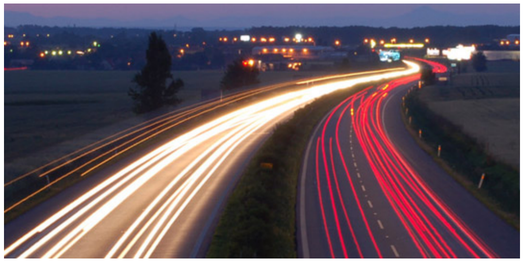
Carnegie Mellon

Tensor Analysis – Applications and Algorithms

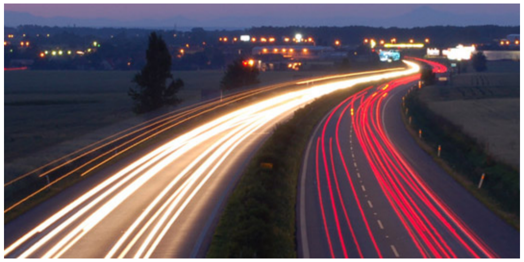
Christos Faloutsos

CMU



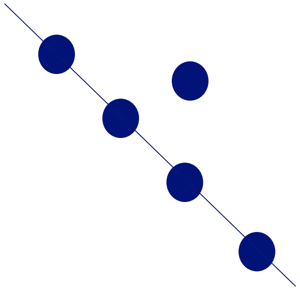
Roadmap

- Applications – pattern discovery
  - Brain scans – coupled matrix-tensor factorization
  - Power grid
- Applications – anomaly detection
- Algorithms
- Conclusions



Roadmap

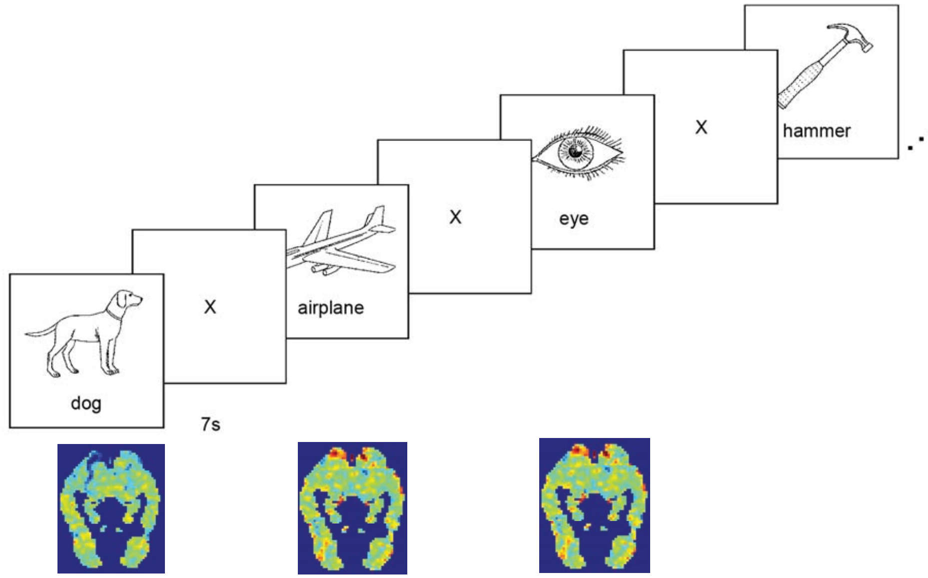
- Applications – pattern discovery
  - Brain scans – coupled matrix-tensor factorization
  - Power grid
- Applications – anomaly detection
- Algorithms
- Conclusions



Carnegie Mellon

Neuro-semantics

- **Brain Scan Data***
  - 9 persons
  - 60 nouns
- **Questions**
  - 218 questions
  - ‘is it alive?’, ‘can you eat it?’

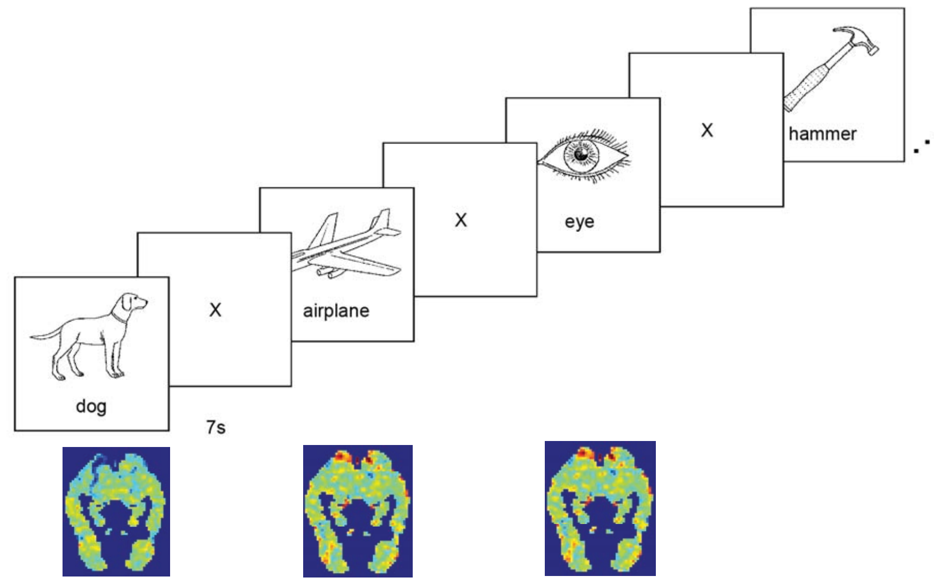


*Mitchell et al. *Predicting human brain activity associated with the meanings of nouns*. Science,2008. Data@ www.cs.cmu.edu/afs/cs/project/theo-73/www/science2008/data.html

Carnegie Mellon

Neuro-semantics

- **Brain Scan Data***
  - 9 persons
  - 60 nouns
- **Questions**
  - 218 questions
  - ‘is it alive?’, ‘can you eat it?’

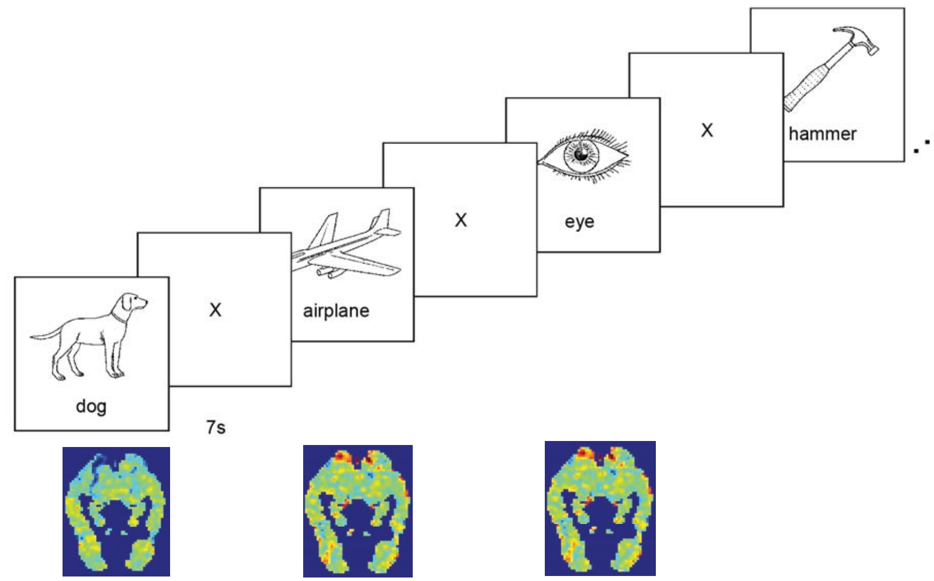


Patterns?

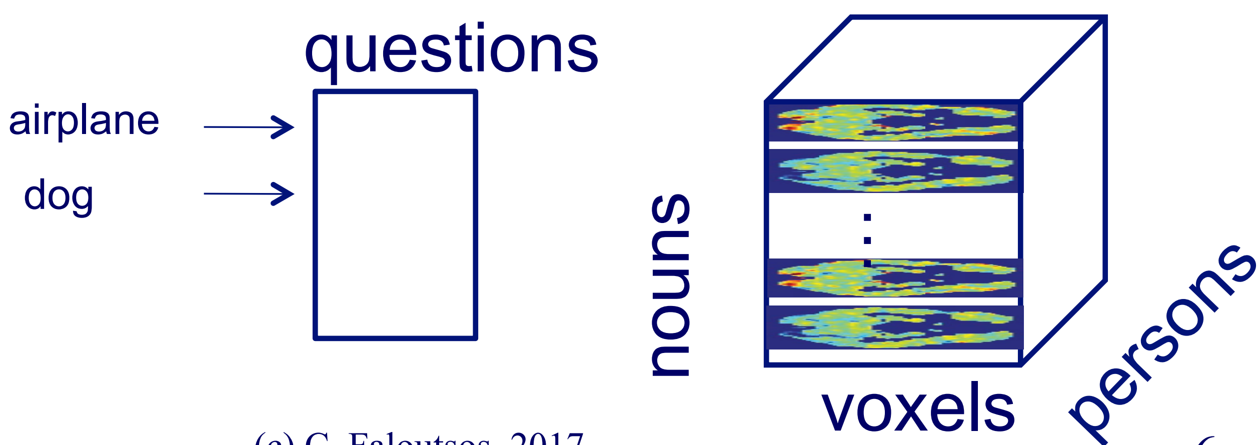
Carnegie Mellon

Neuro-semantics

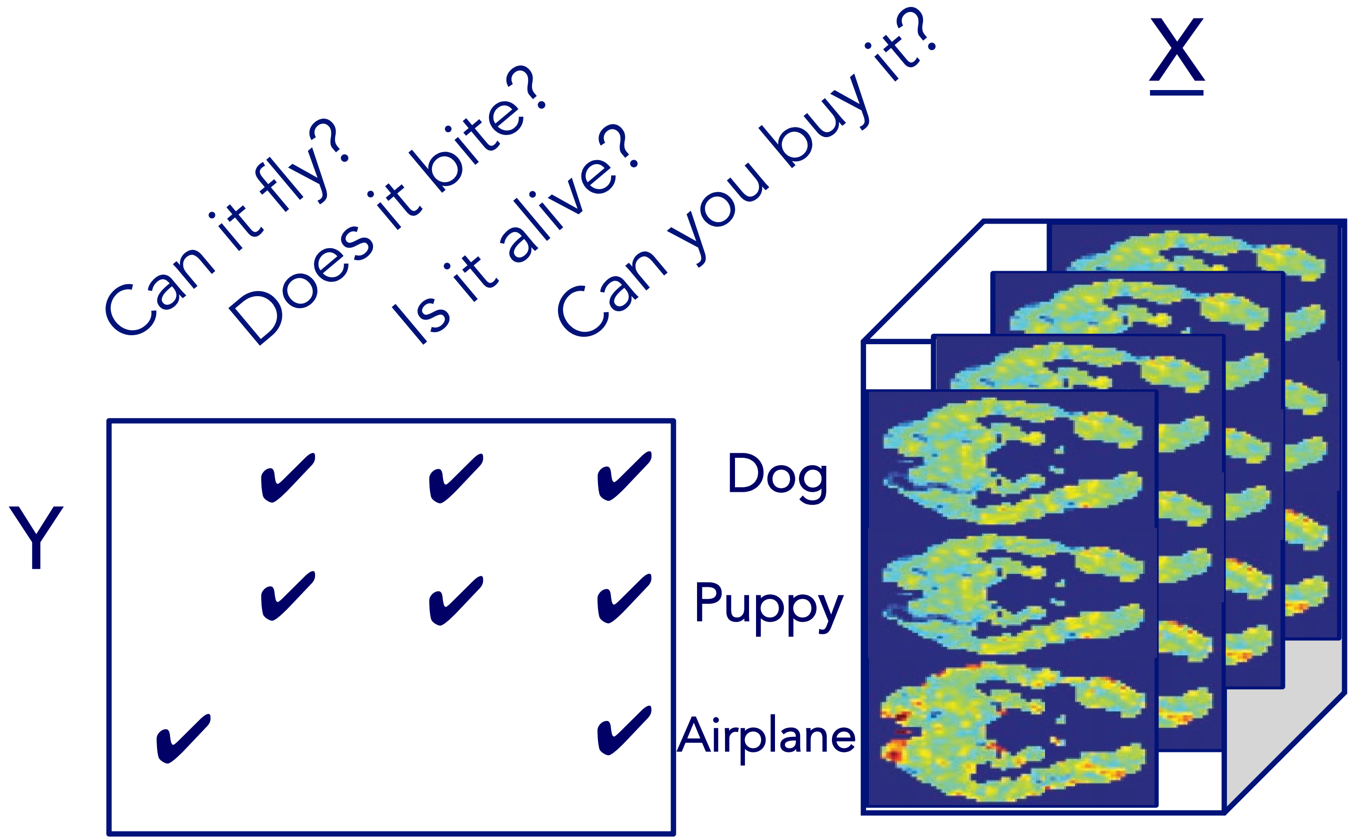
- **Brain Scan Data***
  - 9 persons
  - 60 nouns
- **Questions**
  - 218 questions
  - 'is it alive?', 'can you eat it?'



