

# CatchTartan: Representing and Summarizing Dynamic Multicontextual Behaviors



Meng Jiang<sup>1</sup>, Christos Faloutsos<sup>2</sup>, Jiawei Han<sup>1</sup>

<sup>1</sup> Department of Computer Science, University of Illinois at Urbana-Champaign, IL, USA

<sup>2</sup> Computer Science Department, SCS, Carnegie Mellon University, PA, USA

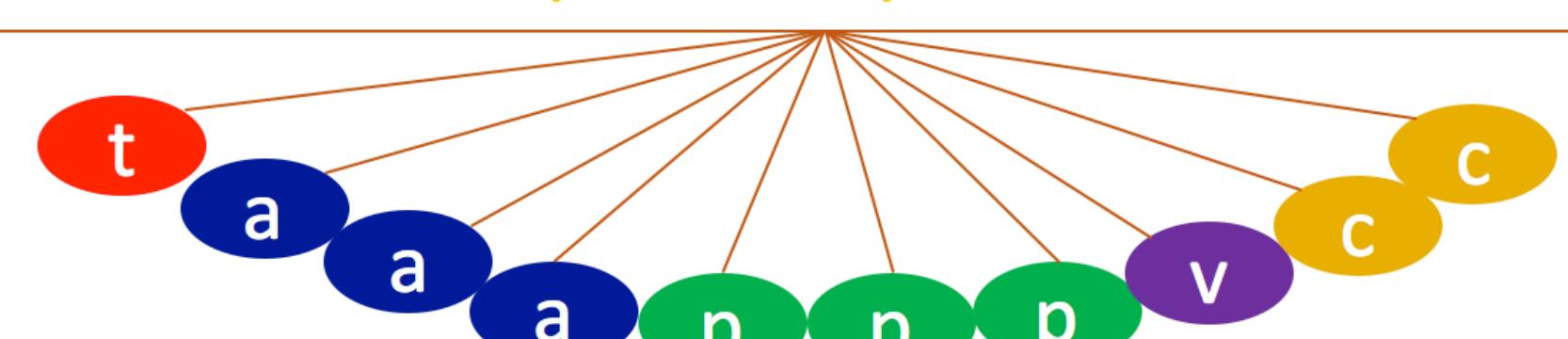


## Behavior and Behavioral Summary

What is Behavior?

