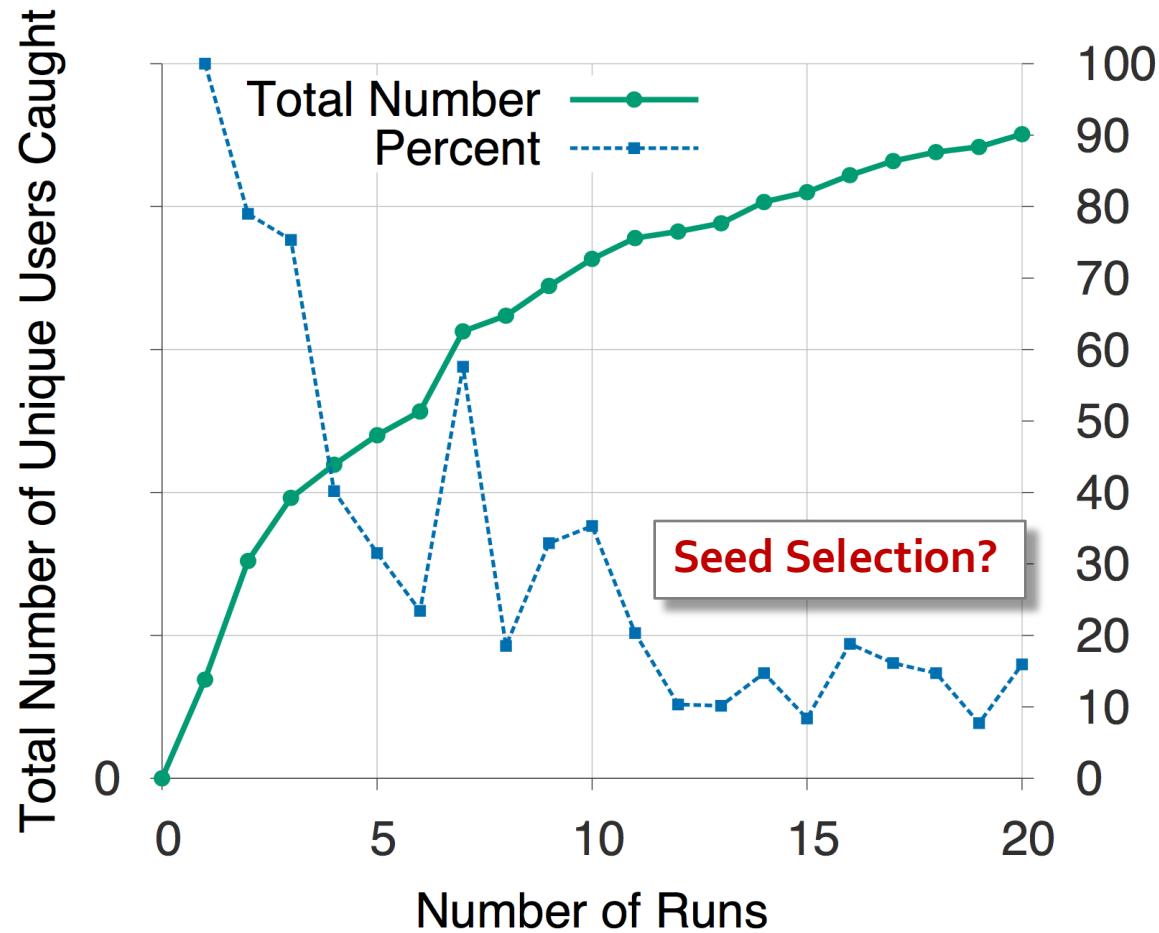


# CopyCatch: Seed + Search

- “Temporal Bipartite Core”:  $n$  users,  $m$  pages,  $\rho$ ,



# CopyCatch: Deployed in Facebook



A: One option: Blocks/Staircases caught by LockInfer  
(spectral-based method).

## 2. Twitter's Zombie Followers



# Catching Zombie Followers



engineers



product managers

Knowledge  
from  
manual  
inspection:

#followees,  
#followers, #tweets,  
#hashtags, #urls...

Learning models (classifiers)



Poor accuracy  
**(serious complaints** from users)

Fake account detection [Egele and Stringhini et al. NDSS'13; Yang and Wilson et al. TKDD'14; Viswanath and Bashir et al. USENIX Security Symposium'14]

# Is this account a zombie follower???

Aisling Walsh  
@xAsherz

Joined April 2009

[Tweet to Aisling Walsh](#)

Who to follow · Refresh · View all

- John Legere @JohnLe...  
[Follow](#)  
Promoted
- Dong Zhou @dongz9  
Followed by Peng Wang 王鹏 and others  
[Follow](#)
- Justin Zeus @askzy9  
Followed by Ruizhe, LI and others  
[Follow](#)

Find friends

Trends · Change

- #ThatsContinental  
Allowing curiosity to chart your course.  
Promoted by LincolnMotorCompany
- #2017in3words  
26.1K Tweets
- #nationalbaconday  
5,915 Tweets
- #NewYearsEveEve  
0 Follower

FOLLOWING 20 FOLLOWERS 3 0 tweet

Rachel Maddow MSN... @maddow I see political people... (Retweets do not imply endorsement.)

Trent Reznor @trent\_reznor Nine Inch Nails, How To Destroy Angels and other things.

Guardian Tech @guardiantech

richard bacon @richardpbacon

Jason Sweeney @sween limited edition, macaroni and glitter on construction paper.

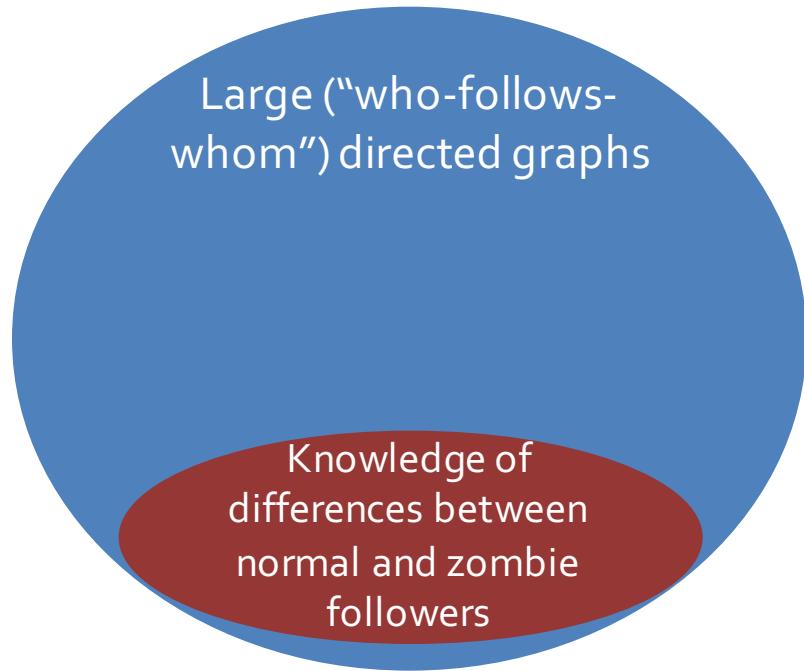
Terry Moran @TerryMoran Chief Foreign Correspondent, ABC News.

Hoppy New Year @markhoppus person

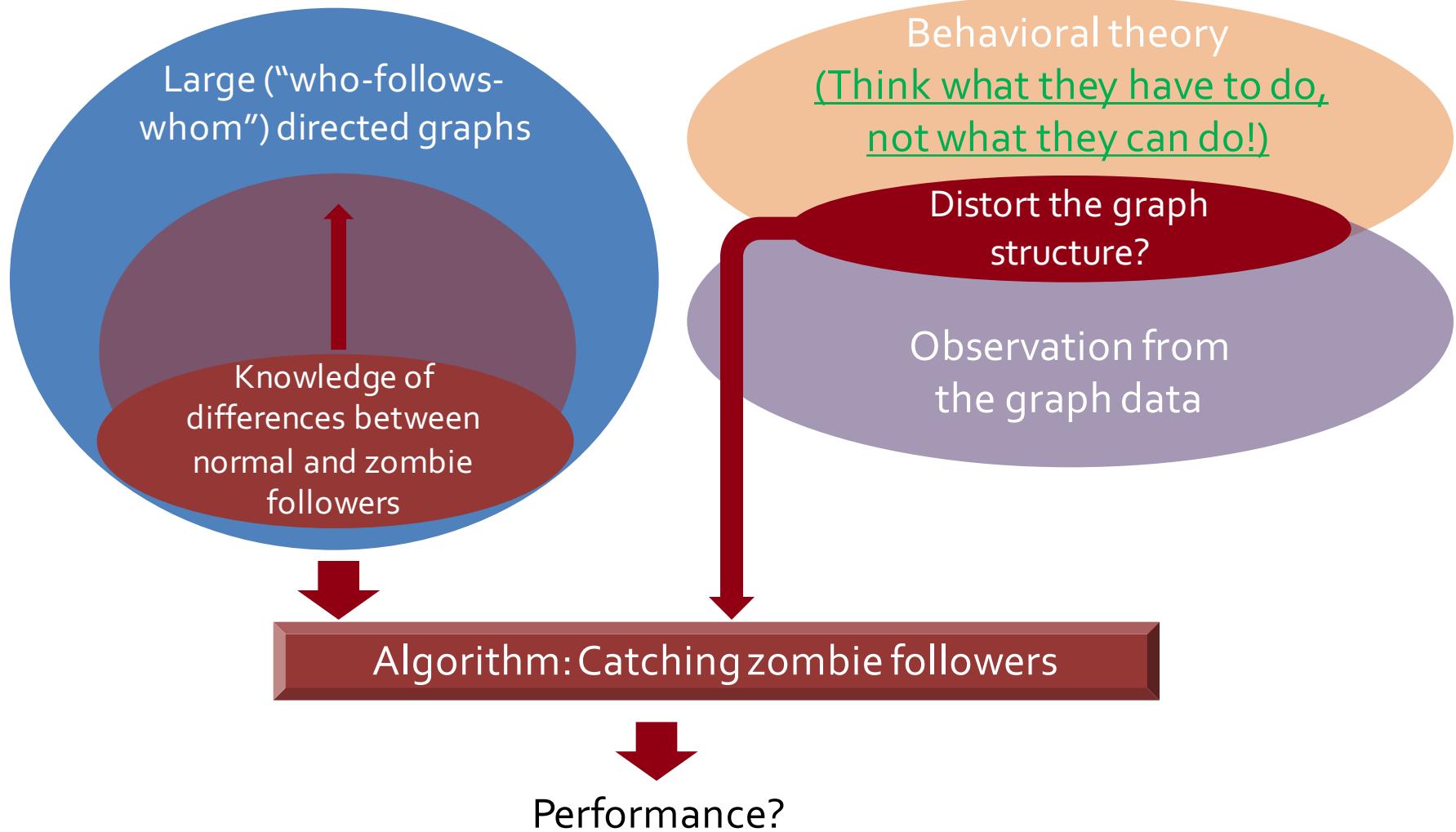
woot.com @woot Check out who we're following for other Woot accounts, and follow us on Facebook for extra excitement: [facebook.com/woot](http://facebook.com/woot)

CBOE @CBOE

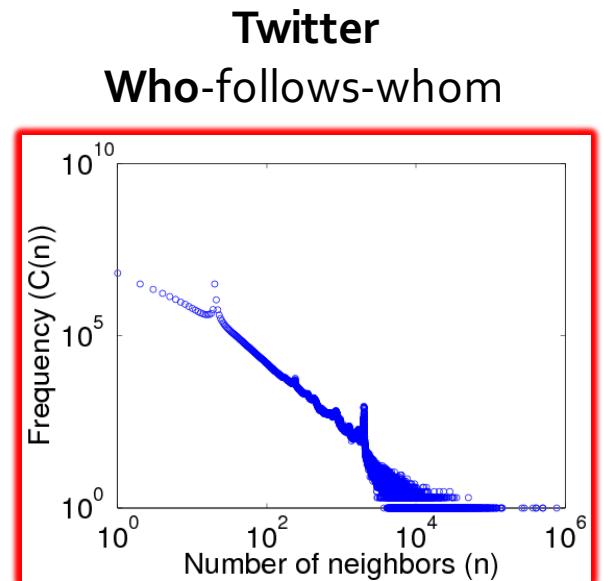
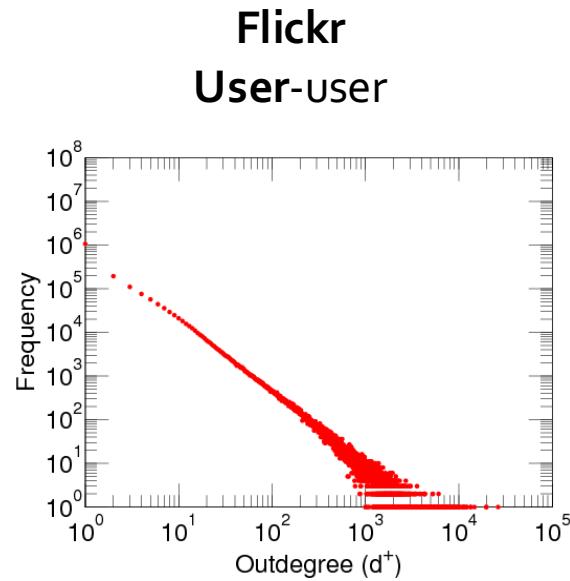
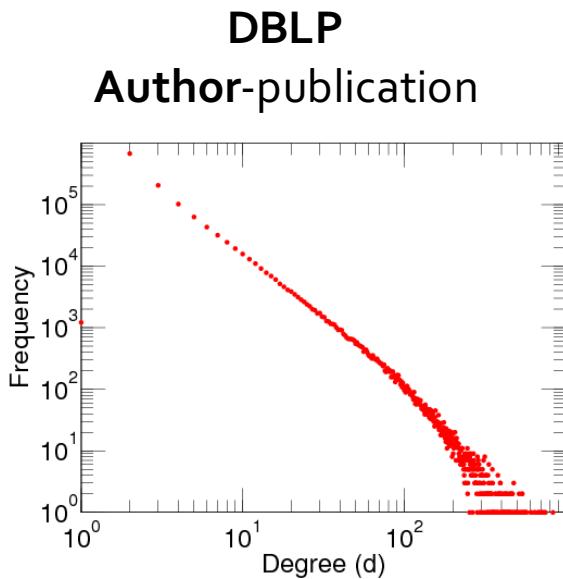
# “Density” in Directed Graph