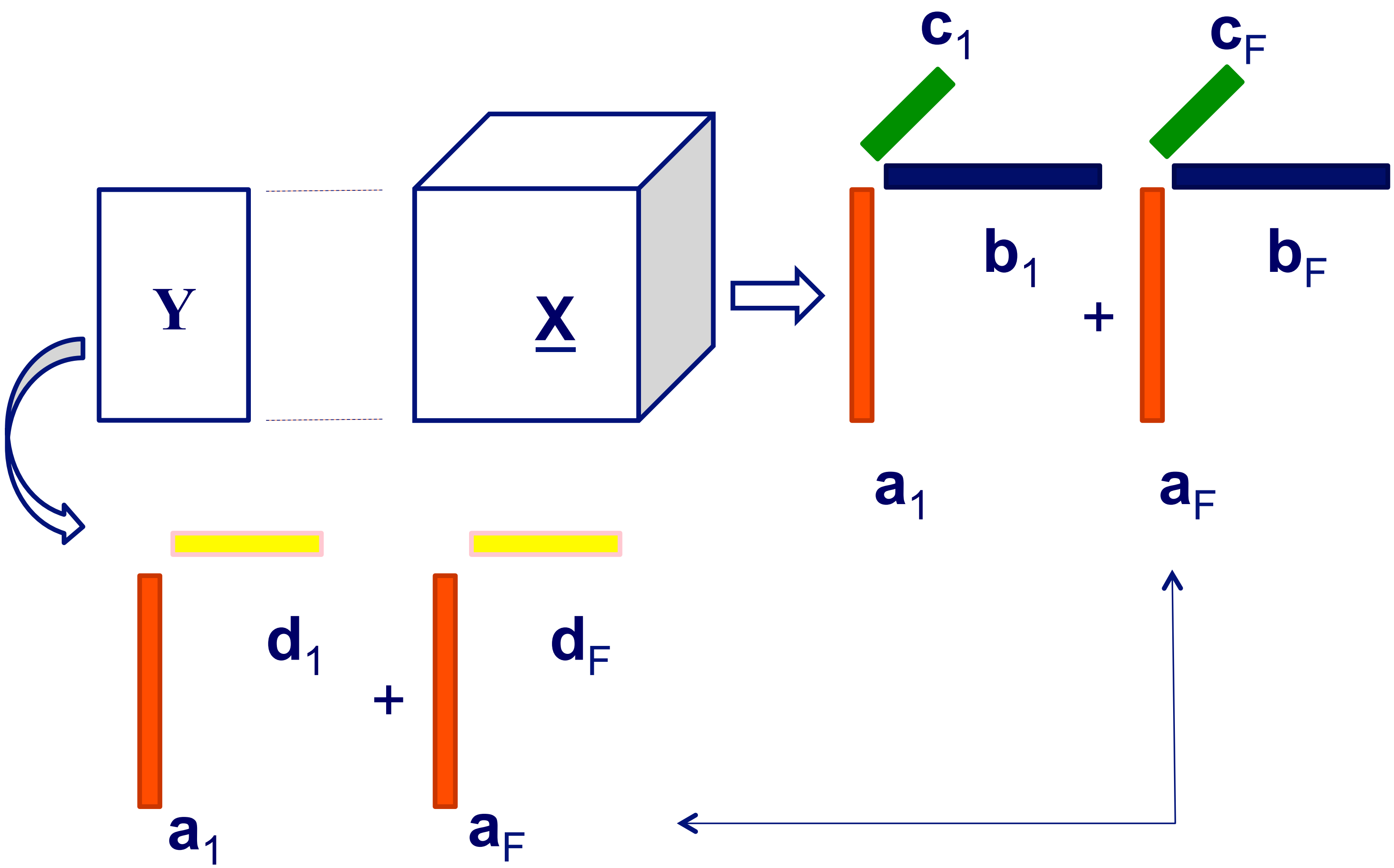
Patterns?



Neuro-semantics

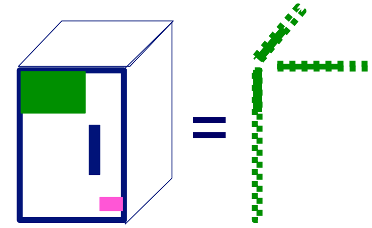


Coupled Matrix-Tensor Factorization (CMTF)



Carnegie Mellon

Neuro-semantics

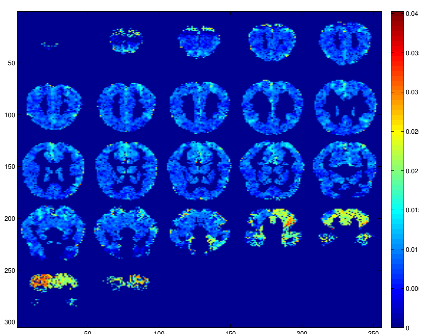


Nouns

beetle
pants
bee

Questions

can it cause you pain?
do you see it daily?
is it conscious?



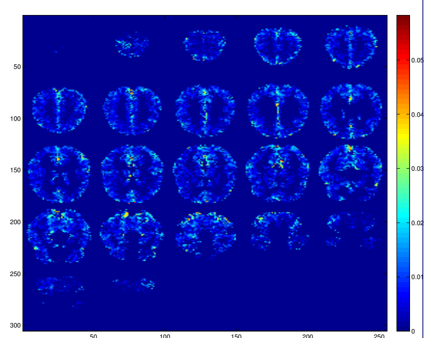
Group1

Nouns

bear
cow
coat

Questions

does it grow?
is it alive?
was it ever alive?



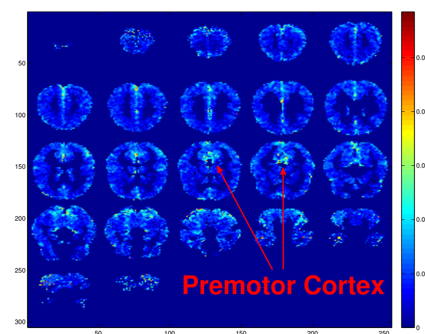
Group 2

Nouns

glass
tomato
bell

Questions

can you pick it up?
can you hold it in one hand?
is it smaller than a golfball?



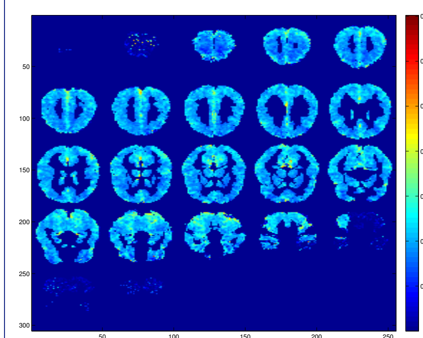
Group 3

Nouns

bed
house
car

Questions

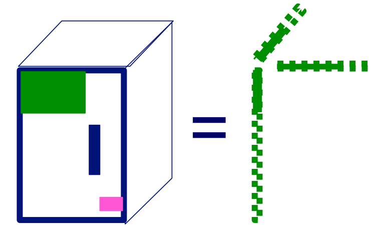
does it use electricity?
can you sit on it?
does it cast a shadow?



Group 4

Carnegie Mellon

Neuro-semantics



Nouns

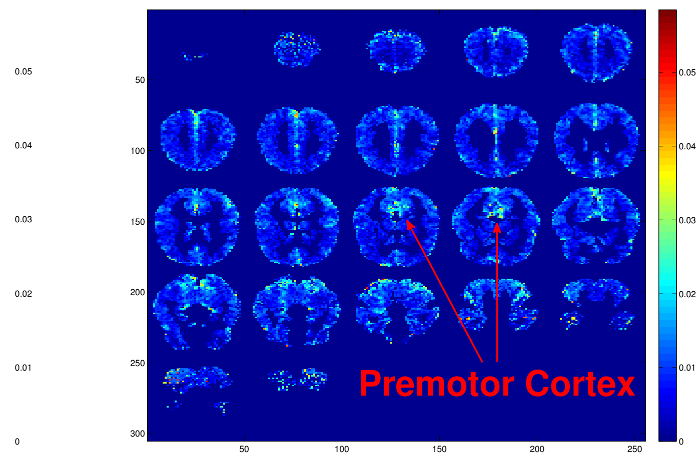
glass
tomato
bell

Questions

can you pick it up?
can you hold it in one hand?
is it smaller than a golfball?'

Small items -> Premotor cortex

- ✔ Unsupervised
- ✔ Matches intuition



Group 3

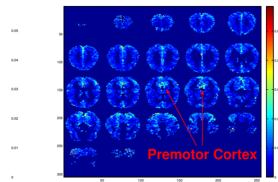
Carnegie Mellon

Neuro-semantics

Small items -> Premotor cortex

Nouns
glass
tomato
bell

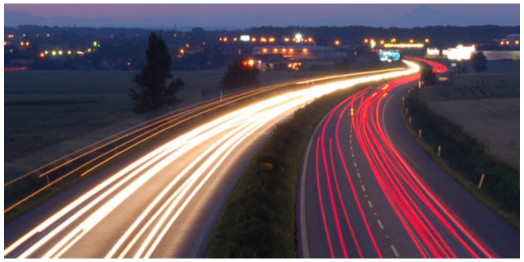
Questions
can you pick it up?
can you hold it in one hand?
is it smaller than a golfball?'