# “Density” in Directed Graph



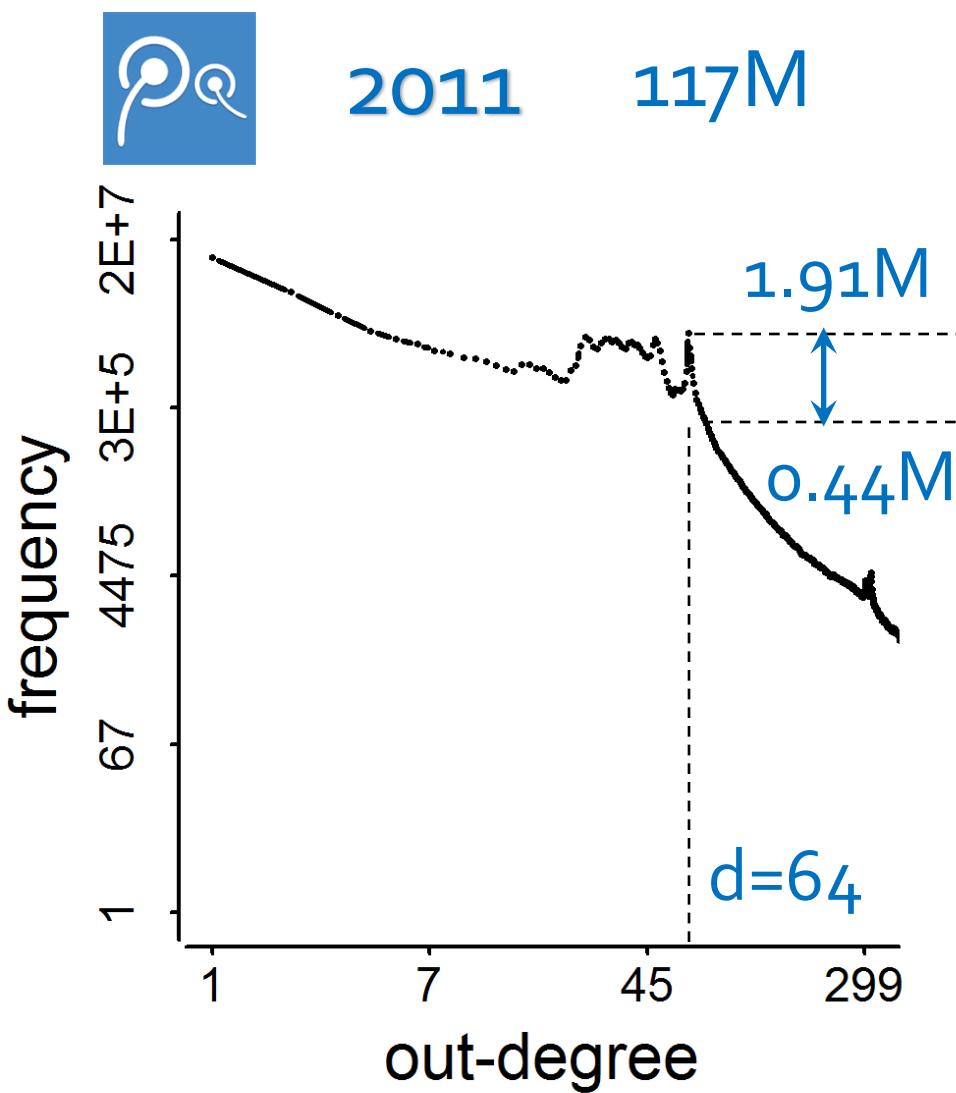
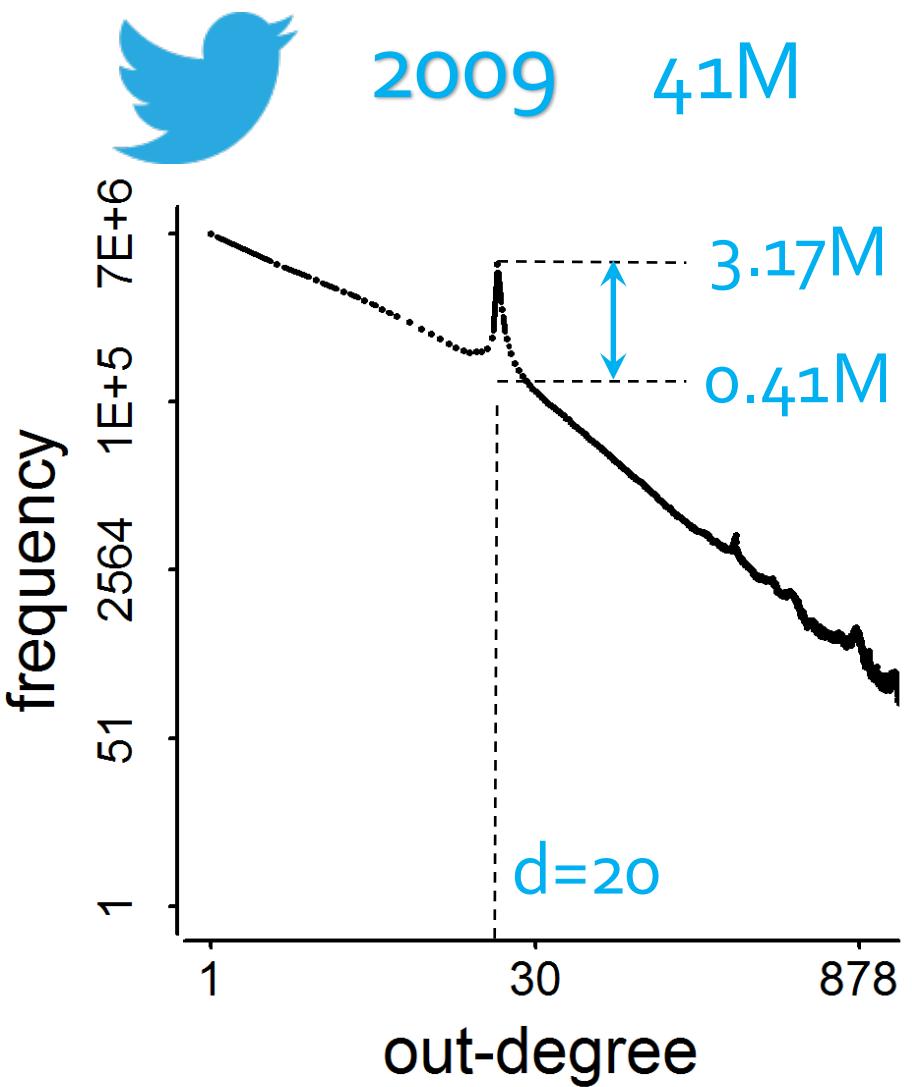
# Out-Degree Distributions: Power Law Expected



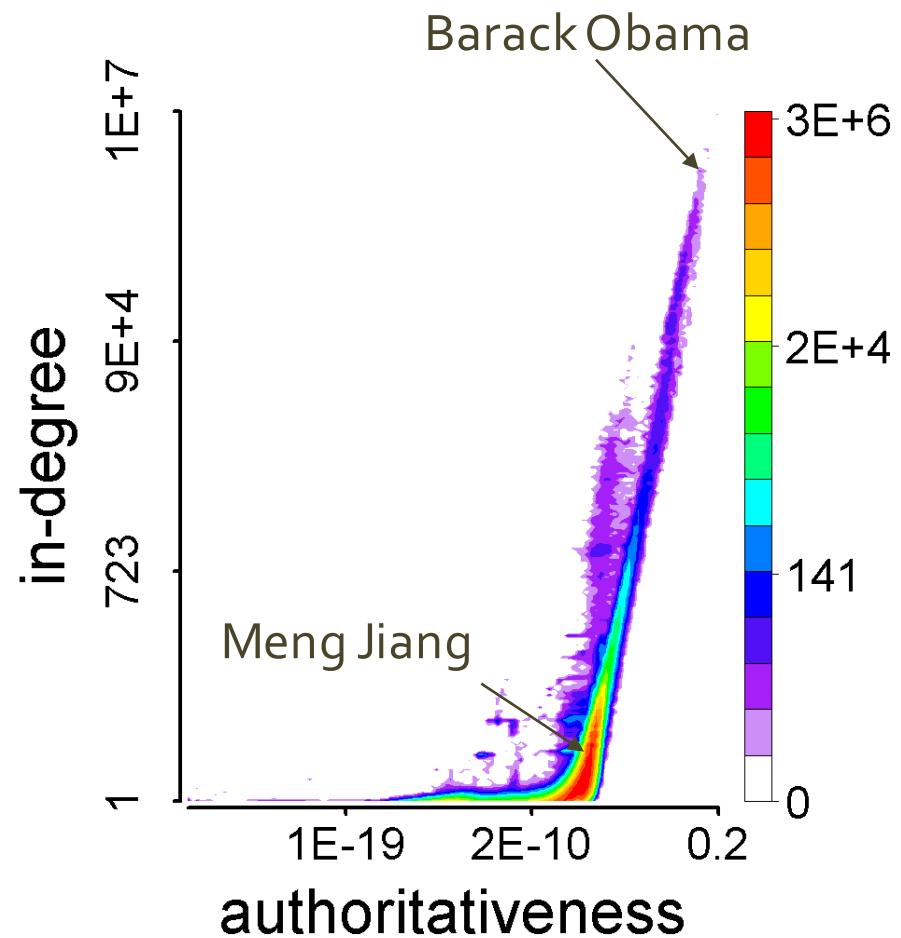
[[konect.uni-koblenz.de/networks/](http://konect.uni-koblenz.de/networks/)]

Power-law distributions in networks

# Spikes!

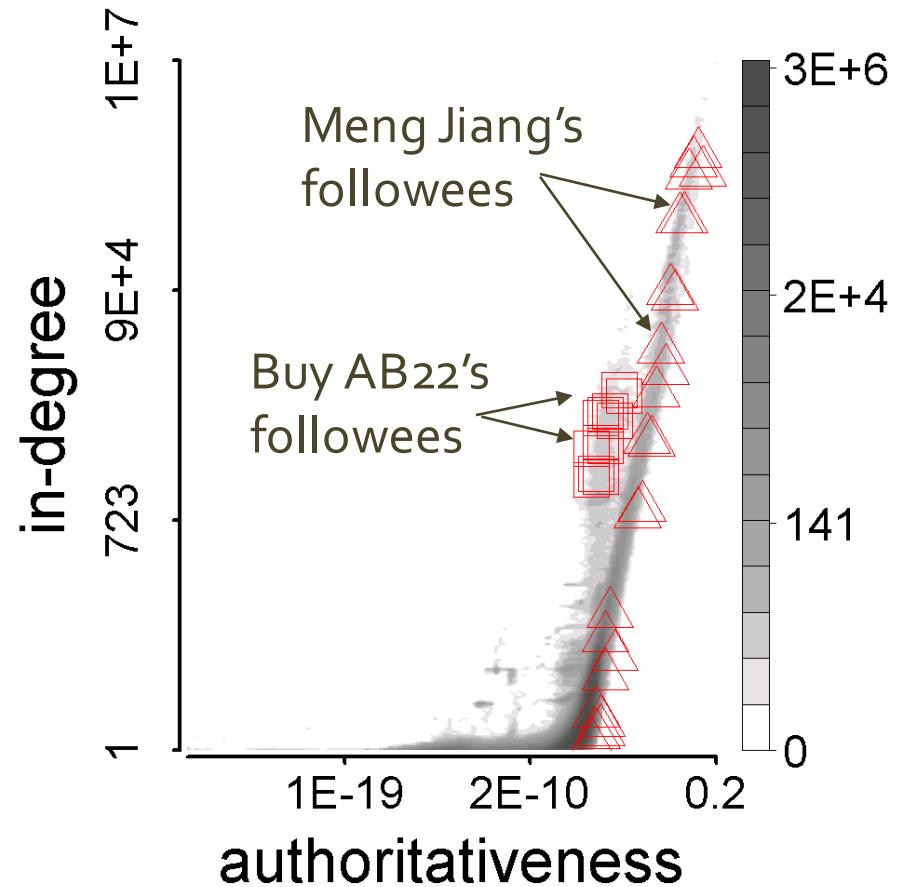


# How We/They Connect to Our/Their Followees

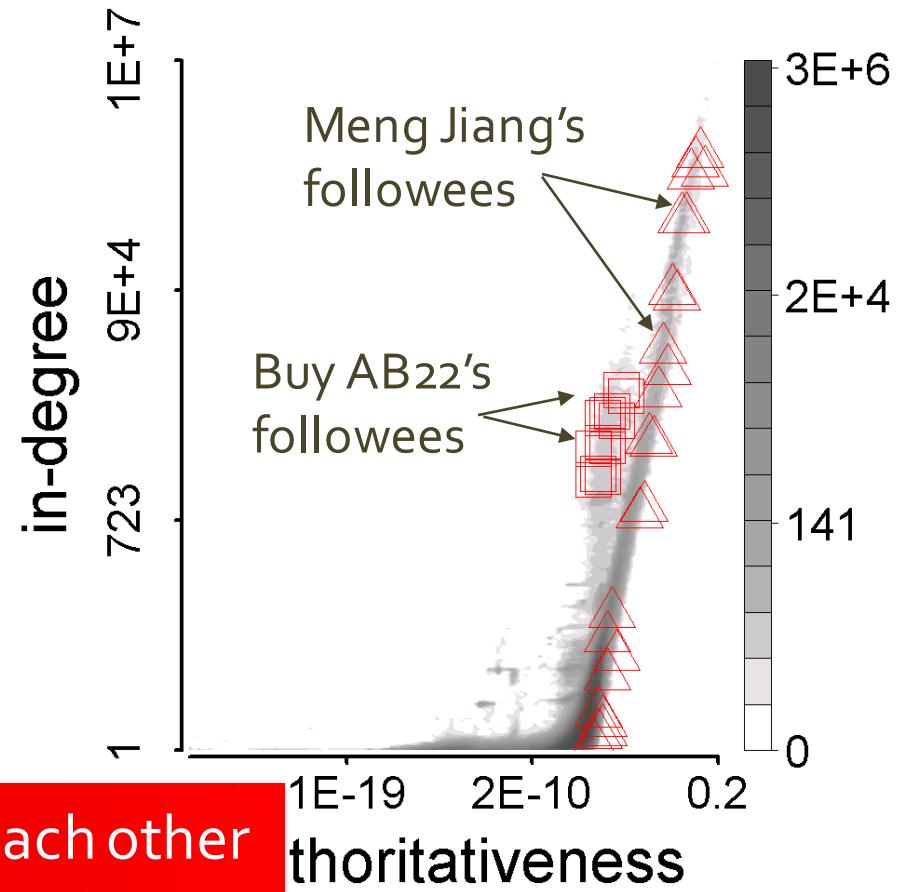


The HITS algorithm. Kleinberg. “Authoritative sources in a hyperlinked environment.” JACM’99.

# How We/They Connect to Our/Their Followees



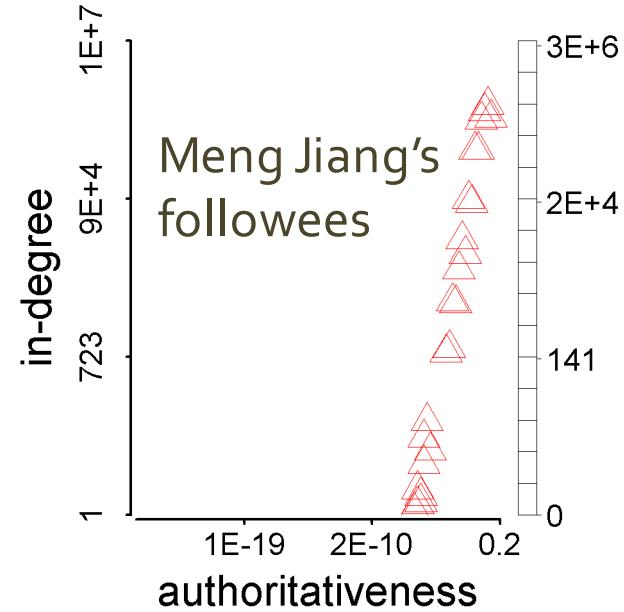
# How We/They Connect to Our/Their Followees



Synchronized: too similar with each other  
Abnormal: too different from the majority

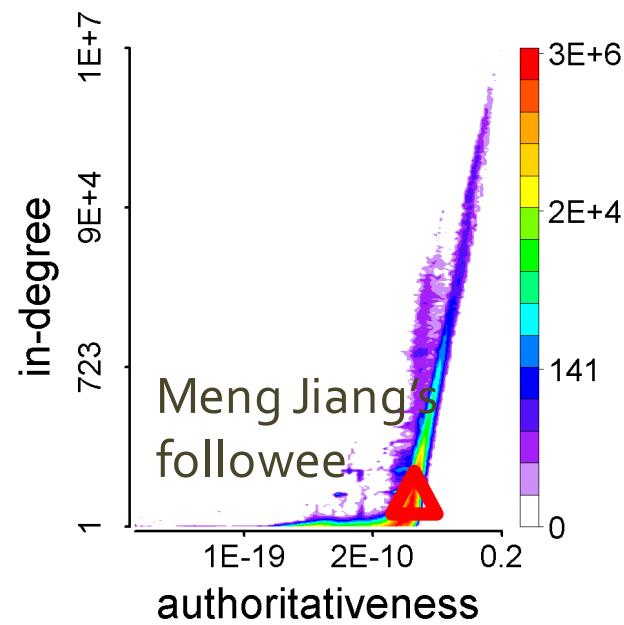
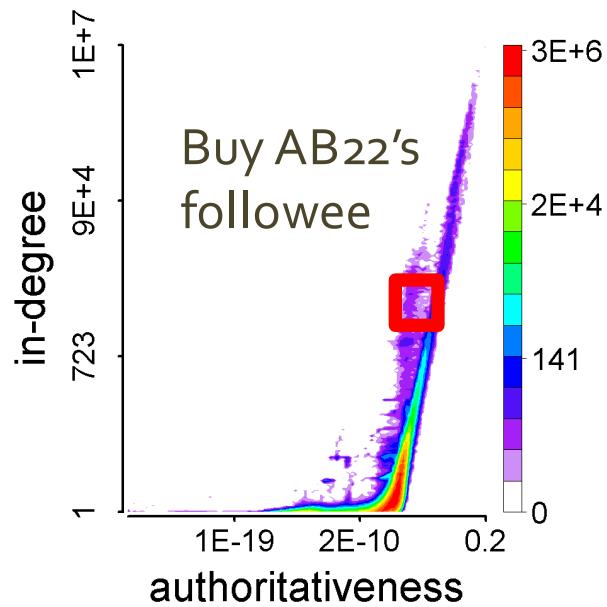
# Definition: Synchronicity

$$sync(u) = \frac{\sum_{(v, v') \in \mathcal{F}(u) \times \mathcal{F}(u)} \mathbf{p}(v) \cdot \mathbf{p}(v')}{d(u) \times d(u)}$$



# Definition: Normality

$$norm(u) = \frac{\sum_{(v,v') \in \mathcal{F}(u) \times \mathcal{U}} \mathbf{p}(v) \cdot \mathbf{p}(v')}{d(u) \times N}$$



# When is the Synchronicity Too High?

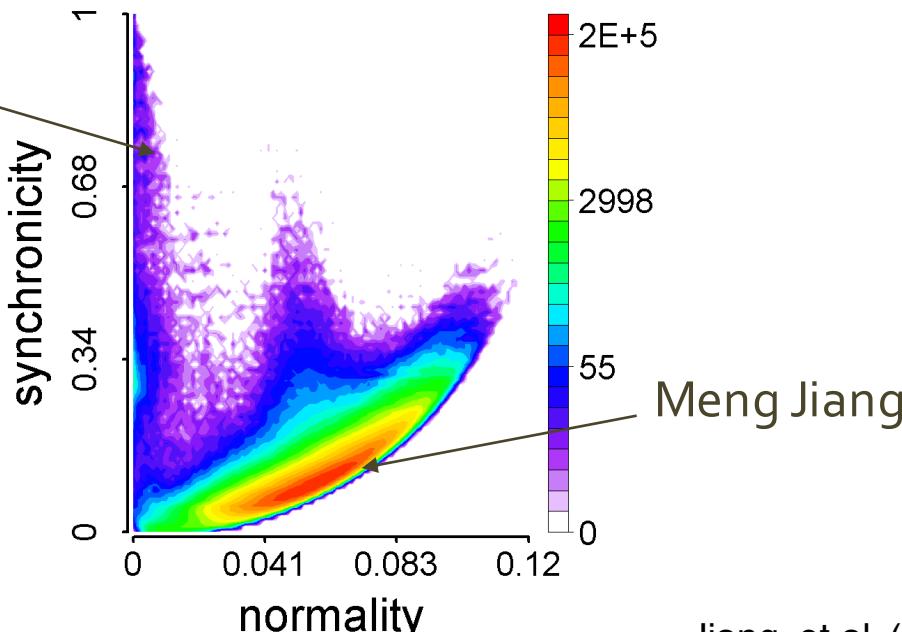
**Problem:** Given a normality value ( $n$ ) of a follower, find the minimal synchronicity value ( $s_{\min}$ ).

**Theorem:**

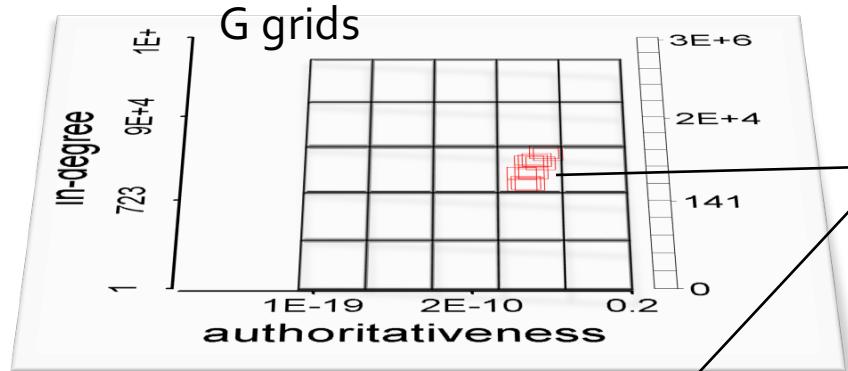
$$s_{\min} = \frac{-G n^2 + 2 n - s_b}{1 - G s_b} \quad (\textit{parabolic lower limit})$$

**Our CatchSync:**

Buy AB22 &  
Aisling Walsh



# Proof



$fp_g$ : #foreground points in grid g  
 $\sum fp_g = F = d(u)$  (#followees of u)



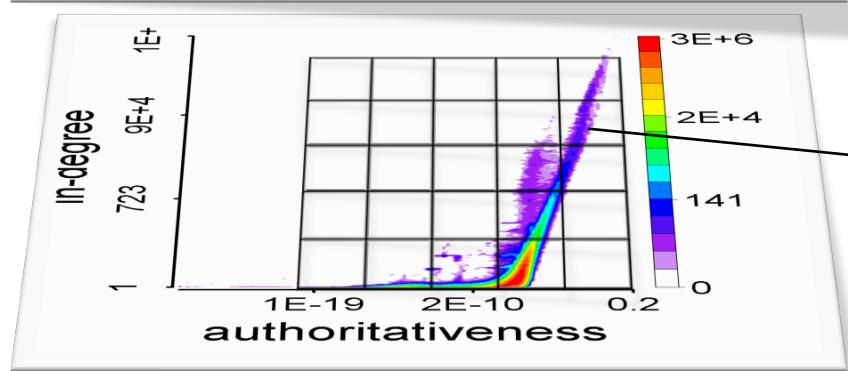
Given normality

$n = \sum (fp_g/F) (bp_g/B) = \sum f_g b_g$ ,  
 find minimal synchronicity

$$s_{\min} = \sum (fp_g/F) (fp_g/F) = \sum f_g^2$$

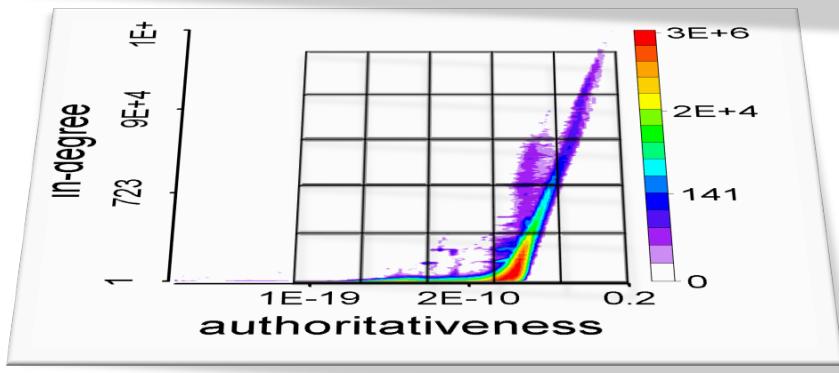
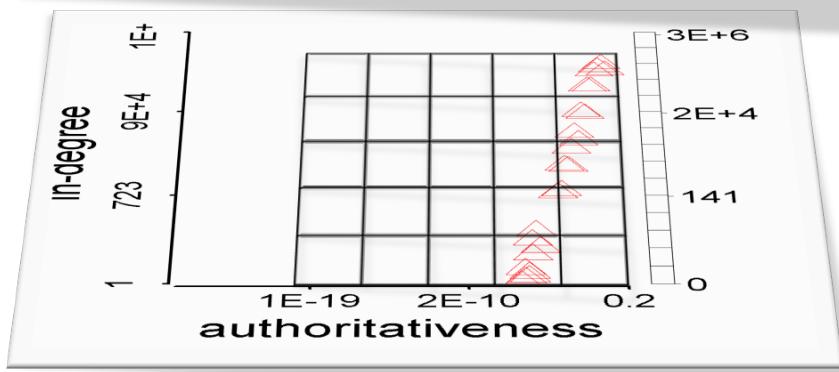
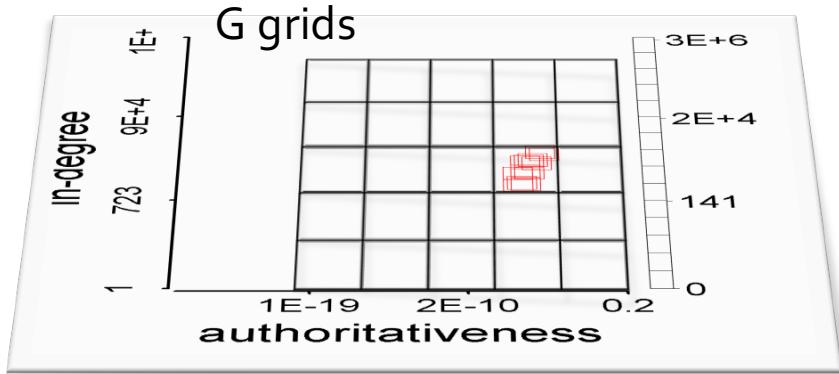
where

$$\sum f_g = 1, \sum b_g = 1$$



$bp_g$ : #background points in grid g  
 $\sum bp_g = B = N$  (#all users)

# Proof



**Lagrange multiplier:**

$$\text{minimize } s(f_g) = \sum f_g^2$$

$$\text{subject to } \sum f_g = 1, \sum f_g b_g = n$$

**Lagrange function:**

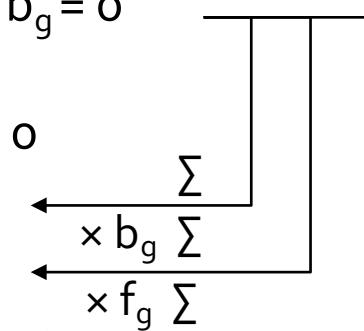
$$F(f_g, \lambda, \mu) = (\sum f_g^2) + \lambda (\sum f_g - 1) + \mu (\sum f_g b_g - n)$$

**Gradients:**

$$\left\{ \begin{array}{l} \nabla_{f_g} F = 2 f_g + \lambda + \mu b_g = 0 \\ \nabla_{\lambda} F = \sum f_g - 1 = 0 \\ \nabla_{\mu} F = \sum f_g b_g - n = 0 \end{array} \right.$$

$$\left\{ \begin{array}{l} 2 + \lambda G + \mu = 0 \\ 2 n + \lambda + \mu s_b = 0 \end{array} \right.$$

$$\left\{ \begin{array}{l} 2 s_{\min} + \lambda + \mu n = 0 \\ \sum b_g \sum \\ \times b_g \sum \\ \times f_g \sum \end{array} \right.$$