Group 3

Evangelos Papalexakis, Tom Mitchell, Nicholas Sidiropoulos, Christos Faloutsos, Partha Pratim Talukdar, Brian Murphy, *Turbo-SMT: Accelerating Coupled Sparse Matrix-Tensor Factorizations by 200x*, SDM 2014

Roadmap

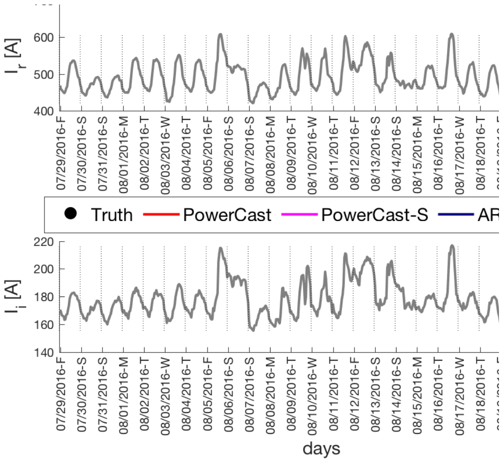


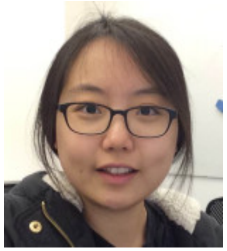
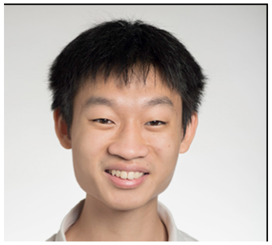
- Applications – pattern discovery
  - Brain scans – coupled matrix-tensor factorization
  - Power grid
- Applications – anomaly detection
- Algorithms
- Conclusions

Carnegie Mellon

PowerCast

Mining electric power data

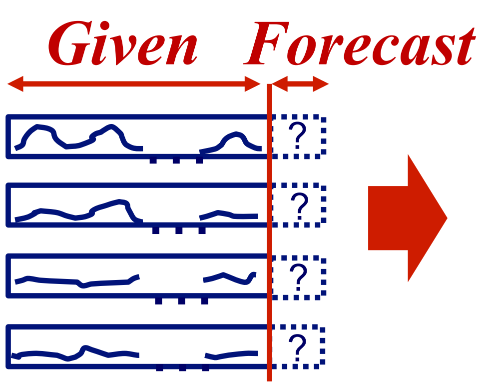



Hyun Ah Song, Bryan Hooi, Marko Jereminov, Amritanshu Pandey, Larry Pileggi, and CF, *PowerCast: Mining and Forecasting Power Grid Sequences*, PKDD’17, Skopje, FYROM

Problem definition

- **Given**: real and imaginary current and voltage
- **Forecast**: power demand in the future, and
- **Guess**: how the forecasts will change under various scenarios (e.g. population drops in half, etc)



PowerCast

What-if: *normal condition?*

What-if: *population drops?*

What-if: *more factories are built?*

Domain knowledge

$$I_r(t) = G(t)V_r(t) - B(t)V_i(t) + \alpha_r(t)$$

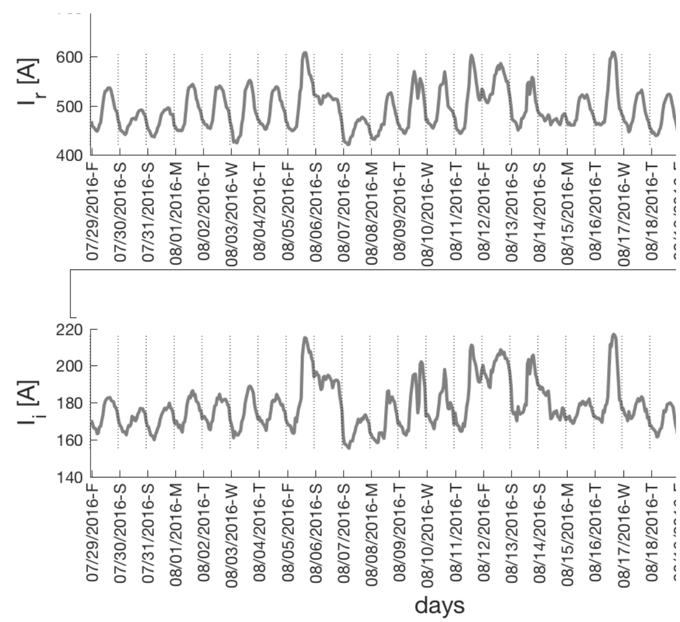
$$I_i(t) = B(t)V_r(t) + G(t)V_i(t) + \alpha_i(t)$$

resistors

Coils/capacitors

I_r

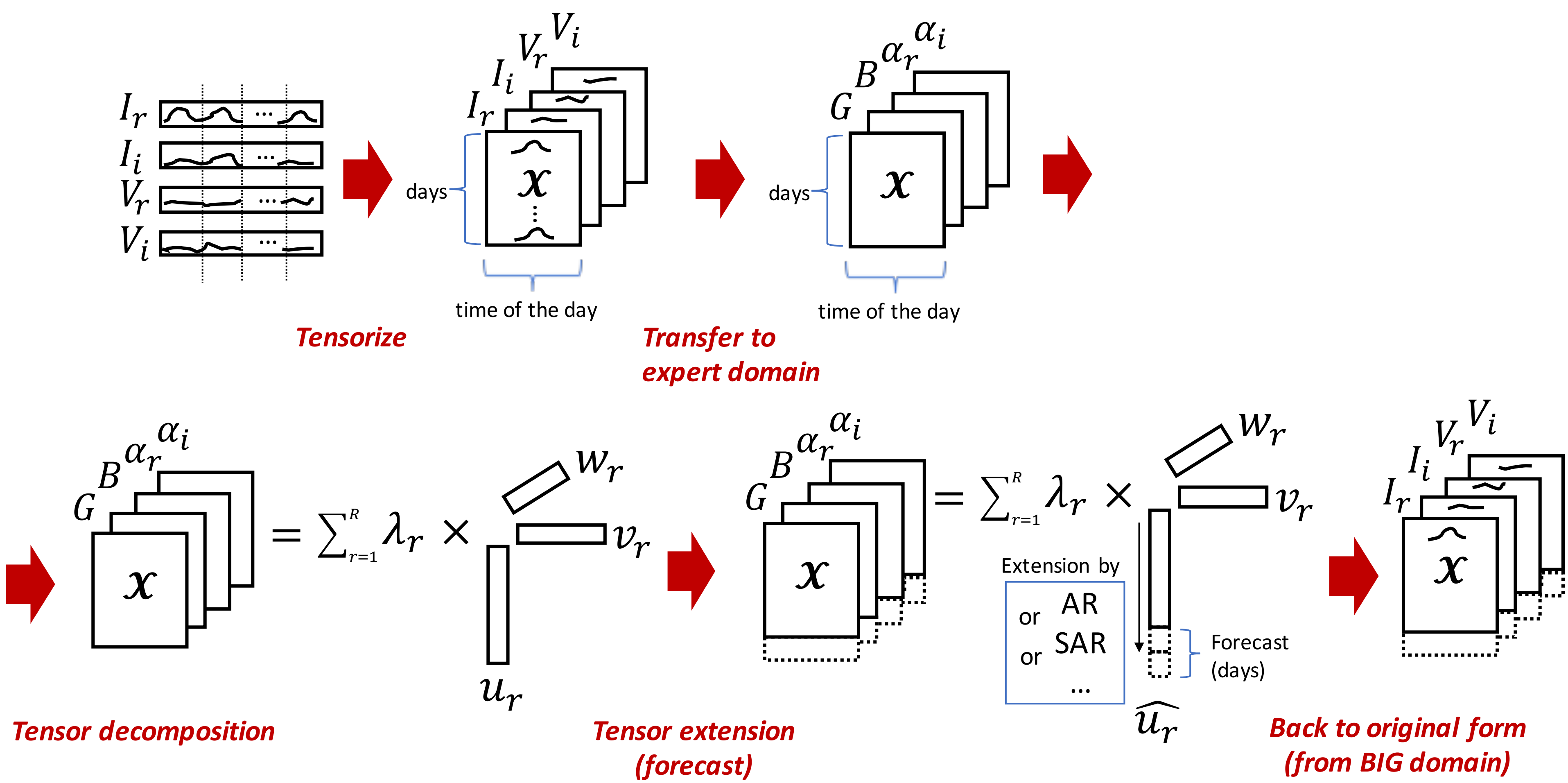
I_i



DETAILS

PowerCast

$$I_r(t) = G(t)V_r(t) - B(t)V_i(t) + \alpha_r(t)$$
$$I_i(t) = B(t)V_r(t) + G(t)V_i(t) + \alpha_i(t)$$

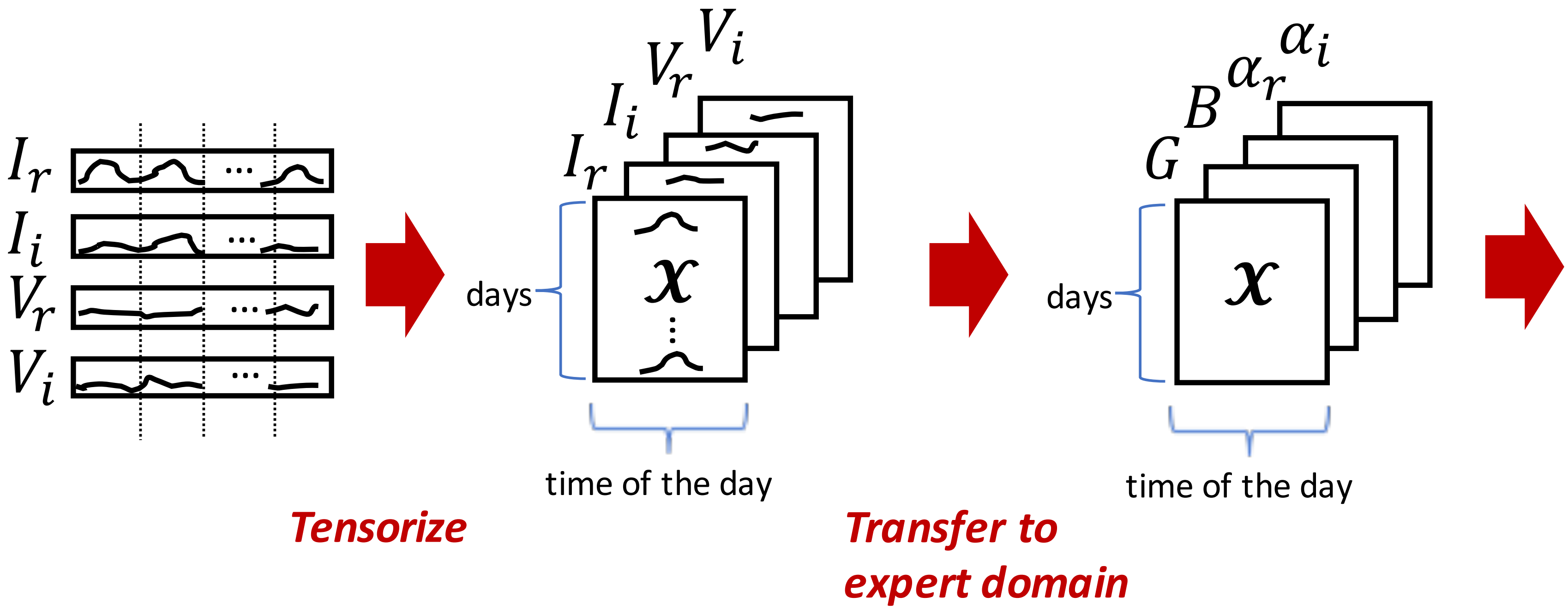


Carnegie Mellon

DETAILS

PowerCast

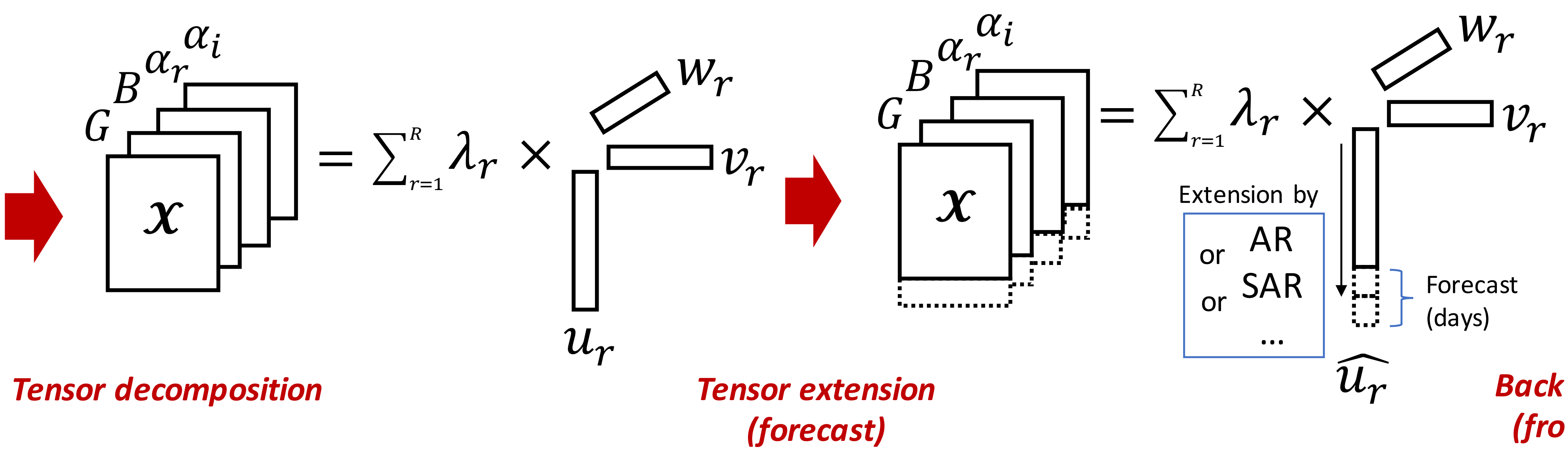
$$I_r(t) = G(t)V_r(t) - B(t)V_i(t) + \alpha_r(t)$$
$$I_i(t) = B(t)V_r(t) + G(t)V_i(t) + \alpha_i(t)$$



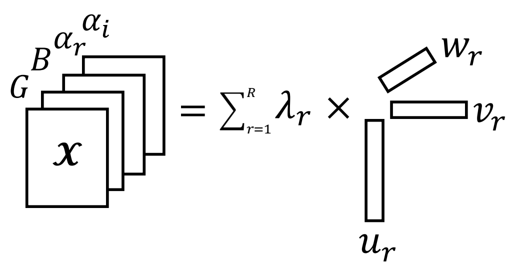
Carnegie Mellon

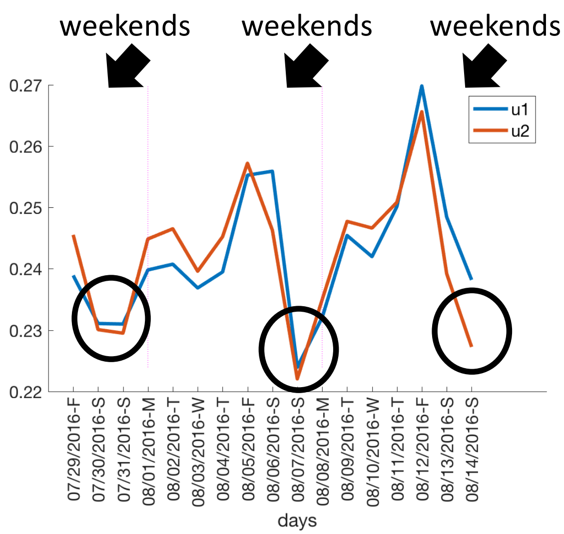
DETAILS

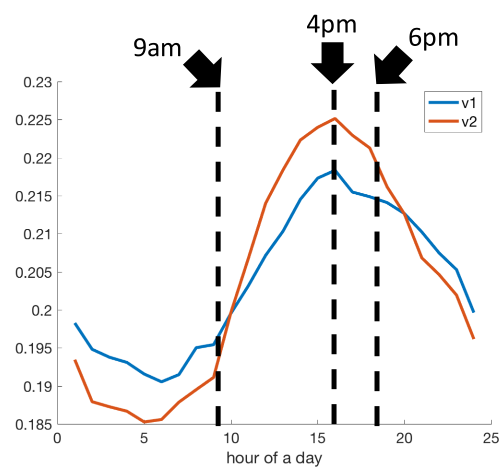
PowerCast

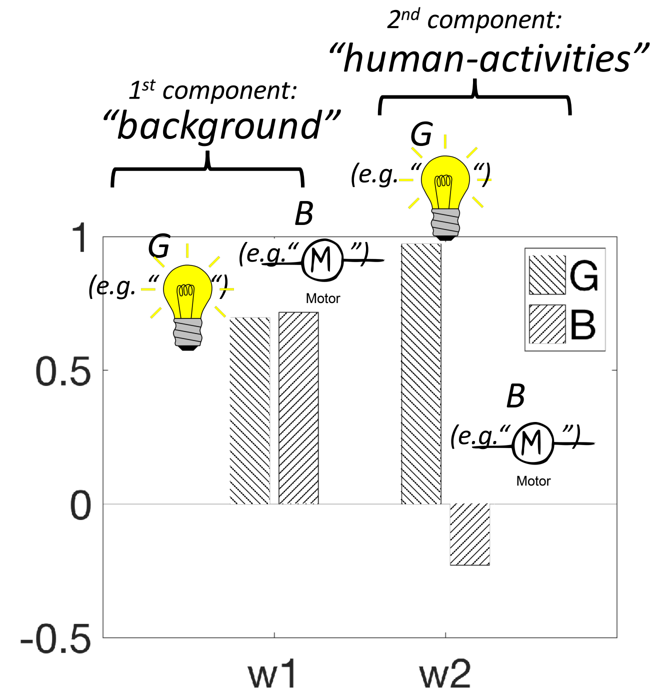


Tensor factors (concepts)







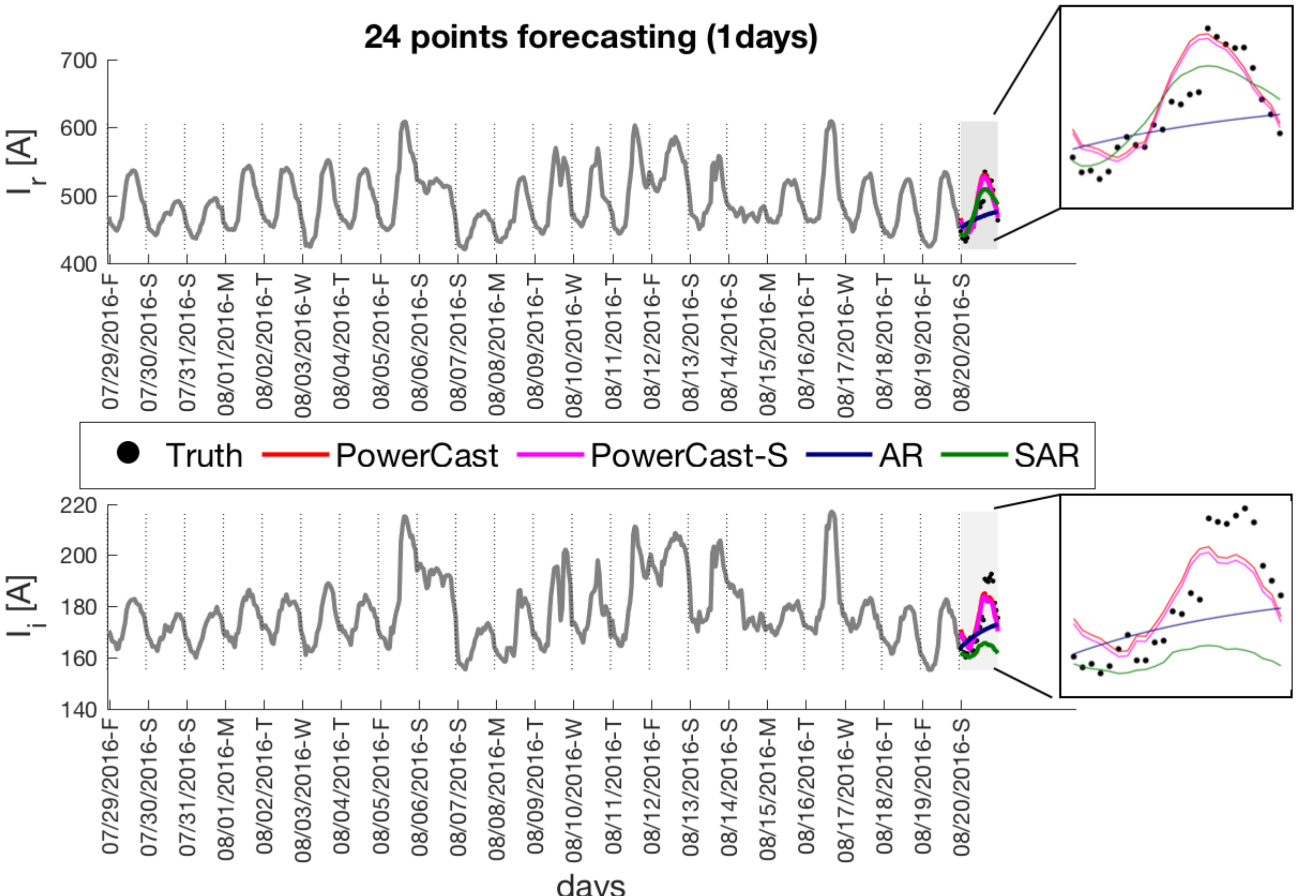


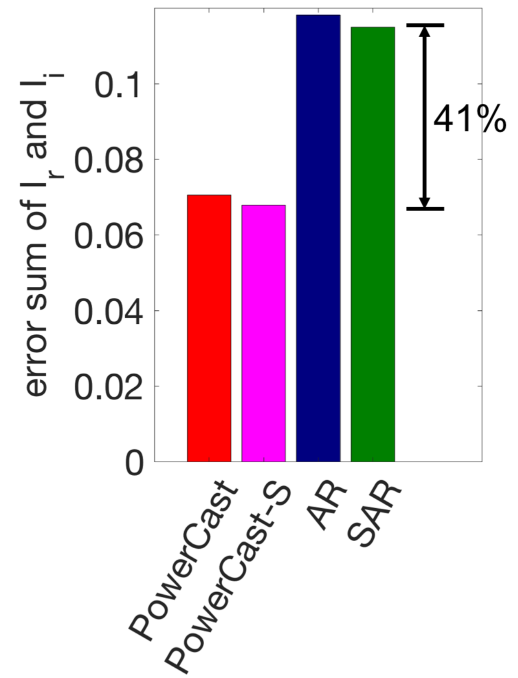
Long-term-concepts (u_r)

Daily-concepts (v_r)

User-profile-concepts (w_r)

Forecast



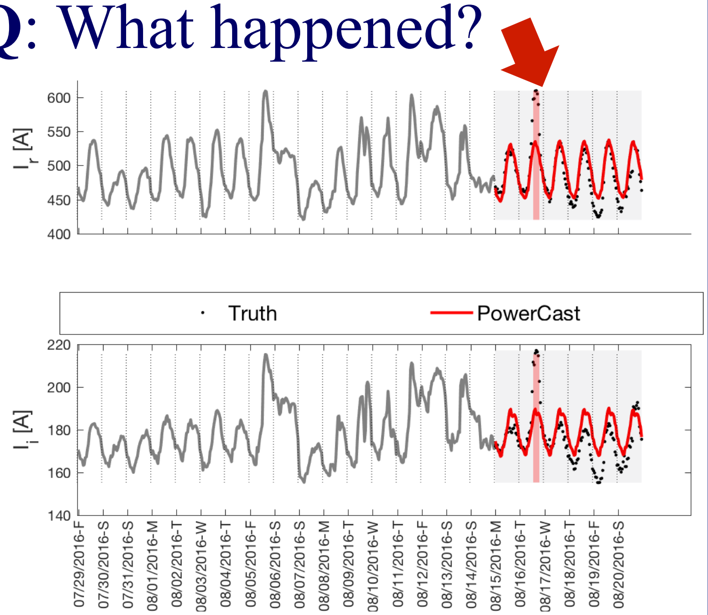


PowerCast forecasts 24-steps (1-day) ahead on CMU data more accurate than other competitors

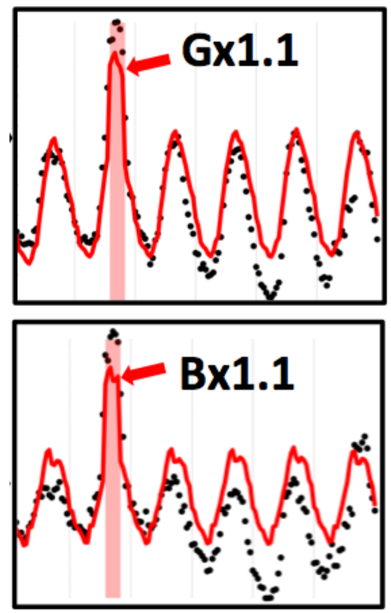
Anomaly detection & explanation

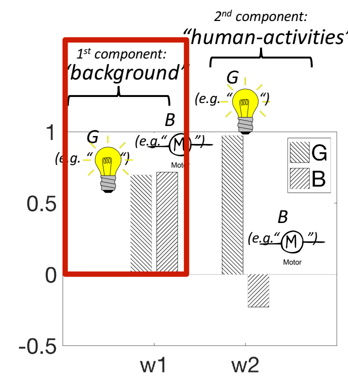
Anomaly!