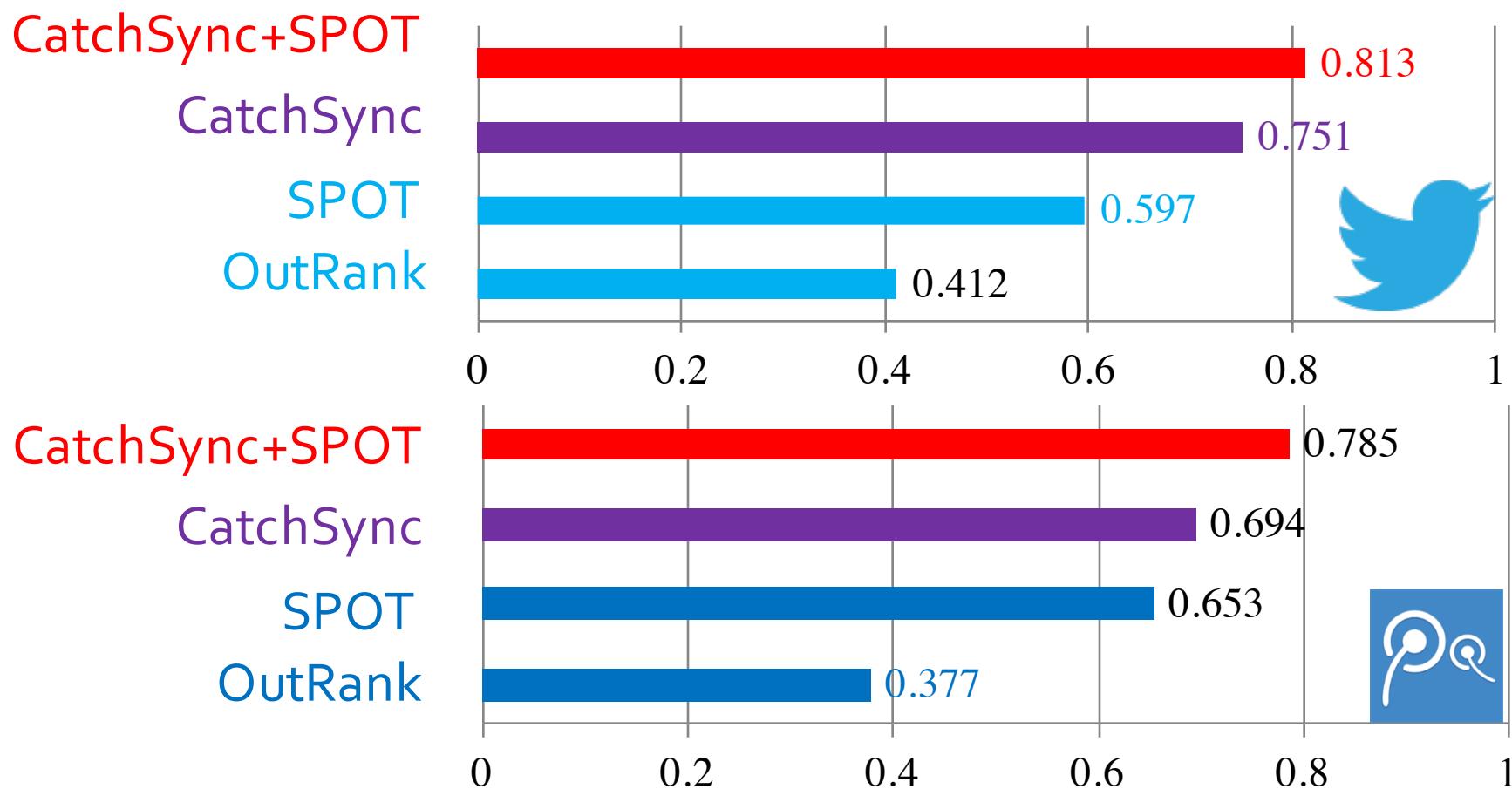


$$\text{where } s_b = \sum b_g^2.$$

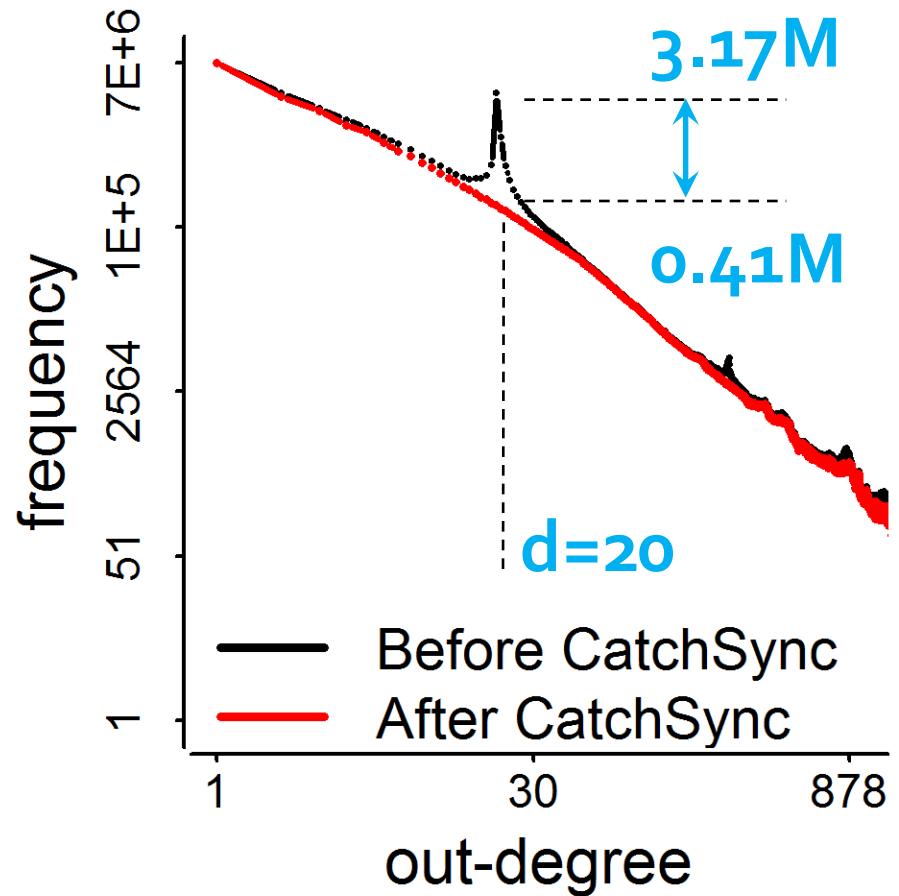
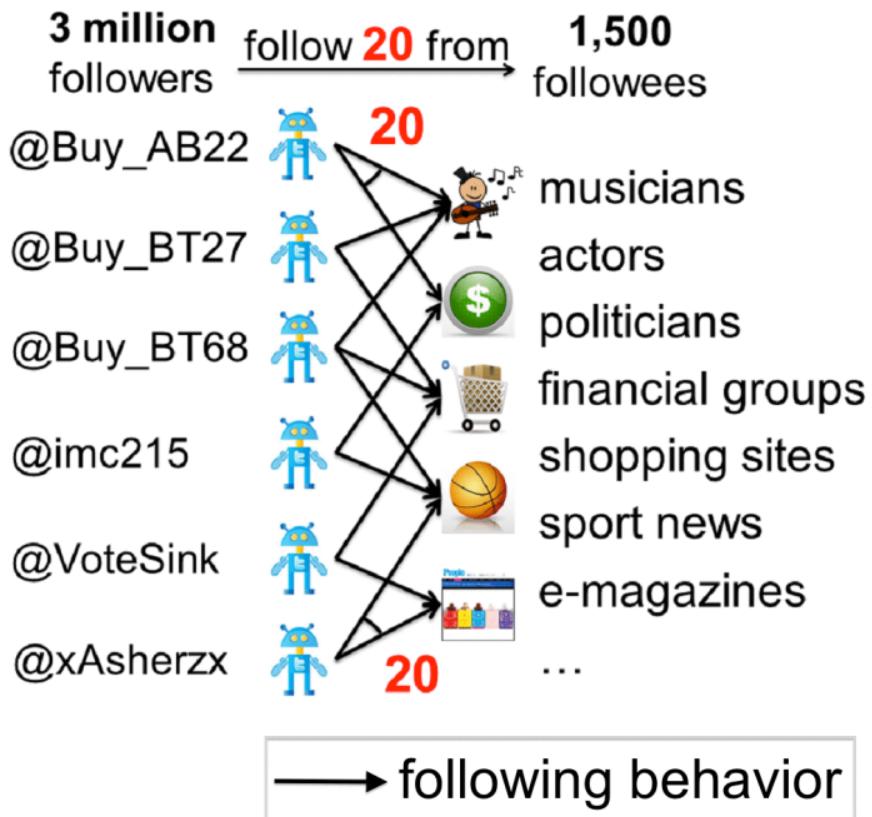
Therefore,

$$s_{\min} = \frac{-G n^2 + 2 n - s_b}{1 - G s_b}$$

# Accuracy: Complementary with Content-based Methods (SPOT)



# The Distribution was Recovered!

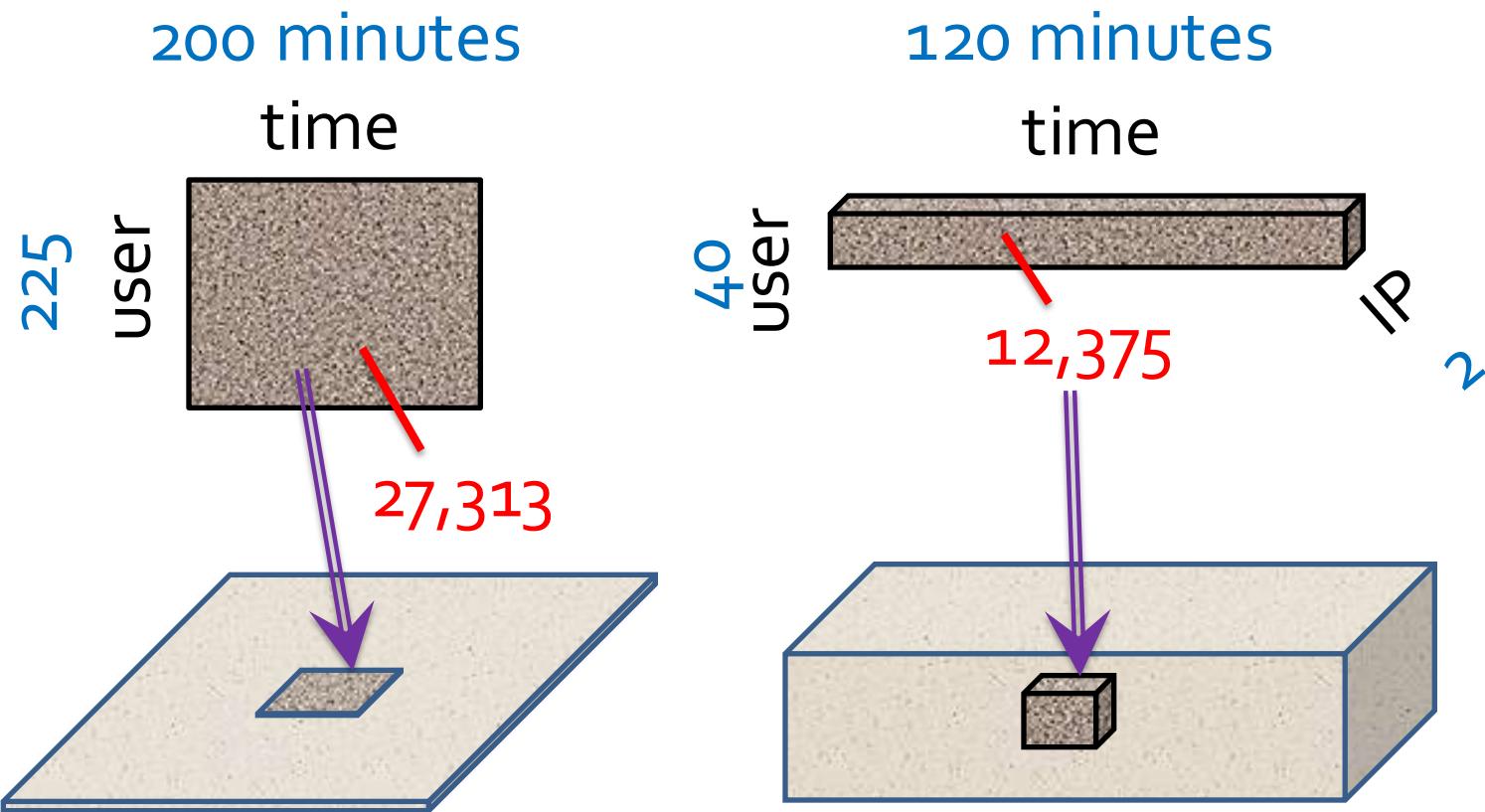


# 3. Social Spam

- **Problem definition:** Given **multidimensional behavioral data of spatiotemporal contexts**, find suspicious behaviors.

Dataset	Dimension/Mode				Mass
Weibo's Retweeting	User	Root ID	IP	Time (min)	#retweet
	29.5M	19.8M	27.8M	56.9K	211.7M
Weibo's Trending (Hashtag)	User	Hashtag	IP	Time (min)	#tweet
	81.2M	1.6M	47.7M	56.9K	276.9M
Network attacks (LBNL)	Src-IP	Dest-IP	Port	Time (sec)	#packet
	2,345	2,355	6,055	3,610	230,836

# Suspiciousness in Multi-dimensional Data

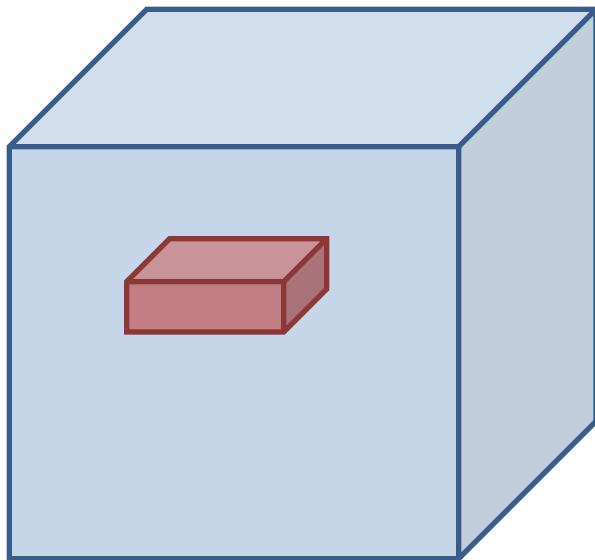


Q: Which is more suspicious?

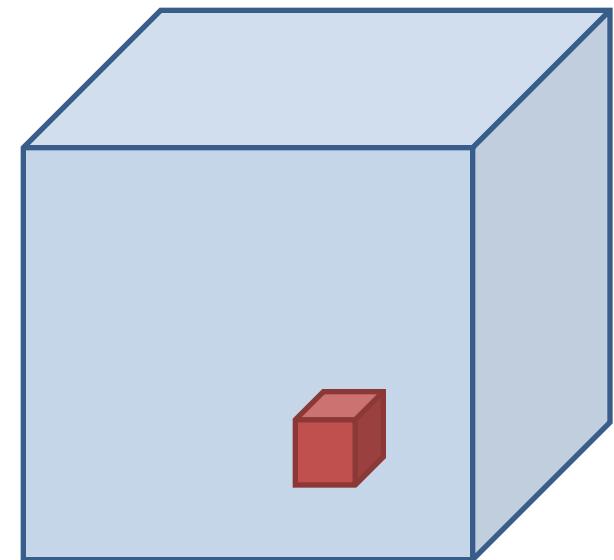
We need a metric to evaluate the suspiciousness.

# Criteria for Suspiciousness Metric

What properties are required of a good metric?



$N_1 \times N_2 \times N_3$   
Count data with  
total “mass”  $C$



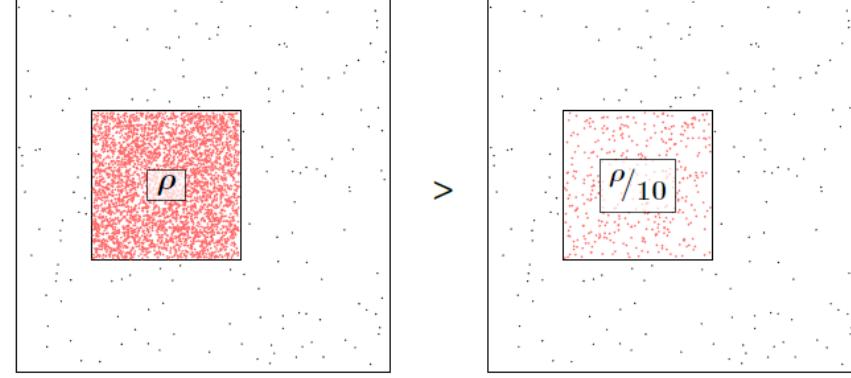
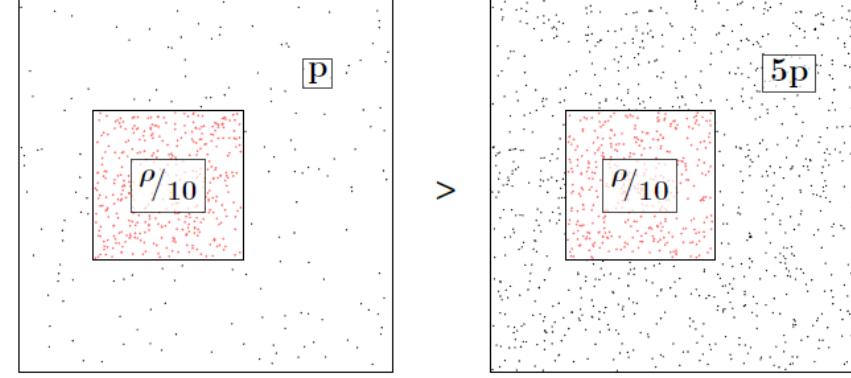
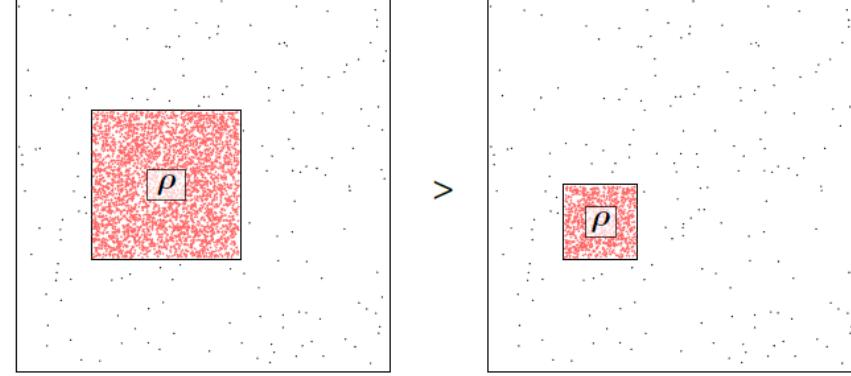
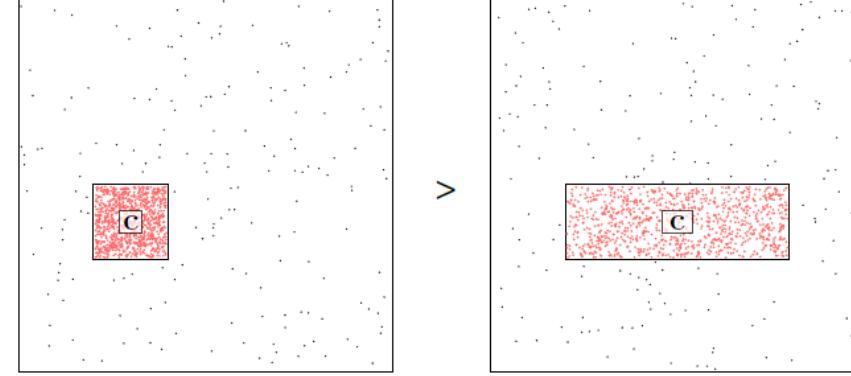
$$f( \begin{array}{c} n_1 \times n_2 \times n_3 \\ \text{mass } c \\ \text{density } \rho \end{array} )$$

vs

$$f( \begin{array}{c} n'_1 \times n'_2 \times n'_3 \\ \text{mass } c' \\ \text{density } \rho' \end{array} )$$

# Axioms: 1 to 4

$$c_1 > c_2 \iff f(\mathbf{n}, c_1, \mathbf{N}, C) > f(\mathbf{n}, c_2, \mathbf{N}, C)$$