Q: What happened?



A:

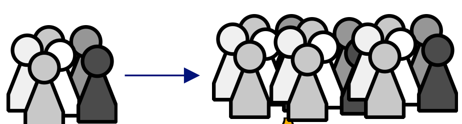




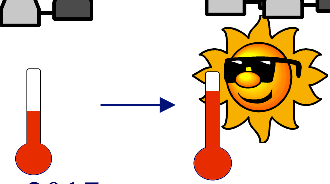
'background' component increased!

?:

Case 1?



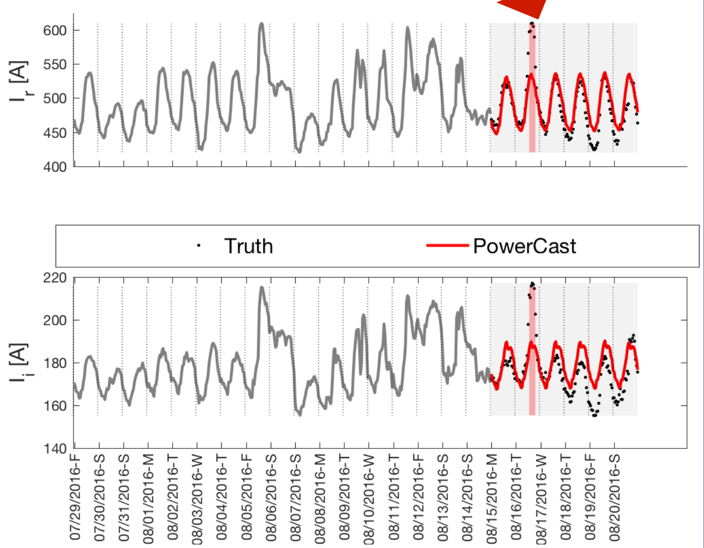
Case 2?



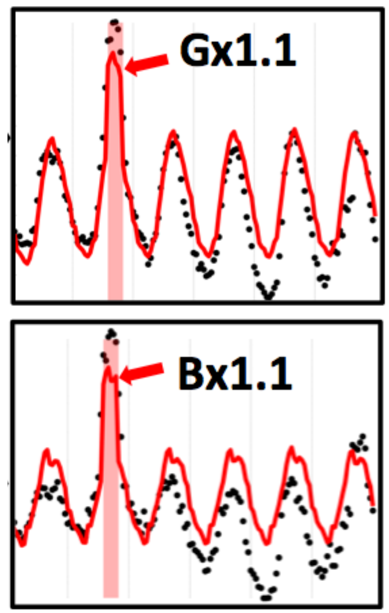
Anomaly detection & explanation

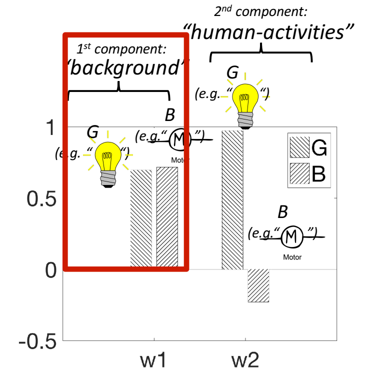
Anomaly!

Q: What happened?



A:



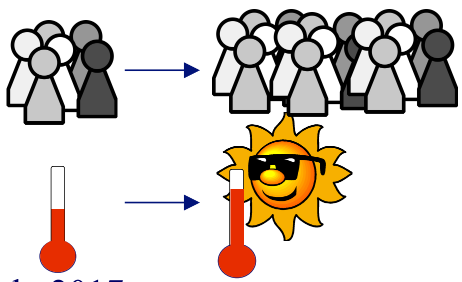


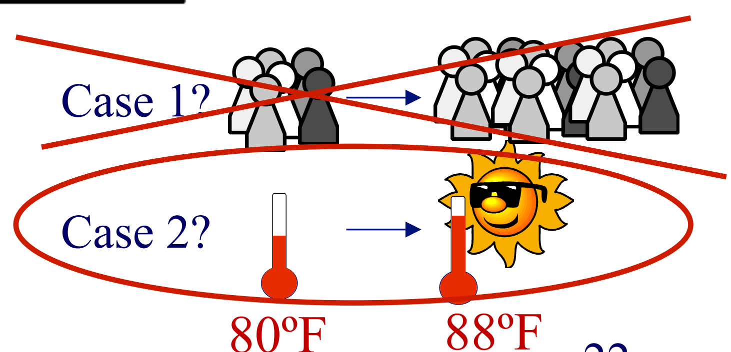
'background' component increased!

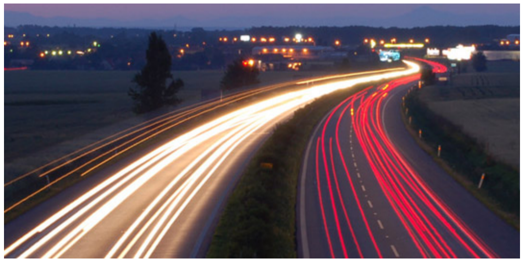
?:

Case 1?

Case 2?







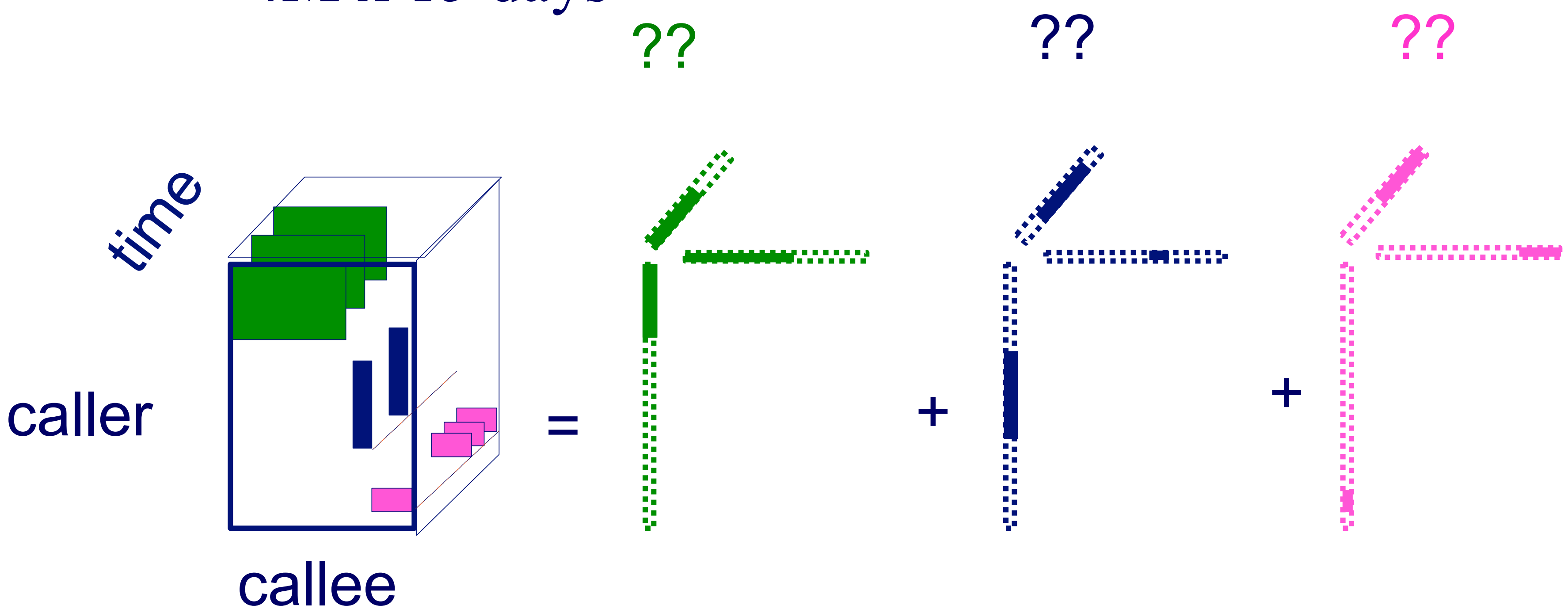
Roadmap

- Applications – pattern discovery
- Applications – anomaly detection
  - Phone-call data
  - Intrusion detection
- Algorithms
- Conclusions

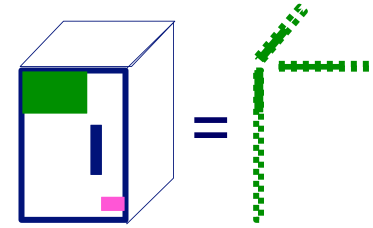
Carnegie Mellon

Anomalies in phone-call data

- PARAFAC decomposition
- Results for who-calls-whom-when
  - 4M x 15 days

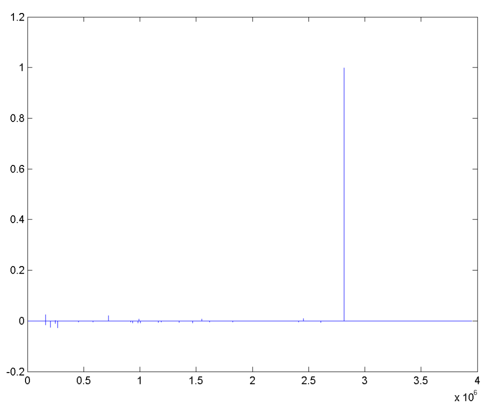


Anomaly detection in time-evolving graphs

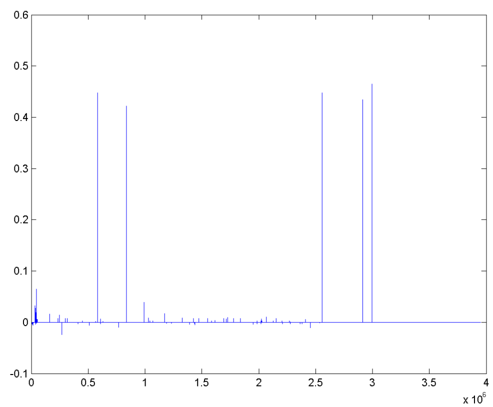


- Anomalous communities in phone call data:
  - European country, 4M clients, data over 2 weeks

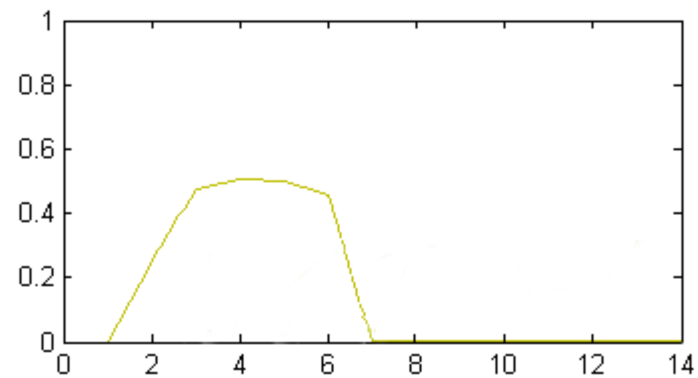
1 caller



5 receivers

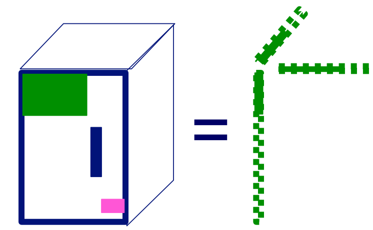


4 days of activity

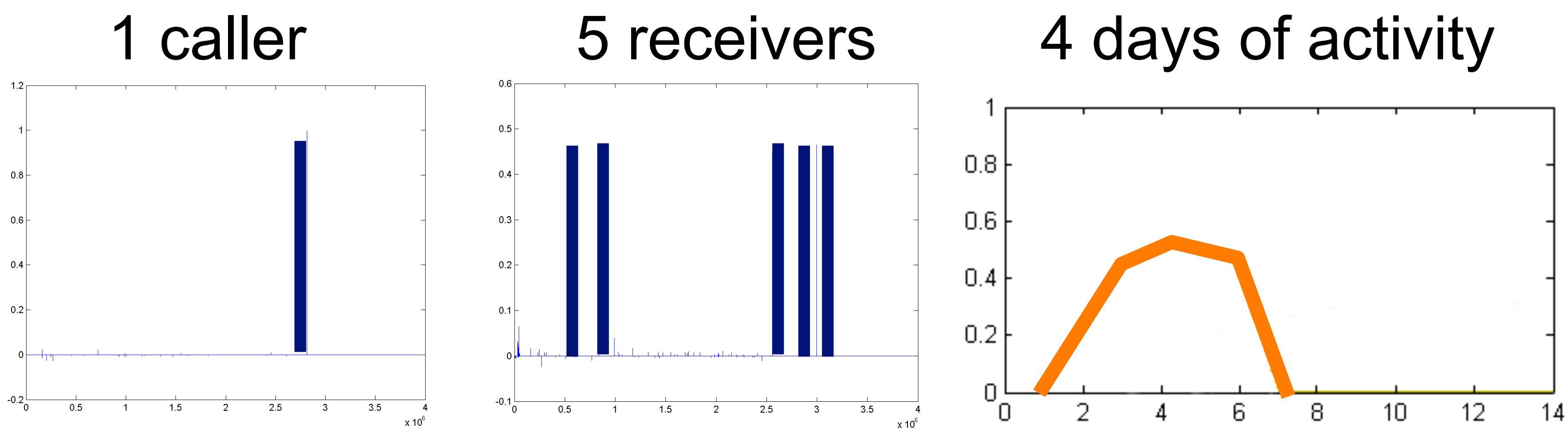


~200 calls to EACH receiver on EACH day!

Anomaly detection in time-evolving graphs



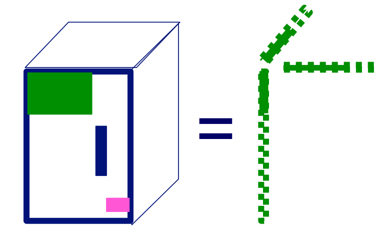
- Anomalous communities in phone call data:
  – European country, 4M clients, data over 2 weeks



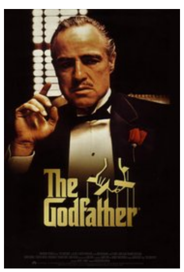
~200 calls to EACH receiver on EACH day!

Carnegie Mellon

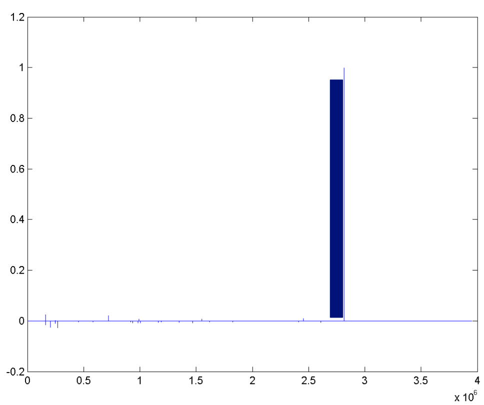
Anomaly detection in time-evolving graphs



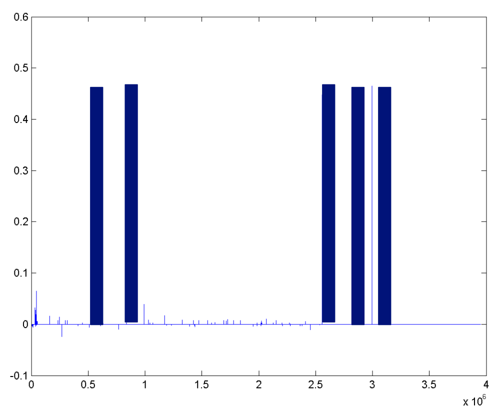
- Anomalous communities in phone call data:
  - European country, 4M clients, data over 2 weeks



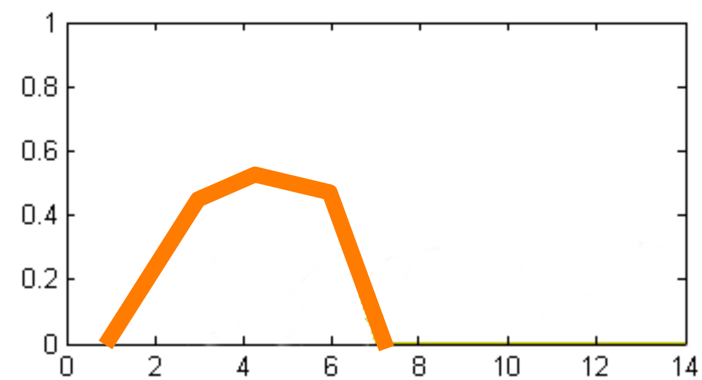
1 caller



5 receivers



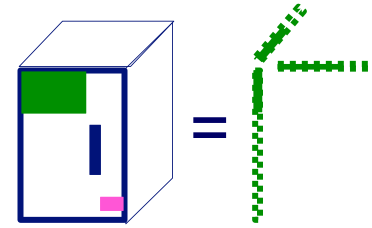
4 days of activity



~200 calls to EACH receiver on EACH day!

Carnegie Mellon

Anomaly detection in time-evolving graphs

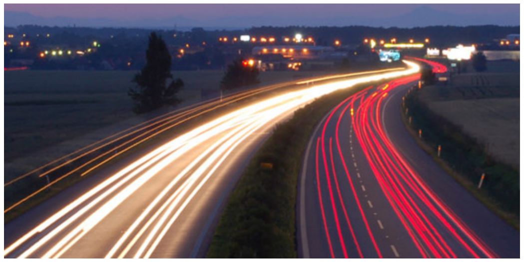


- Anomalous communities in phone call data:
  - European country, 4M clients, data over 2 weeks





Miguel Araujo, Spiros Papadimitriou, Stephan Günnemann, Christos Faloutsos, Prithwish Basu, Ananthram Swami, Evangelos Papalexakis, Danai Koutra. *Com2: Fast Automatic Discovery of Temporal (Comet) Communities*. PAKDD 2014, Tainan, Taiwan.

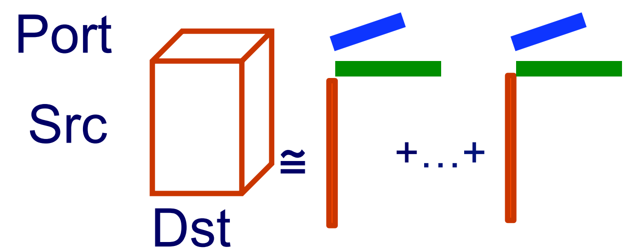


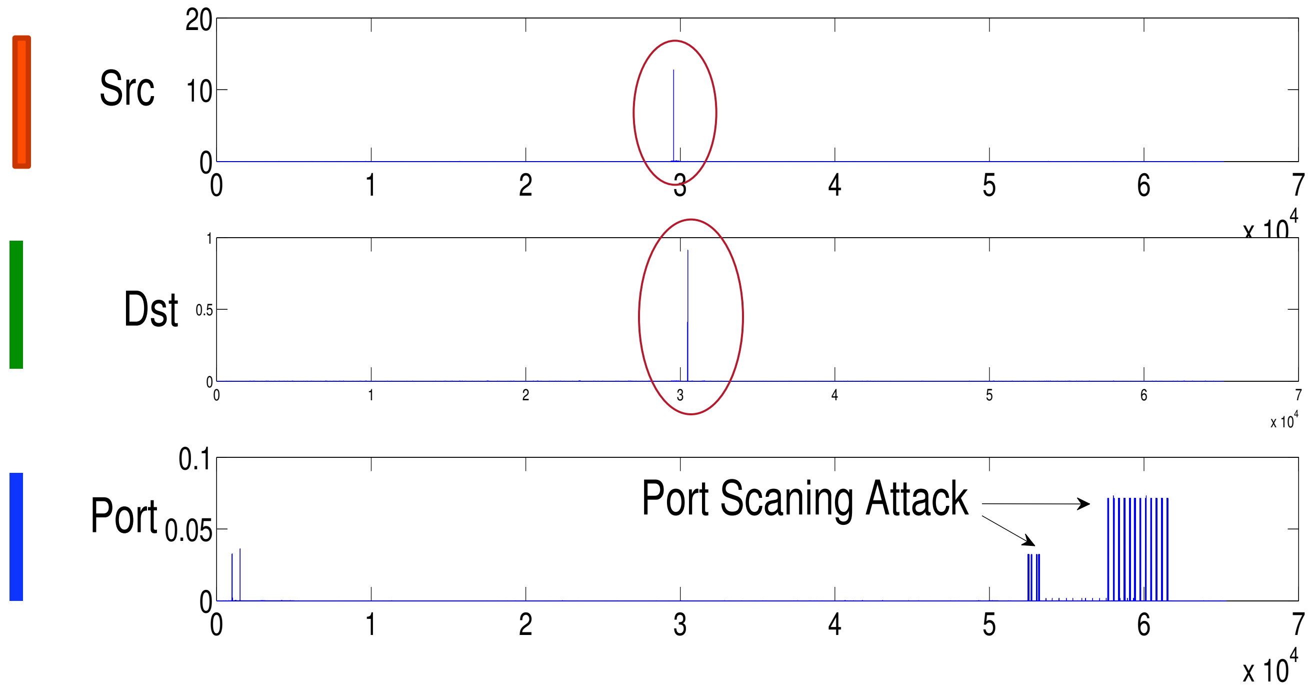
Roadmap

- Applications – pattern discovery
- Applications – anomaly detection
  - Phone-call data
  - Intrusion detection
- Algorithms
- Conclusions