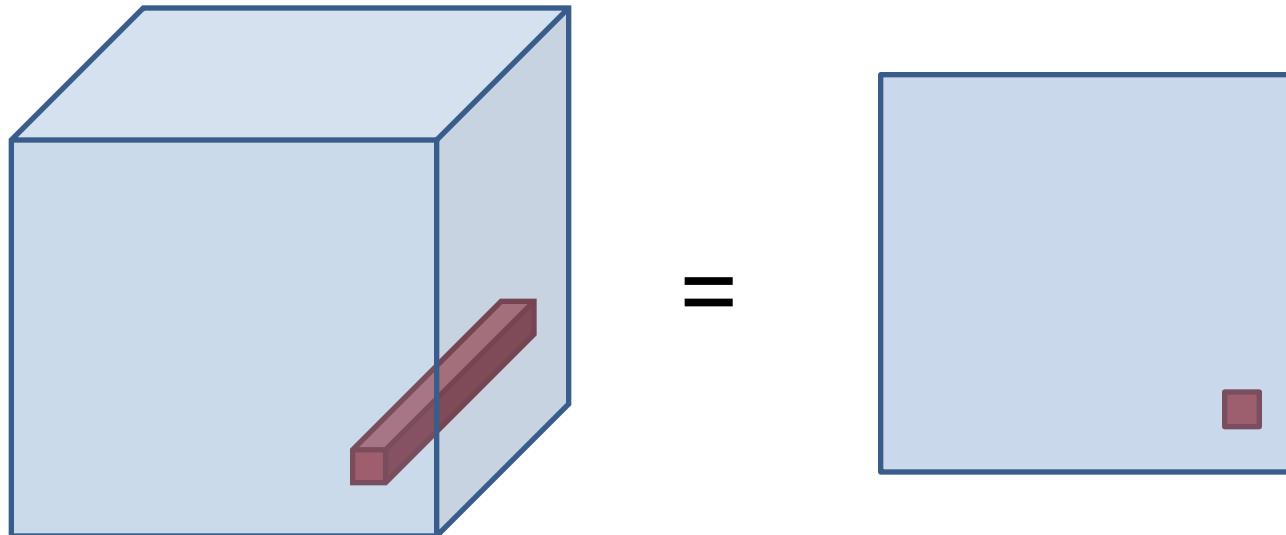
Density Axiom	Contrast Axiom
	
Size Axiom	Concentration Axiom
	

$$p_1 < p_2 \iff \hat{f}(\mathbf{n}, \rho, \mathbf{N}, p_1) > \hat{f}(\mathbf{n}, \rho, \mathbf{N}, p_2)$$

# Axiom 5: Cross Dimensions

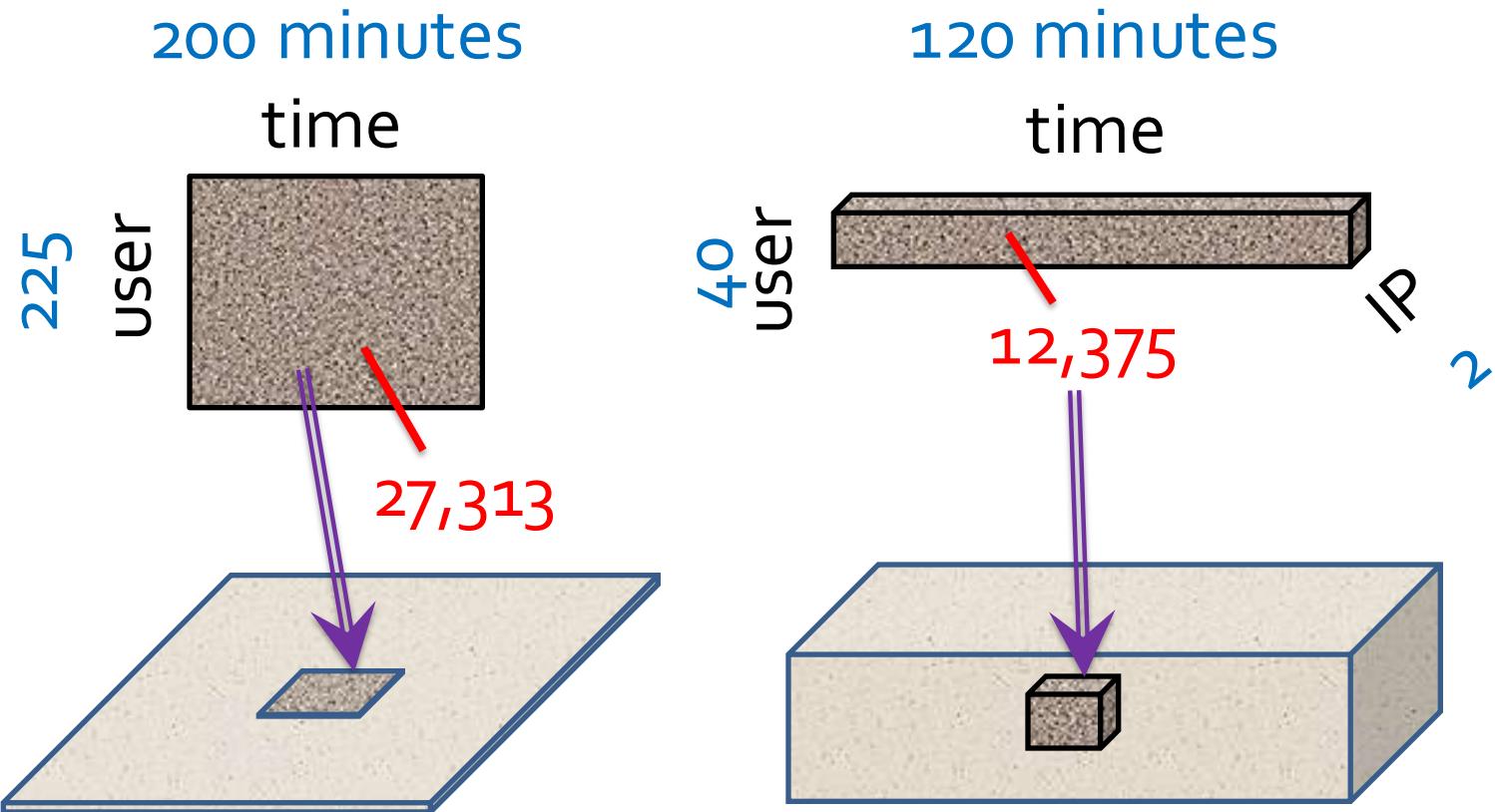
$$f_{K-1} \left( [n_k]_{k=1}^{K-1}, c, [N_k]_{k=1}^{K-1}, C \right) = f_K \left( ([n_k]_{k=1}^{K-1}, N_K), c, [N_k]_{k=1}^K, C \right)$$

Not including a mode is the same as including all values for that mode.



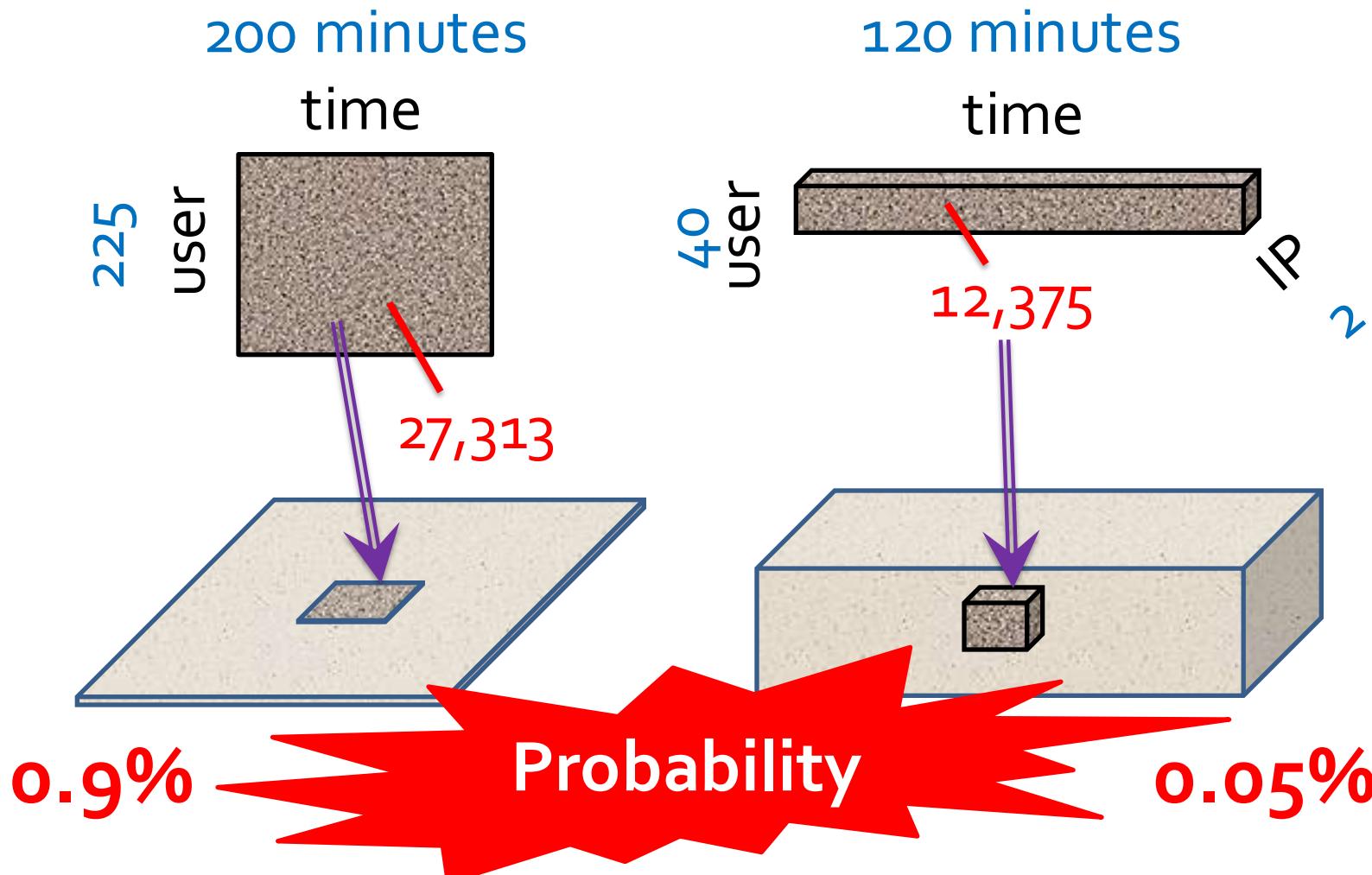
- ▶ New information (more modes) can only make our blocks more suspicious

# Scoring Suspiciousness



Q: Which is more suspicious?

# Scoring Suspiciousness



# A General Suspiciousness Metric

- Negative log likelihood of block's probability

$$f(n, c, N, C) = -\log [Pr(Y_n = c)]$$

**Lemma** Given an  $n_1 \times \cdots \times n_K$  block of mass  $c$  in  $N_1 \times \cdots \times N_K$  data of total mass  $C$ , the suspiciousness function is

$$f(\mathbf{n}, c, \mathbf{N}, C) = c \left( \log \frac{c}{C} - 1 \right) + C \prod_{i=1}^K \frac{n_i}{N_i} - c \sum_{i=1}^K \log \frac{n_i}{N_i}$$

Using  $\rho$  as the block's density and  $p$  is the data's density, we have the simpler formulation

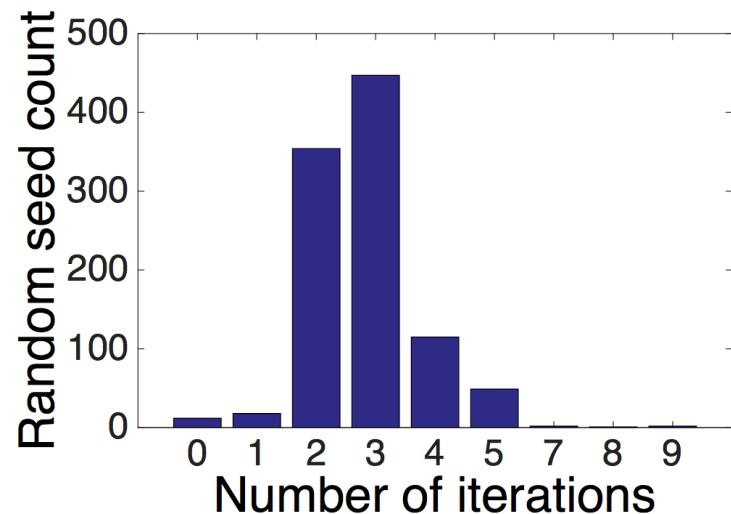
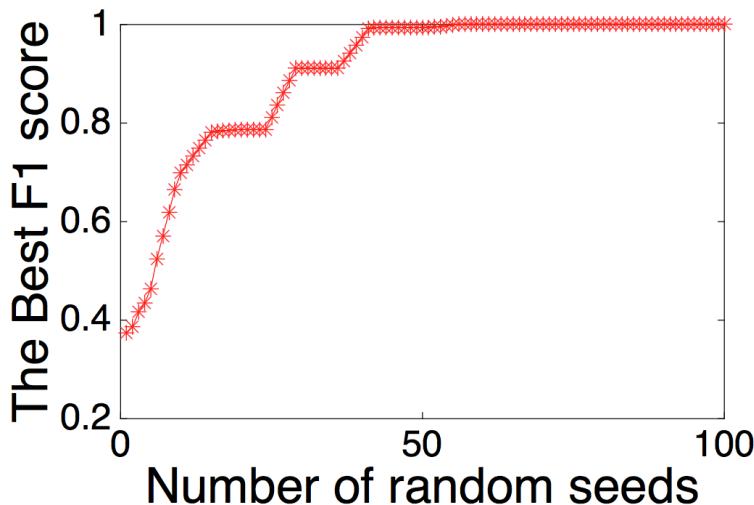
$$\hat{f}(\mathbf{n}, \rho, \mathbf{N}, p) = \left( \prod_{i=1}^K n_i \right) D_{KL}(\rho || p)$$

# Advantages

		Axioms				
		Density	Size	Concentration	Contrast	Multi-modal
Method		Scores				
		1	2	3	4	5
Metrics	<b>SUSPICIOUSNESS</b>	✓	✓	✓	✓	✓
	Mass	✓	✓	✗	✗	✗
	Density	✓	✓	✗	✓	✗
	Average Degree [9]	✓	✓	✗	✗	N/A
	Singular Value [10]	✓	✓	✓	✓	✗
Methods	<b>CROSSSPOT</b>	✓	✓	✓	✓	✓
	Subgraph [30, 10, 36]	✓	✓	✓	✓	N/A
	CopyCatch [6]	✓	✓	✓	✓	N/A
	EigenSpokes [31]	✗				
	TrustRank [14, 8]	✗				
	BP [28, 1]	✗				

# CrossSpot Algorithm

- Greedy algorithm by maximizing the metric
  - Start with seed blocks
  - Parameter-free: iteratively update the blocks
  - Scalable: parallelize to multiple machines



# Hijacked Hashtags

User × hashtag × IP × minute	Mass $c$	Suspiciousness
$582 \times 3 \times 294 \times \mathbf{56,940}$	5,941,821	111,799,948
$188 \times 1 \times 313 \times \mathbf{56,943}$	2,344,614	47,013,868
$75 \times 1 \times 2 \times 2,061$	689,179	19,378,403

User ID	Time	IP address (city, province)	Tweet text with hashtag
USER-D	11-18 12:12:51	IP-1 (Deyang, Shandong)	#Snow# the Samsung GALAXY SII QQ Service customized version...
USER-E	11-18 12:12:53	IP-1 (Deyang, Shandong)	#Snow# the Samsung GALAXY SII QQ Service customized version...
USER-F	11-18 12:12:54	IP-2 (Zaozhuang, Shandong)	#Snow# the Samsung GALAXY SII QQ Service customized version...
USER-E	11-18 12:17:55	IP-1 (Deyang, Shandong)	#Li Ning - a weapon with a hero# good support activities!
USER-F	11-18 12:17:56	IP-2 (Zaozhuang, Shandong)	#Li Ning - a weapon with a hero# good support activities!
USER-D	11-18 12:18:40	IP-1 (Deyang, Shandong)	#Toshiba Bright Daren# color personality test to find out your sense...
USER-E	11-18 17:00:31	IP-2 (Zaozhuang, Shandong)	#Snow# the Samsung GALAXY SII QQ Service customized version...
USER-D	11-18 17:00:49	IP-2 (Zaozhuang, Shandong)	#Toshiba Bright Daren# color personality test to find out your sense...
USER-F	11-18 17:00:56	IP-2 (Zaozhuang, Shandong)	#Li Ning - a weapon with a hero# good support activities!