ParCube at Work: Port-Scanning

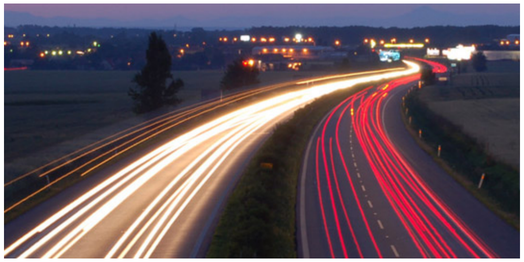
LBNL Network Data





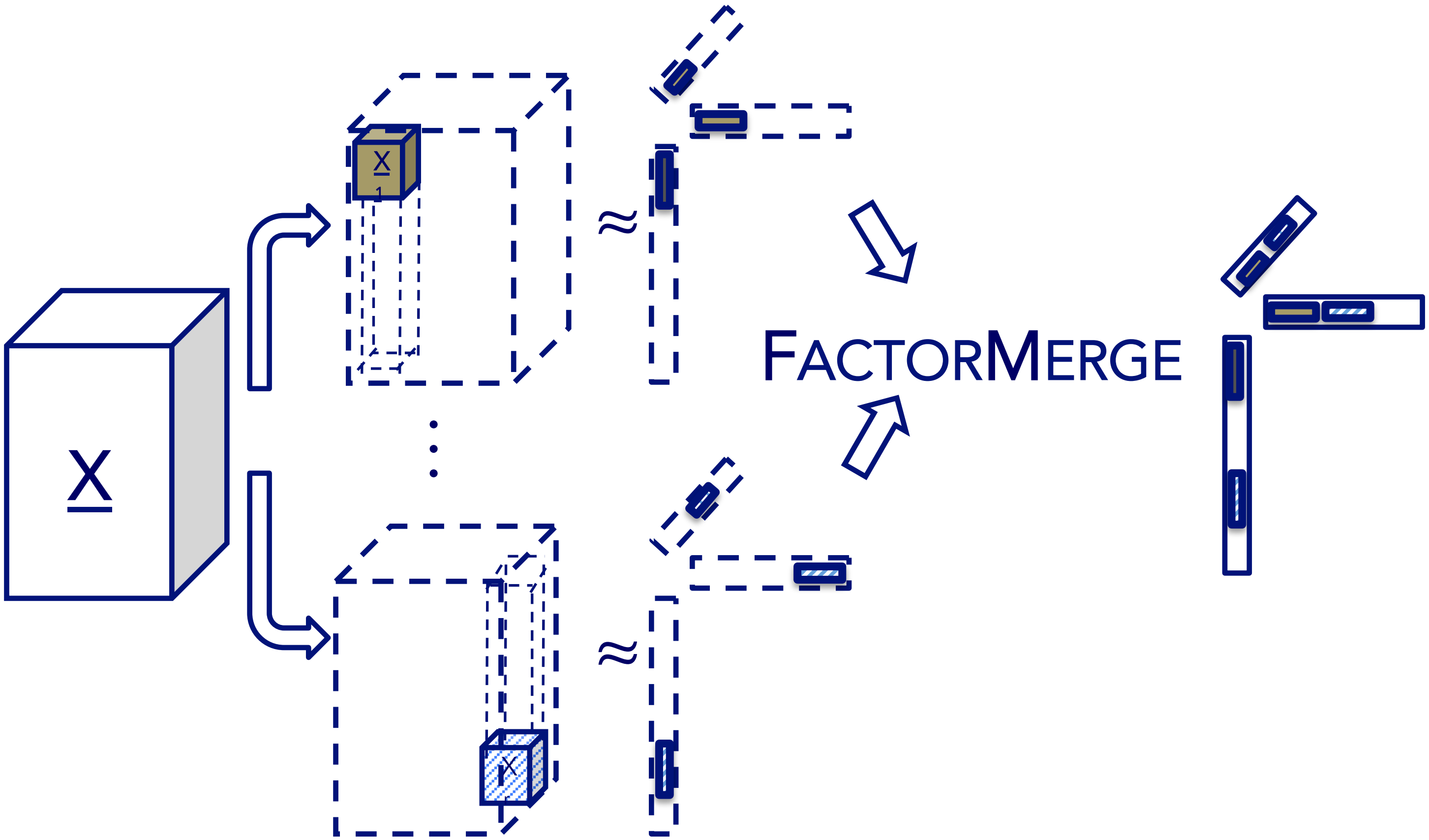
Papalexakis et al. ECML-PKDD 2012

Roadmap



- Applications – pattern discovery
- Applications – anomaly detection
- Algorithms
  - ParCube (and TurboSMT)
  - S-HOT for higher order Tucker
- Conclusions

ParCube/Turbo-SMT: Triple-sparse Parallel Tensor & Coupled Decomposition



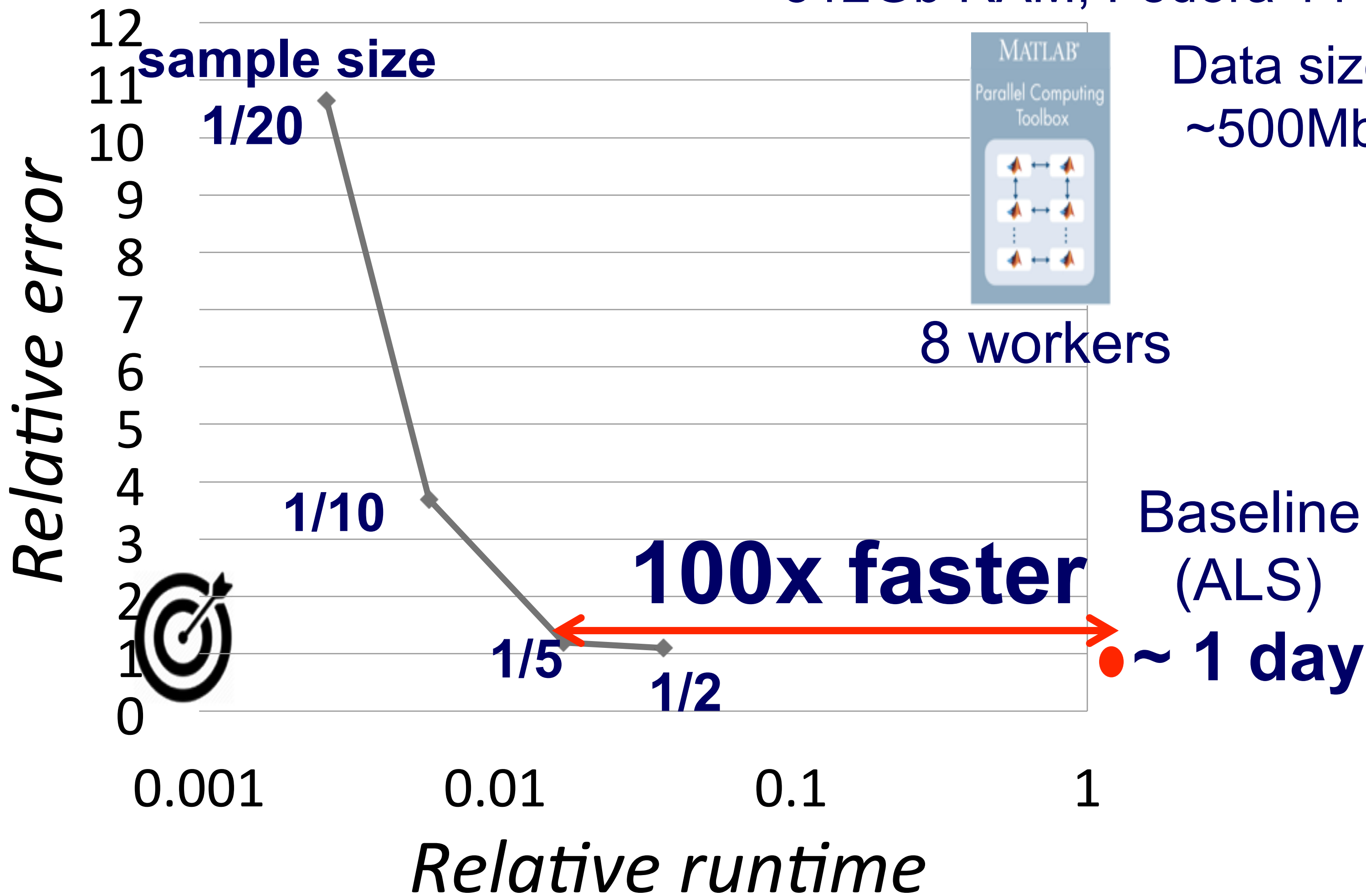
Papalexakis et al. ECML-PKDD 2012 / SDM 2014

Speedup

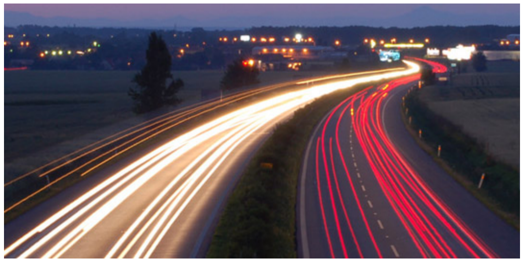
4 Intel Xeon E74850
512Gb RAM, Fedora 14

Data size
~500Mb

8 workers



Roadmap

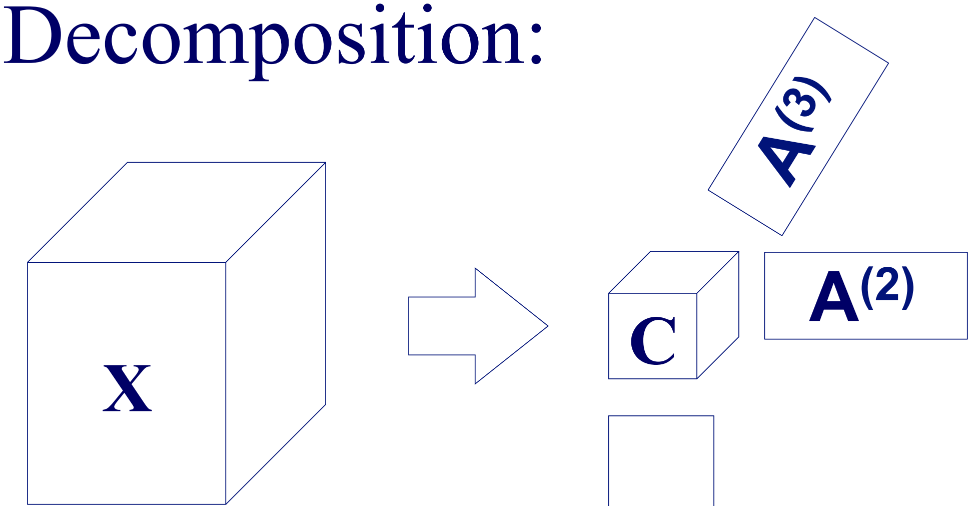


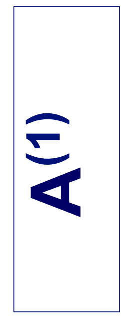
- Applications – pattern discovery
- Applications – anomaly detection
- Algorithms
  - ParCube (and TurboSMT)
  - S-HOT for higher order Tucker
- Conclusions

Carnegie Mellon

S-HOT: Scalable High-Order Tucker Decomposition

- Tucker Decomposition:







Jinoh Oh, Kijung Shin, Evangelos E. Papalexakis, Christos Faloutsos, and Hwanjo Yu, *S-HOT: Scalable High-Order Tucker Decomposition*, WSDM 2017

High-Order Tucker Decomposition (Example)

- Input tensor
  - **X**: a 5-way sparse tensor

    (size: 1M … 1M, #non-zeros: 100M)
- Output tensors:
  - **C**: a 5-way core tensor

    (size: 10 … 10, #non-zeros : ~100K)
  - **A, …, A⁽⁵⁾**: factor matrices

    (size: 1M 10, #non-zeros : ~10M)

Carnegie Mellon

High-Order Tucker Decomposition (Main idea)

Algorithm 1: Tucker-ALS

Input : $\mathcal{X}$, a N-order tensor of $\mathbb{R}^{I_1 \times \cdots \times I_N}$.
$J_1, \ldots, J_N$, the rank size of each mode.
T, the number of iterations.

Output: $\{\mathbf{A}\}$, a set of factor matrices $\{\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(N)}\}$ where $\mathbf{A}^{(n)} \in \mathbb{R}^{I_n \times J_n}$.
$\mathcal{G}$, a N-order core tensor of $\mathbb{R}^{J_1 \times \cdots \times J_N}$.

Initialize all $\mathbf{A}^{(n)}$
for $t \leftarrow 1..T$ **do**
  for $n \leftarrow 1..N$ **do**
    $\mathbf{Y}_{(n)} \leftarrow [\mathcal{X} \times_{-n} \{\mathbf{A}^T\}]_{(n)}$
    $\mathbf{A}^{(n)} \leftarrow$ top-J_n left singular vectors of $\mathbf{Y}_{(n)}$
$\mathcal{G} \leftarrow \mathbf{Y}_{(N)} \times_N \mathbf{A}^{(N)T}$
return $\mathcal{G}$, $\{\mathbf{A}\}$

Huge, & dense

High-Order Tucker Decomposition (Main idea)

Algorithm 1: Tucker-ALS

Input : $\mathcal{X}$, a N-order tensor of $\mathbb{R}^{I_1 \times \cdots \times I_N}$.
$J_1, \ldots, J_N$, the rank size of each mode.
T, the number of iterations.

Output: $\{\mathbf{A}\}$, a set of factor matrices $\{\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(N)}\}$ where $\mathbf{A}^{(n)} \in \mathbb{R}^{I_n \times J_n}$.
$\mathcal{G}$, a N-order core tensor of $\mathbb{R}^{J_1 \times \cdots \times J_N}$.

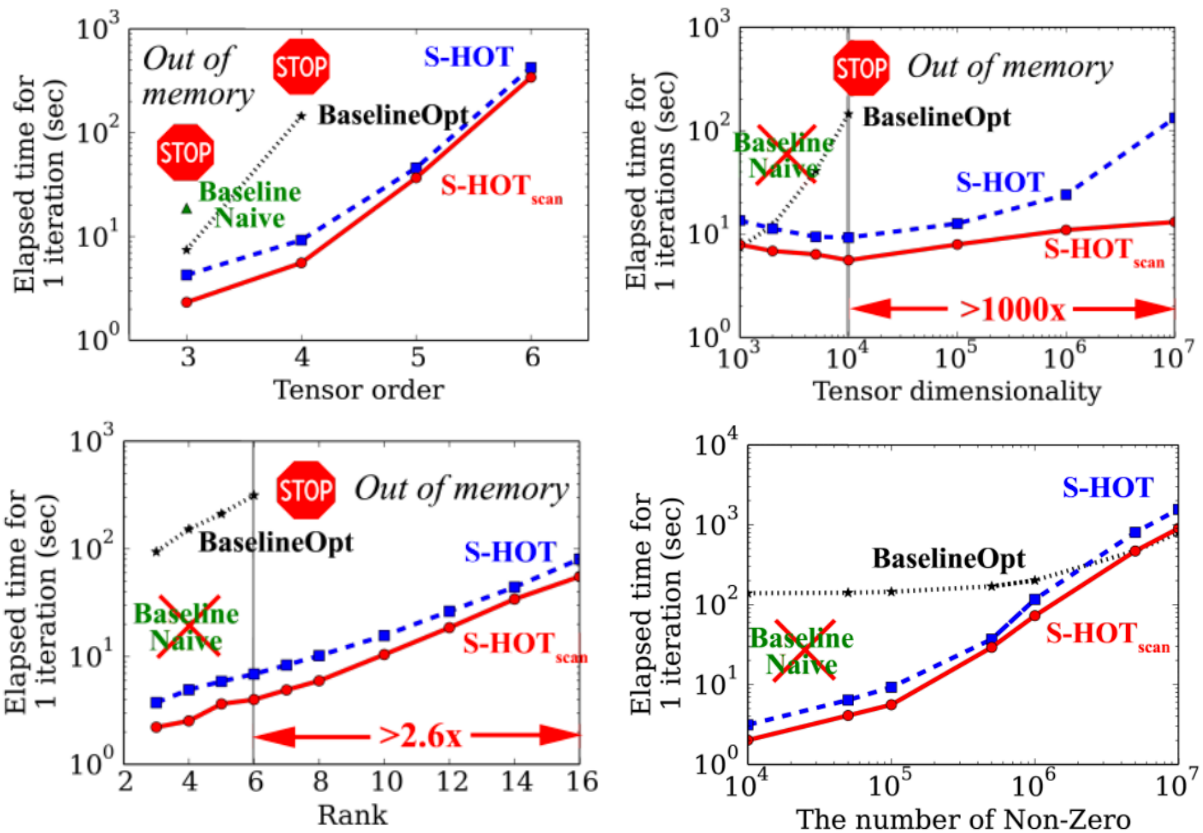
```
Initialize all A^(n)
for t ← 1..T do
    for n ← 1..N do
        Y_(n) ← [X ×_{-n} {A^T}]_(n)
        A^(n) ← top-J_n left singular vectors of Y_(n)
G ← Y_(N) ×_N A^(N)T
return G, {A}
```

Huge, & dense

-> DON'T materialize it

Scalability of S-HOT

- Baselines: Standard ALS (Naïve), & (opt)



Scalability of S-HOT

- Baselines: Standard ALS (Naïve), & (opt)
- I = 1M; J=10; M=1B; N=5 modes

Method	Space Requirements for Intermediate Data (in Theory)	(in Example)
BaselineNaive	MJ^{N-1}	$\sim$ 40TB
BaselineOpt [12]	IJ^{N-1}	$\sim$ 40GB
HaTen2 [8]	$\max(IJ^{N-1}, M(N-1)J)$	$\sim$ 160GB
S-HOT	$\max(I, J^{N-1})$	$\sim$ 4MB
S-HOT$_{scan}$	J^{N-1}	$\sim$ 40KB

Discovery using S-HOT

- Microsoft Academic Graph

(42M papers × 25K venues × 115M authors × 54K keywords)

CS-related	International Conference on Networking(ICN), Wired/Wireless Internet Communications(WWIC), Database and Expert Systems Applications(DEXA), Data Mining and Knowledge Discovery, IEEE Transactions on Robotics, ...
Nanotech.	Nature Nanotechnology, PLOS ONE, Journal of Experimental Nanoscience, Journal of Nanoscience and Nanotechnology, Journal of Semiconductors, Trends in Biotechnology, ...
Clinical	European Journal of Cancer, PLOS Biology, Clinical and Applied Thrombosis-Hemostasis, Journal of Infection Prevention, RBMC Clinical Pharmacology, Regional Anesthesia and Pain ...

Carnegie Mellon

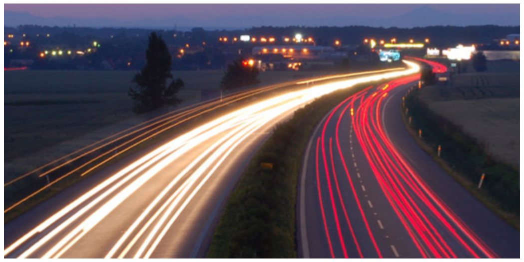
Discovery using S-HOT

- Microsoft Academic Graph

(42M papers × 25K venues × 115M authors × 54K keywords)

CS-related	International Conference on Networking(ICN), Wired/Wireless Internet Communications(WWIC), Database and Expert Systems Applications(DEXA), Data Mining and Knowledge Discovery, IEEE Transactions on Robotics, . . .

Carnegie Mellon



Roadmap

- Applications – pattern discovery
- Applications – anomaly detection
- Algorithms
  - ParCube (and TurboSMT)
  - S-HOT for higher order Tucker
- Conclusions

Carnegie Mellon

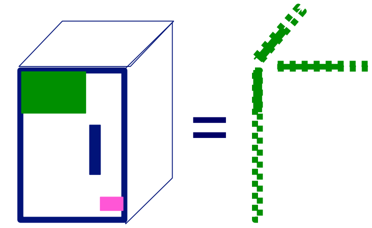
Thanks

Disclaimer: All opinions are mine; not necessarily reflecting the opinions of the funding agencies

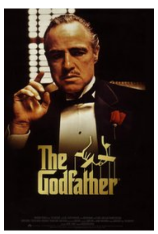
Thanks to: NSF IIS-0705359, IIS-0534205, CTA-INARC; Yahoo (M45), LLNL, IBM, SPRINT, Google, INTEL, HP, iLab

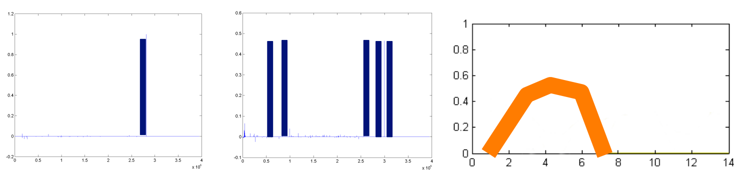
Carnegie Mellon

CONCLUSION#1



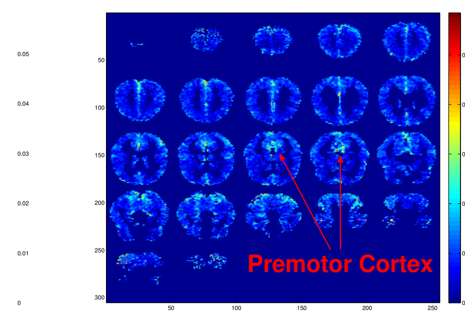
- **MANY** applications for tensors





Nouns
glass
tomato
bell

Questions
can you pick it up?
can you hold it in one hand?
is it smaller than a golfball?'

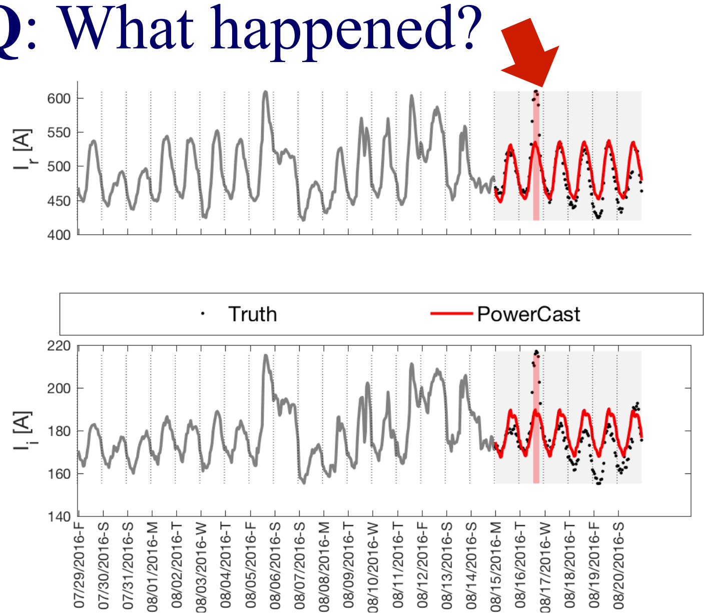


Carnegie Mellon

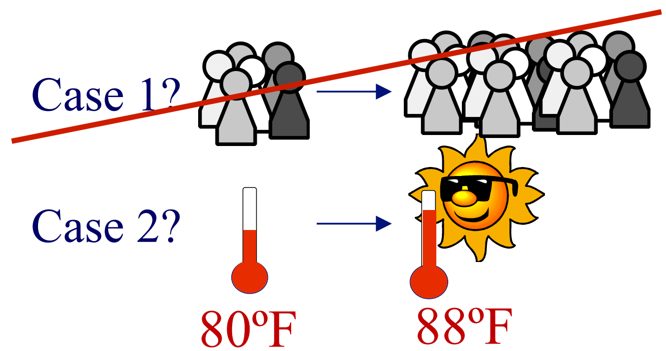
CONCLUSION#2

- Domain-experts: valuable

Q: What happened?



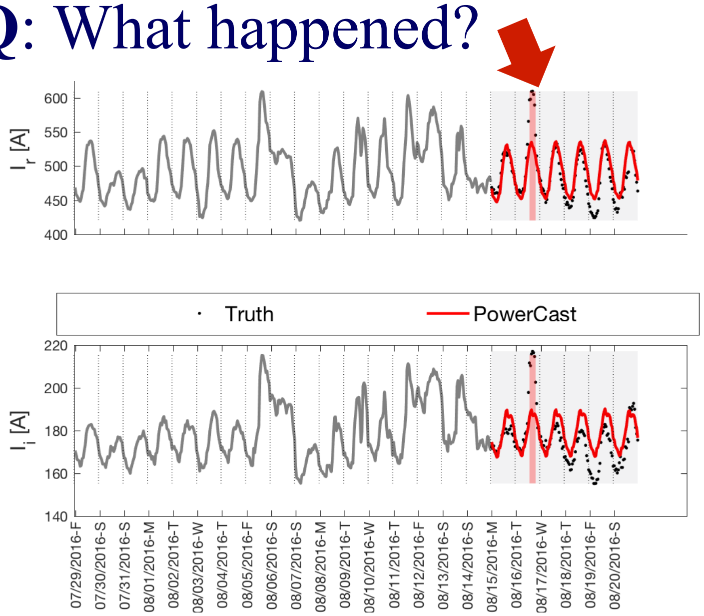
$$I_r(t) = G(t)V_r(t) - B(t)V_i(t) + \alpha_r(t)$$
$$I_i(t) = B(t)V_r(t) + G(t)V_i(t) + \alpha_i(t)$$



Thank you!

- Domain-experts: valuable

Q: What happened?



$$I_r(t) = G(t)V_r(t) - B(t)V_i(t) + \alpha_r(t)$$
$$I_i(t) = B(t)V_r(t) + G(t)V_i(t) + \alpha_i(t)$$