# Network Attacks (LBNL)

	#	Src-IP $\times$ dst-IP $\times$ port $\times$ second	Mass $c$	Suspiciousness
CROSSSPOT	1	$411 \times 9 \times 6 \times \mathbf{3,610}$	47,449	552,465
	2	$533 \times 6 \times 1 \times \mathbf{3,610}$	30,476	400,391
	3	$5 \times 5 \times 2 \times \mathbf{3,610}$	18,881	317,529
	4	$11 \times 7 \times 7 \times \mathbf{3,610}$	20,382	295,869
HOSVD	1	$15 \times 1 \times 1 \times 1,336$	4,579	80,585
	2	$1 \times 2 \times 2 \times 1,035$	1,035	18,308
	3	$1 \times 1 \times 1 \times 1,825$	1,825	34,812
	4	$1 \times 13 \times 6 \times 181$	1,722	29,224

# 4. Events or Advertising Campaigns

- Density in **multi-contextual** behavioral data

20:03:09 @ebekahwsm  
this better be the best halftime show ever  
in the history of halftimes shows. ever.  
#SuperBowl

Contextual factors:

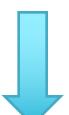
*One-guaranteed  
value*

*Empty (set  
of) value*

*Set value*

*Empty (set  
of) value*

*Dynamic*

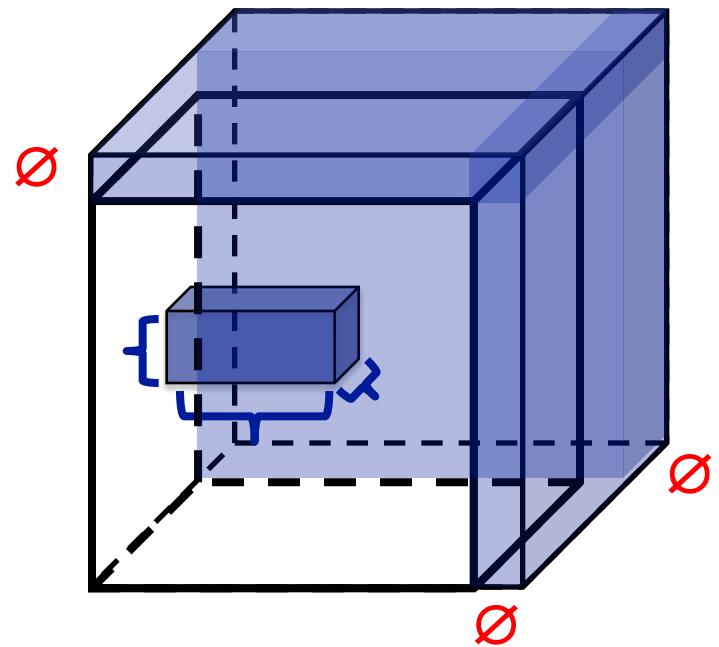
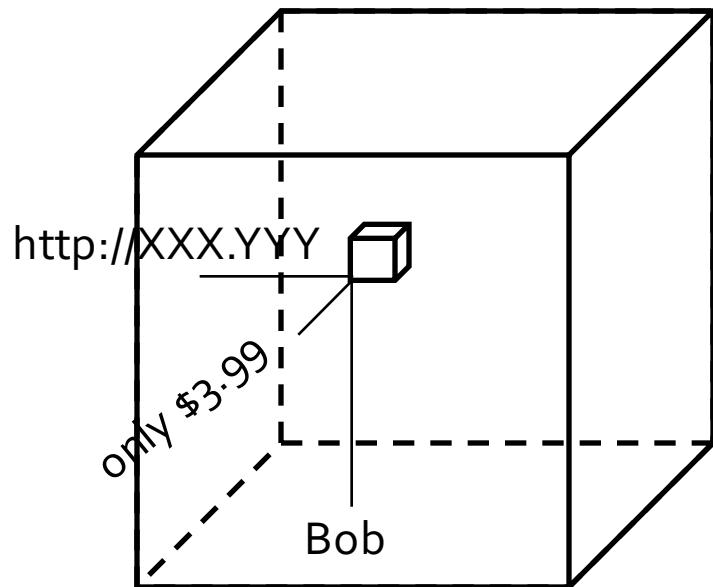


Time slice	User	Location	Phrase	Hashtag	URL
20:00-20:30	@ebekahwsm	∅	{best halftime show, in the history, halftimes shows}	{#SuperBowl}	∅

# Given tweets in Super Bowl 2013, can we find events (score prediction, halftime show, advertising campaigns, etc.)?

16:30	<p>16:30:31 <u>My prediction</u> Ravens 34 Niners 31          16:30:57 Ready for the big game :D, <u>my prediction</u> 24-20 SF #SuperBowl          16:31:14 <u>My prediction for superbowl..</u> 48.. Jets over Bears 17-13 Mark Sanchez MVP          16:32:24 I predict Baltimore Ravens will win 27 to 24 or 25 or 26. Basically it will be a close game.</p>	“my prediction”	user	phrase	hashtag	URL	3,397 tweets	Tartan #1: (1 dim) 16:30-17:30
17:00								
17:30	<p>17:30:51 RT @LMAOTWITPCIS: <u>Make Your Prediction</u>. Retweet For 49ers <a href="http://t.co/KKksEist">http://t.co/KKksEist</a>          17:31:01 RT @LMAOTWITPCIS: <u>Make Your Prediction</u>. Retweet For 49ers <a href="http://t.co/KKksEist">http://t.co/KKksEist</a>          17:31:16 RT @LMAOTWITPCIS: <u>Make Your Prediction</u>. Retweet For 49ers <a href="http://t.co/KKksEist">http://t.co/KKksEist</a>          17:31:19 RT @LMAOTWITPCIS: <u>Make Your Prediction</u>. Retweet For 49ers <a href="http://t.co/KKksEist">http://t.co/KKksEist</a></p>	“make your prediction”	user	phrase	RT @user	URL	196 tweets	Tartan #2: (3 dims) 17:00-18:00
18:00	<p>18:55:03 RT @49ers: Kaepernick is sacked on 3rd and goal. #49ers K David Akers makes 36-yard FG. Baltimore leads 7-3 with 3:58 left in 1st Qtr. #SB47          18:55:04 RT @49ers: Kaepernick is sacked on 3rd and goal. #49ers K David Akers makes 36-yard FG. Baltimore leads 7-3 with 3:58 left in 1st Qtr. #SB47          18:55:44 RT @Ravens: <u>David Akers</u> is good from 36 yards to make the score 7-3 Ravens.  <u>Nice job by the defense to tighten up in the red zone.</u></p>	“7-3”, “1st Qtr”	user	phrase	RT @user	URL	215 tweets	Tartan #3: (2 dims) 18:30-19:30
18:30			(213)	21	3	(0)		
19:00								
19:30	<p>20:20:01 RT @ExtraGrumpyCat: No Superbowl halftime show will ever surpass this. <a href="http://t.co/0VSy7Cv6">http://t.co/0VSy7Cv6</a>          20:20:02 RT @WolfpackAlan: No Superbowl halftime show will ever surpass this. <a href="http://t.co/6Bll0PXs">http://t.co/6Bll0PXs</a>          20:20:04 RT @ExtraGrumpyCat: No Superbowl halftime show will ever surpass this. <a href="http://t.co/0VSy7Cv6">http://t.co/0VSy7Cv6</a>          20:20:05 RT @WolfpackAlan: No Superbowl halftime show will ever surpass this. <a href="http://t.co/6Bll0PXs">http://t.co/6Bll0PXs</a></p>	halftime show”	user	phrase	RT @user	URL	617 tweets	Tartan #4: (3 dims) 20:00-21:00
20:00			(617)	11	4	4		
20:30	<p>20:20:47 (Manhattan, NY)...and <u>every one of those girls</u> took #ballet #Beyonce #superbowl          20:22:01 (New York, NY) I have <u>the biggest lady boner</u> for Beyonce #BeyonceBowl #DestinyBowl #DestinysChild #SuperBowl          20:24:32 (Manhattan, NY) No one can ever <u>top that performance</u> by Beyonce. EVER. #Beyonce #superbowl #halftimeshow</p>	“beyonce”, #beyonce, #superbowl, #DestinysChild	location	phrase	hashtag	URL	166 tweets	Tartan #5: (3 dims) 20:00-21:00
21:00								
21:30	<p>21:44:42 Ahora si pff #49ers 23-28 #Ravens          21:44:44 Baltimore #Ravens 28-23 San Francisco #49ers          21:44:50 FG Akers #49ers 23-28 #Ravens 3Q 3:10 #SuperBowlXLVII #SuperBowl #NFL</p>	“28-23”, #49ers, #Ravens	user	phrase	hashtag	URL	653 tweets	Tartan #6: (2 dims) 21:00-22:00
22:00			(650)	69	11	(0)		
	<p>22:42:27 Congratulations Ravens!!!!          22:42:43 Congratulations Ray Lewis and the Ravens.          22:42:43 Game over! <u>Ravens</u> won ray got his retirement ring now all y'all boys and girls go to sleep !          22:42:52 @LetThatBoyTweet: Game over. <u>Ravens</u> win the Super Bowl.”</p>	“congratulations”, “game over”	user	phrase	hashtag	URL	1,950 tweets	Tartan #7: (1 dim) 22:00-23:30
			(1942)	248	(0)	(0)		

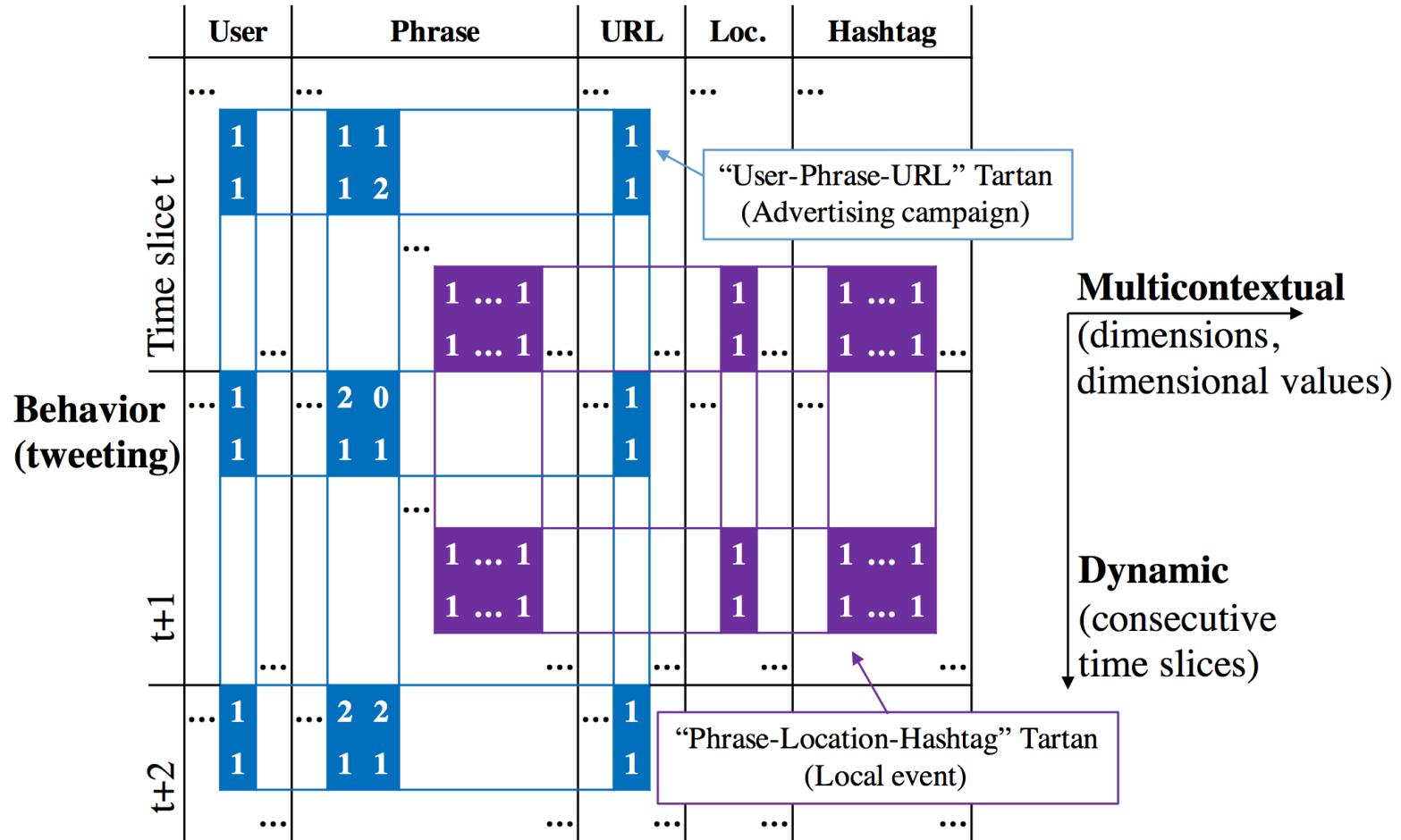
# Tensor Fails: Representation



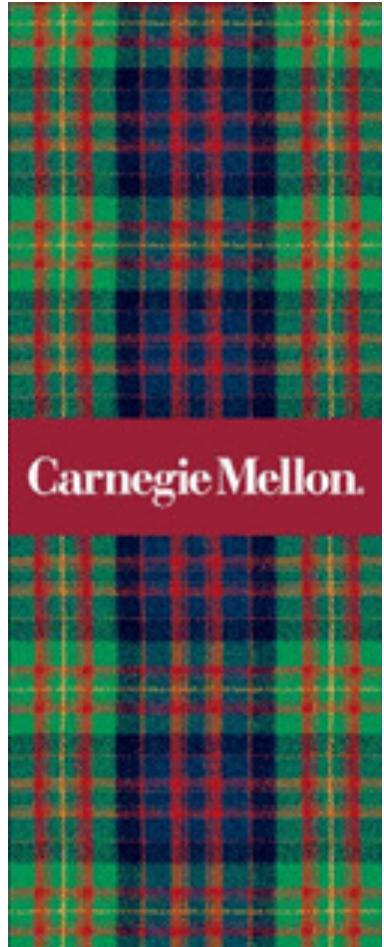
# “Two-Level Matrix” and Tartans

Behavior representation

Behavior summaries



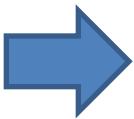
# What is Tartan?



**GO TARTANS!**



Visited CMU in 2012-13



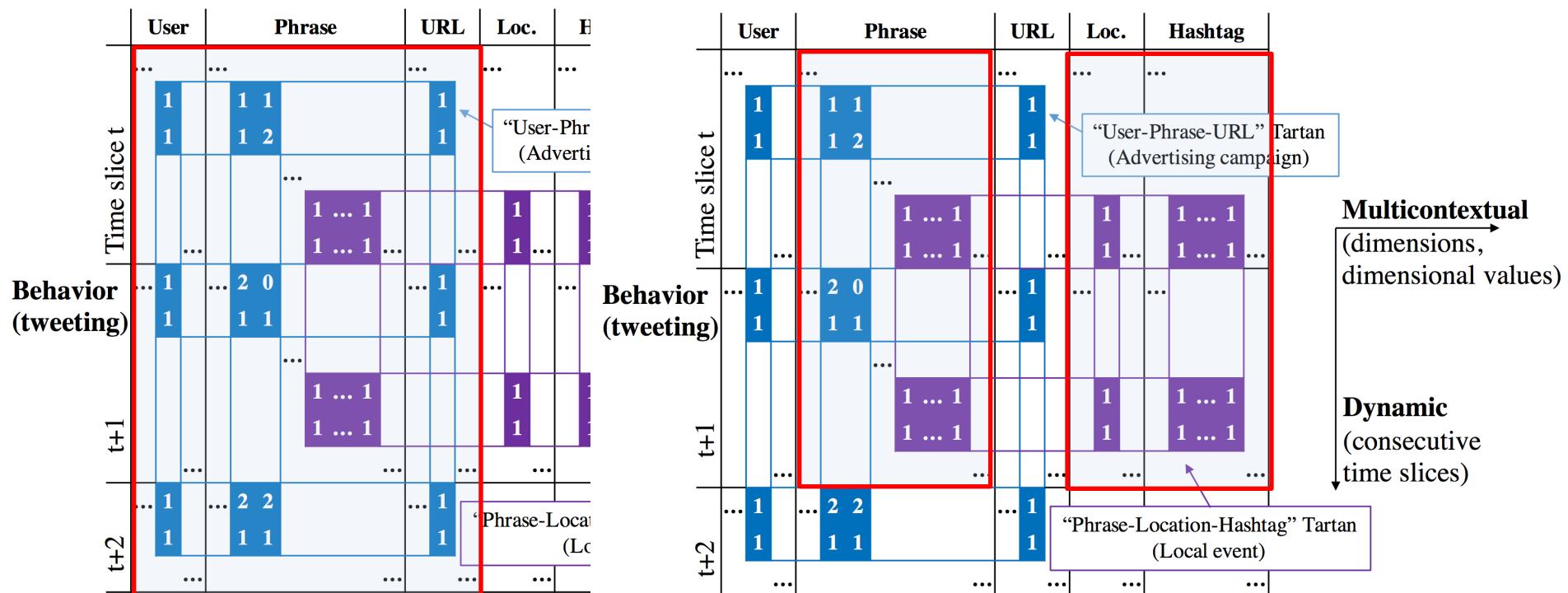
Watched lots of  
Tartans' games...



# Objective Function to Maximize (Minimum Description Length principle)

$$f(\mathcal{A}, \mathcal{X}) = L(\mathcal{X}^{\mathcal{A}}) - L(\mathcal{A}) - L(\mathcal{X}^{\mathcal{A}} \setminus \mathcal{A}).$$

Tartan Data First-level matrix Individual entries



$$\mathcal{X}^{\mathcal{A}} = \{\mathcal{X}_d^{(t)}(b, i) | d \in \mathcal{D}, t \in \mathcal{T}, i \in \{1, \dots, N_d\}, b \in \{1, \dots, E^{(t)}\}\}.$$

# Objective Function to Maximize (cont.)

$$f(\mathcal{A}, \mathcal{X}) = L(\mathcal{X}^{\mathcal{A}}) - L(\mathcal{A}) - L(\mathcal{X}^{\mathcal{A}} \setminus \mathcal{A}).$$

$$V = (\sum_{d \in \mathcal{D}} N_d) (\sum_{t \in \mathcal{T}} E^{(t)}).$$

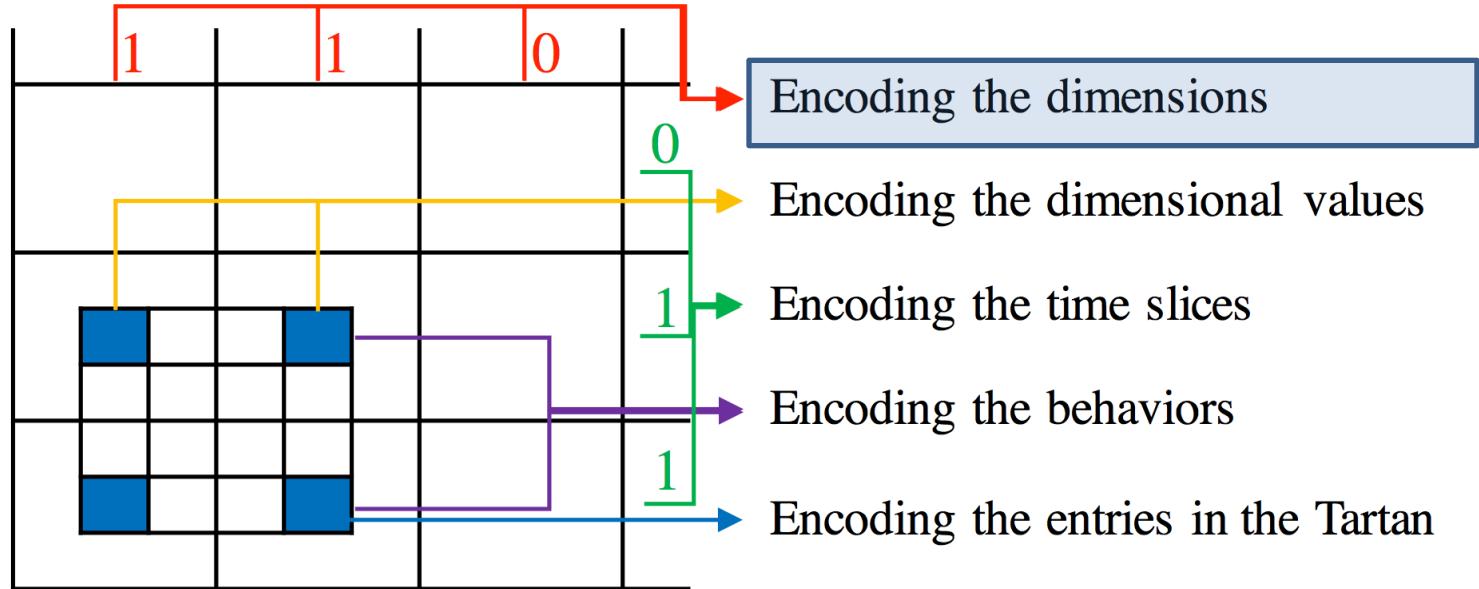
$$C = \sum_{d \in \mathcal{D}, t \in \mathcal{T}} \sum_{b \in \{1, \dots, E^{(t)}\}, i \in \{1, \dots, N_d\}} \mathcal{X}_d^{(t)}(b, i).$$

$$\begin{aligned} L(\mathcal{X}^{\mathcal{A}}) &= g(V + C, C) + L_{\mathcal{D}}(\mathcal{A}) + L_{\mathcal{T}}(\mathcal{A}) \\ &\quad + \sum_{d \in \mathcal{D}} \log^* N_d + \sum_{t \in \mathcal{T}} \log^* E^{(t)}. \end{aligned}$$

$$L(\mathcal{A}) = L_{\mathcal{D}}(\mathcal{A}) + L_{\mathcal{V}}(\mathcal{A}) + L_{\mathcal{T}}(\mathcal{A}) + L_{\mathcal{B}}(\mathcal{A}) + L_{\mathcal{A}}(\mathcal{A}).$$

$$L(\mathcal{X}^{\mathcal{A}} \setminus \mathcal{A}) = g(V + C - v - c, C - c);$$

# Encoding the Tartan: Dimensions

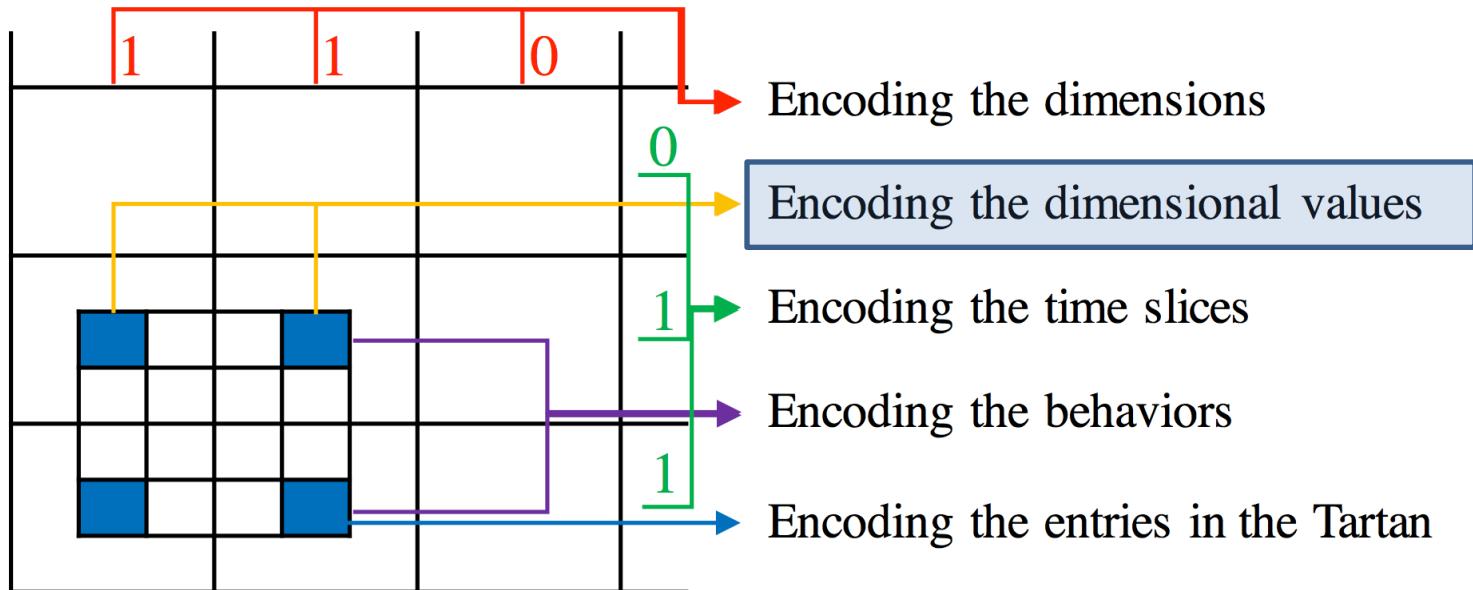


$$H_{\mathcal{D}}(X) = - \sum_{x \in \{0,1\}} P(X = x) \log P(X = x)$$

$$= - \left( \frac{D^{\mathcal{A}}}{D} \log \frac{D^{\mathcal{A}}}{D} + \frac{D - D^{\mathcal{A}}}{D} \log \frac{D - D^{\mathcal{A}}}{D} \right).$$

$$\begin{aligned} L_{\mathcal{D}}(\mathcal{A}) &= \log^* D + \log^* D^{\mathcal{A}} + D \cdot H_{\mathcal{D}}(X) \\ &= \log^* D + \log^* D^{\mathcal{A}} + g(D, D^{\mathcal{A}}), \end{aligned}$$

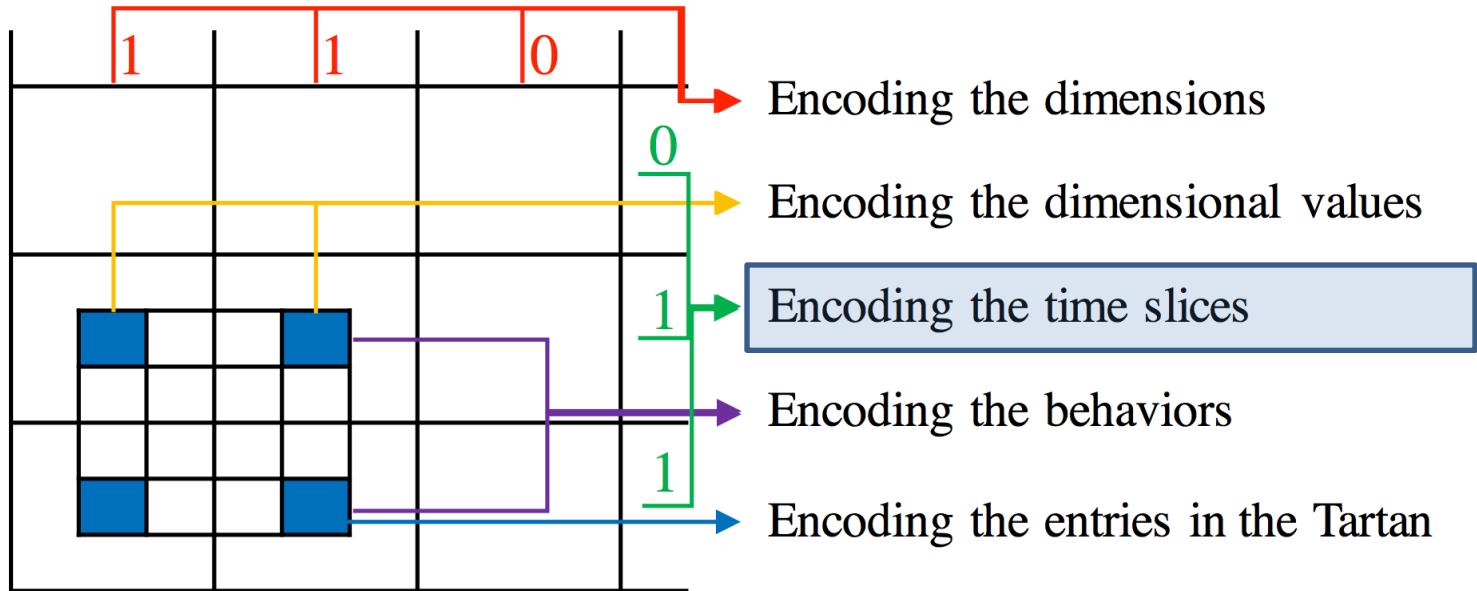
# Encoding the Tartan: Dimensional Values



$$H_{\mathcal{V}_d}(X) = - \left( \frac{n_d}{N_d} \log \frac{n_d}{N_d} + \frac{N_d - n_d}{N_d} \log \frac{N_d - n_d}{N_d} \right).$$

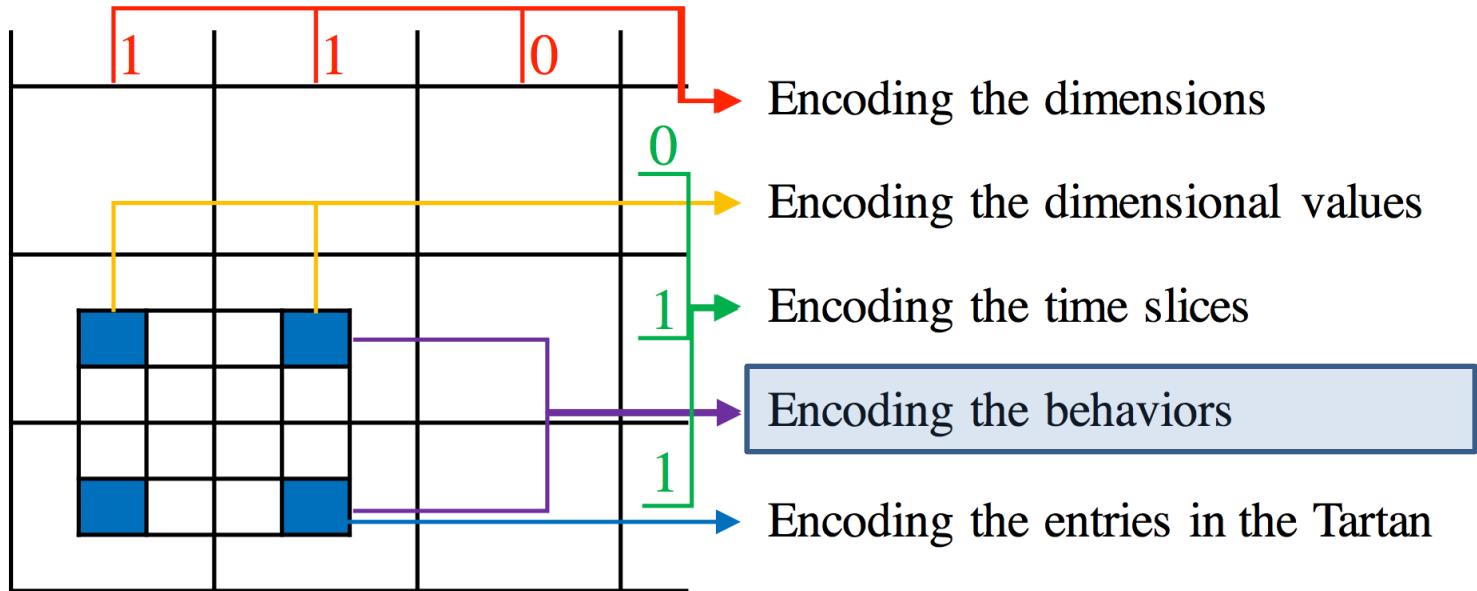
$$L_{\mathcal{V}}(\mathcal{A}) = \sum_{d \in \mathcal{D}} \left( \log^* N_d + \log^* n_d + g(N_d, n_d) \right).$$

# Encoding the Tartan: Time Slices



$$L_{\mathcal{T}}(\mathcal{A}) = \log^* T + \log^* T^{\mathcal{A}} + \log^* t_{start}$$

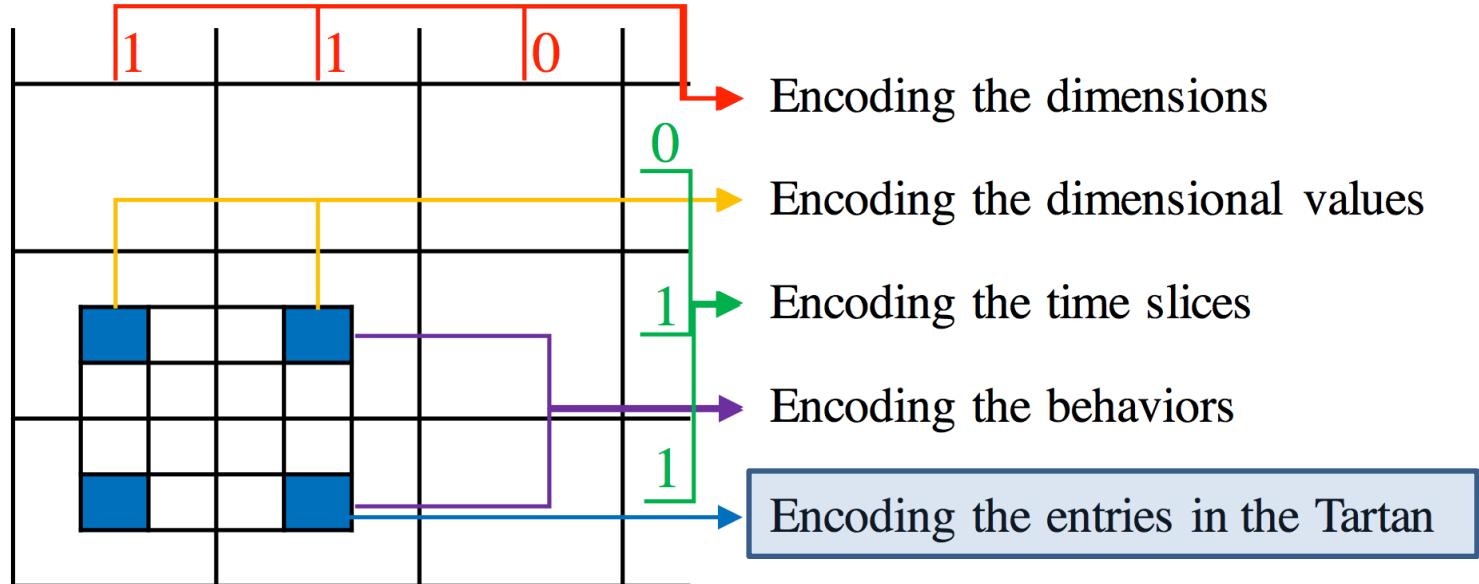
# Encoding the Tartan: Behaviors



$$H_{\mathcal{B}^{(t)}}(X) = - \left( \frac{e^{(t)}}{E^{(t)}} \log \frac{e^{(t)}}{E^{(t)}} + \frac{E^{(t)} - e^{(t)}}{E^{(t)}} \log \frac{E^{(t)} - e^{(t)}}{E^{(t)}} \right).$$

$$L_{\mathcal{B}}(\mathcal{A}) = \sum_{t \in \mathcal{T}} \left( \log^* E^{(t)} + \log^* e^{(t)} + g(E^{(t)}, e^{(t)}) \right).$$

# Encoding the Tartan: Entries



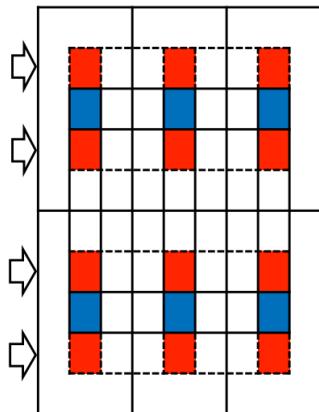
$$v = (\sum_{d \in \mathcal{D}} n_d) (\sum_{t \in \mathcal{T}} e^{(t)}).$$

$$c = \sum_{d \in \mathcal{D}, t \in \mathcal{T}} \sum_{b \in \mathcal{B}^{(t)}, i \in \mathcal{V}_d} \mathcal{X}_d^{(t)}(b, i).$$

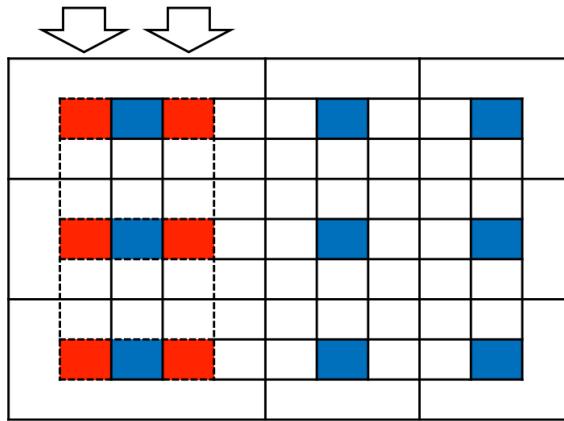
$$H_{\mathcal{A}}(X) = -\left(\frac{c}{v+c} \log \frac{c}{v+c} + \frac{v}{v+c} \log \frac{v}{v+c}\right).$$

$$L_{\mathcal{A}}(\mathcal{A}) = (v + c) H_{\mathcal{A}}(X) = g(v + c, c).$$

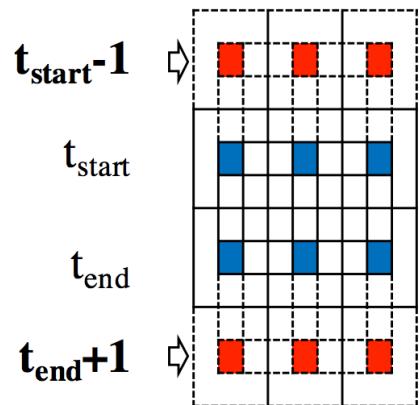
# Greedy Search for the Local Optimum



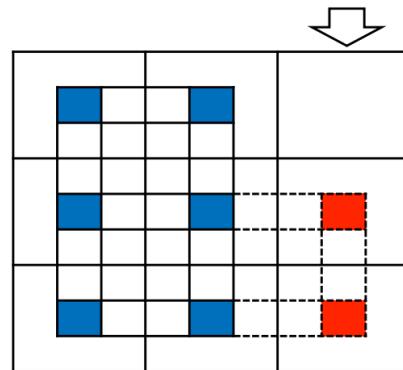
(a) Update the set of behaviors.



(b) Update the set of values.



(c) Update the consecutive time slices.



(d) Update the set of dimensions.

**Time complexity:**

$$\mathcal{O}(\sum_d N_d \log N_d + \sum_t E^{(t)} \log E^{(t)})$$

# Experiments: Events in Tweets

16:30	16:30:31 My prediction Ravens 34 Niners 31 16:30:57 Ready for the big game :D, my prediction 24-20 SF #SuperBowl	“my prediction”	user (3,325)	phrase <b>226</b>	hashtag (0)	URL (0)		<b>3,397</b>	tweets	Tartan #1: (1 dim) 16:30-17:30	
17:00	16:31:14 My prediction for superbowl.. 48.. Jets over Bears 17-13 Mark Sanchez MVP 16:32:24 I predict Baltimore Ravens will win 27 to 24 or 25 or 26. Basically it will be a close game.										
17:30	17:30:51 RT @LMAOTWITPICS: Make Your Prediction. Retweet For 49ers <a href="http://t.co/KKksEist">http://t.co/KKksEist</a> 17:31:01 RT @LMAOTWITPICS: Make Your Prediction. Retweet For 49ers <a href="http://t.co/KKksEist">http://t.co/KKksEist</a> 17:31:16 RT @LMAOTWITPICS: Make Your Prediction. Retweet For 49ers <a href="http://t.co/KKksEist">http://t.co/KKksEist</a> 17:31:19 RT @LMAOTWITPICS: Make Your Prediction. Retweet For 49ers <a href="http://t.co/KKksEist">http://t.co/KKksEist</a>	“make your prediction”	user (196)	phrase <b>RT @user</b> <b>4</b>	RT <b>@user</b> <b>1</b>	URL <b>1</b>		<b>196</b>	tweets	Tartan #2: (3 dims) 17:00-18:00	
18:00	18:55:03 RT @49ers: Kaepernick is sacked on 3rd and goal. #49ers K David Akers makes 36-yard FG. Baltimore leads 7-3 with 3:58 left in 1st Qtr. #SB47 18:55:04 RT @49ers: Kaepernick is sacked on 3rd and goal. #49ers K David Akers makes 36-yard FG. Baltimore leads 7-3 with 3:58 left in 1st Qtr. #SB47 18:55:44 RT @Ravens: David Akers is good from 36 yards to make the score 7-3 Ravens. Nice job by the defense to tighten up in the red zone.	“7-3”, “1 <sup>st</sup> Qtr”	user (213)	phrase <b>RT @user</b> <b>21</b>	RT <b>@user</b> <b>3</b>	URL (0)		<b>215</b>	tweets	Tartan #3: (2 dims) 18:30-19:30	
18:30											
19:00	20:20:01 RT @ExtraGrumpyCat: No Superbowl halftime show will ever surpass this. <a href="http://t.co/0VSy7Cv6">http://t.co/0VSy7Cv6</a> 20:20:02 RT @WolfpackAlan: No Superbowl halftime show will ever surpass this. <a href="http://t.co/6Bll0PXs">http://t.co/6Bll0PXs</a> 20:20:04 RT @ExtraGrumpyCat: No Superbowl halftime show will ever surpass this. <a href="http://t.co/0VSy7Cv6">http://t.co/0VSy7Cv6</a> 20:20:05 RT @WolfpackAlan: No Superbowl halftime show will ever surpass this. <a href="http://t.co/6Bll0PXs">http://t.co/6Bll0PXs</a>	halftime show”	user (617)	phrase <b>RT @user</b> <b>11</b>	RT <b>@user</b> <b>4</b>	URL <b>4</b>		<b>617</b>	tweets	Tartan #4: (3 dims) 20:00-21:00	
19:30											
20:00	20:20:47 (Manhattan, NY)...and every one of those girls took #ballet #Beyonce #superbowl 20:22:01 (New York, NY) I have the biggest lady boner for Beyonce #BeyonceBowl #DestinyBowl #DestinysChild #SuperBowl	“beyonce”, #beyonce,	location <b>2</b>	phrase <b>55</b>	hashtag <b>17</b>	URL (0)		<b>166</b>	tweets	Tartan #5: (3 dims) 20:00-21:00	
20:30	20:24:32 (Manhattan, NY) No one can ever top that performance by Beyonce. EVER. #Beyonce #superbowl #halftimeshow	#superbowl, #DestinysChild									
21:00	21:44:42 Ahora si pff #49ers 23-28 #Ravens 21:44:44 Baltimore #Ravens 28-23 San Francisco #49ers 21:44:50 FG Akers #49ers 23-28 #Ravens 3Q 3:10 #SuperBowlXLVII #SuperBowl #NFL	“28-23”, #49ers, #Ravens	user (650)	phrase <b>hashtag</b> <b>69</b>	URL <b>11</b>	URL (0)		<b>653</b>	tweets	Tartan #6: (2 dims) 21:00-22:00	
21:30											
22:00	22:42:27 Congratulations Ravens!!!! 22:42:43 Congratulations Ray Lewis and the Ravens. 22:42:43 Game over! Ravens won ray got his retirement ring now all y'all boys and girls go to sleep ! 22:42:52 “@LetThatBoyTweet: Game over. Ravens win the Super Bowl.”	“congratulations”, “game over”	user (1942)	phrase <b>hashtag</b> <b>248</b>	URL (0)	URL (0)		<b>1,950</b>	tweets	Tartan #7: (1 dim) 22:00-23:30	

# Experiments: Research Trends in DBLP Data

1997  
2000  
2003  
2006  
2009  
2012

Author	Venue	Keyword	Cited	#Paper
<b>76</b> Cheng-xiang Zhai Hui Fang S. Kambhampati	<b>7</b> SIGIR VLDB TKDE	<b>7</b> “information retrieval” “data integration” “text classification”	<b>68</b> p56743 <sup>1</sup> p62995 p76869	<b>32</b> 2003- 2007

<sup>1</sup> “A language modeling approach to information retrieval”

Venue	Keyword	#Paper
<b>5</b> ICML NIPS ...	<b>6</b> “reinforcement learning” “machine learning”	<b>40</b> 1997- 2002

Author	Venue	Cited	#Paper
<b>6</b> Jiawei Han Xifeng Yan	<b>1</b> SIG- MOD	<b>1</b> p76095 <sup>2</sup>	<b>22</b> 2004- 2010

<sup>2</sup> “Frequent subgraph discovery”

Venue	Keyword	#Paper
<b>3</b> ICDM AAAI TKDE	<b>1</b> “anomaly detection”	<b>25</b> 2005- 2013

Author	Venue	Keyword	#Paper
<b>27</b> C. Faloutsos J. Pei P. S. Yu X. Lin C. Aggarwal...	<b>6</b> KDD ICDM ICDE TKDE ...	<b>12</b> “large graphs” “data streams” “evolving data” “evolving graphs” ...	<b>70</b> 2006- 2013

Author	Venue	Keyword	Cited	#Paper
<b>12</b> Ryen White Hang Li Tie-Yan Liu Zhaohui Zheng...	<b>5</b> SIGIR WWW WSDM CIKM...	<b>3</b> “web search” “click-through data” “sponsored search”	<b>12</b> p82630 <sup>3</sup> p116290 p103899 p106191...	<b>32</b> 2006- 2013

<sup>3</sup> “Optimizing search engines using clickthrough data”

Author	Venue	Keyword	#Paper
<b>8</b> Qiang Yang Dou Shen Sinno Pan...	<b>3</b> KDD PAKDD AAAI	<b>6</b> “transfer learning” “data mining” “localization models”	<b>17</b> 2007- 2010

# Summary

- Density-based methods and applications
  - Density in temporal bipartite graphs
    - CopyCatch (WWW'13): Facebook, ill-gotten likes
  - Density (synchronicity) in large directed graphs
    - CatchSync (KDD'14): Twitter/Weibo, zombie followers
  - Density (suspiciousness) in multidimensional data
    - CrossSpot (TKDE'16): Twitter/Weibo, social spam
  - Density (MDL principle) in multicontextual data
    - CatchTartan (KDD'16): Twitter, events/campaigns

# References

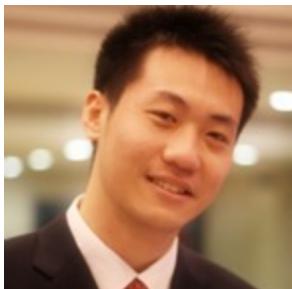
- Meng Jiang, Peng Cui, and Christos Faloutsos. “Suspicious behavior detection: current trends and future directions.” IEEE Intelligent Systems, 2016. (Survey paper)
- Alex Beutel, Wanhong Xu, Venkatesan Guruswami, Christopher Palow, Christos Faloutsos. “Copycatch: stopping group attacks by spotting lockstep behavior in social networks”, WWW 2013.
- Meng Jiang, Peng Cui, Alex Beutel, Christos Faloutsos, Shiqiang Yang. “CatchSync: Catching Synchronized Behavior in Large Directed Graphs”, KDD 2014 Best Paper Finalist.
- Meng Jiang, Alex Beutel, Peng Cui, Bryan Hooi, Shiqiang Yang, Christos Faloutsos. “Spotting Suspicious Behaviors in Multimodal Data: A General Metric and Algorithms”, TKDE 2016.
- Meng Jiang, Christos Faloutsos, Jiawei Han. “CATCHTARTAN: Representing and Summarizing Dynamic Multicontextual Behaviors”, KDD 2016.
- M. Faloutsos, P. Faloutsos, and C. Faloutsos. “On power-law relationships of the internet topology.” SIGCOMM, 1999.
- A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Weiner. “Graph structure in the web.” Computer Networks, 2000.
- F. Chung and L. Lu. “The average distances in random graphs with given expected degrees.” PNAS, 2002.
- J. Kleinberg. “Authoritative sources in a hyperlinked environment.” JACM, 1999.
- H. Kwak, C. Lee, H. Park, and S. Moon. “What is Twitter, a social network or a news media?” WWW, 2010.



KDD 2017

Halifax, Nova Scotia - Canada  
August 13 - 17, 2017

# Tutorial: Data-Driven Approaches towards Malicious Behavior Modeling



Meng Jiang  
University of Notre Dame



Srijan Kumar  
Stanford University



Christos Faloutsos  
Carnegie Mellon University



V.S. Subrahmanian  
University of Maryland, College Park

Tutorial link: <http://bit.ly/kdd2017>