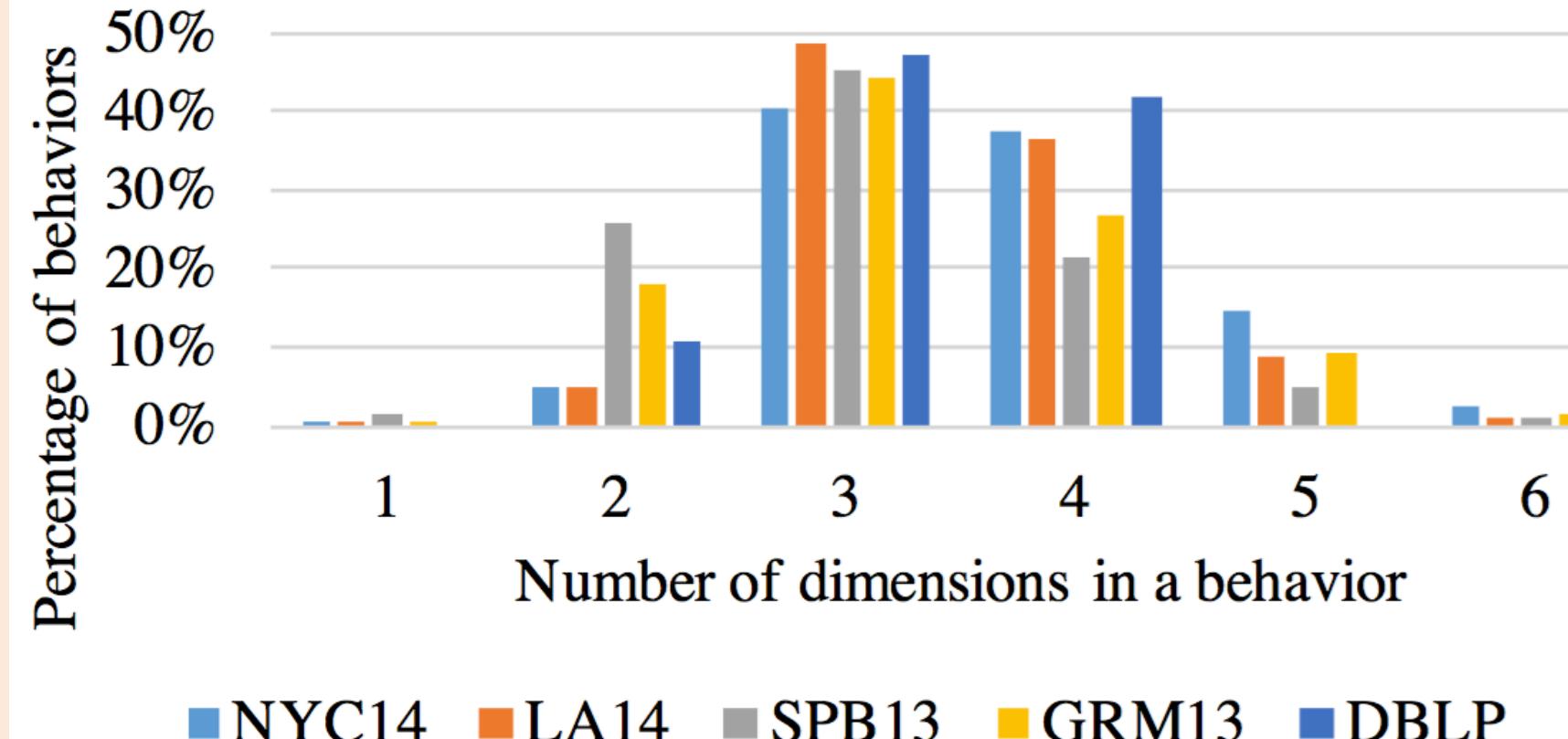
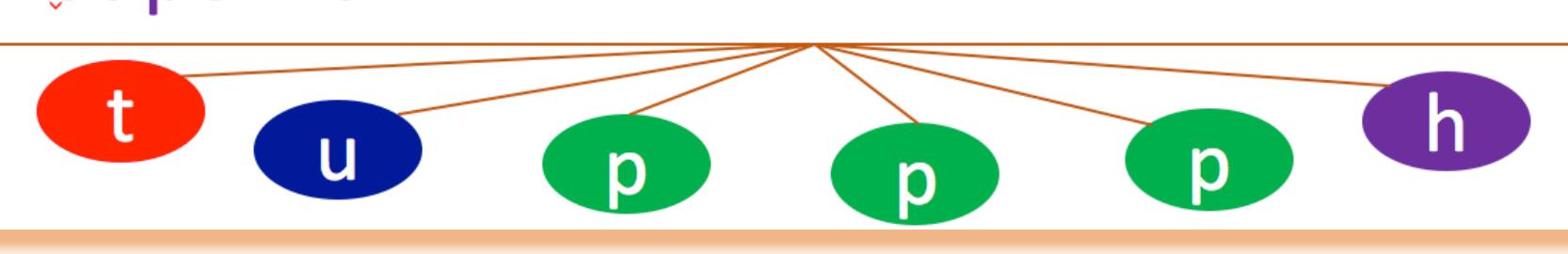
- Publishing-paper behavior

2009 P. Melville, W. Gryc, R. Lawrence,  
“Sentiment analysis of blogs by combining  
lexical knowledge with text classification”,  
KDD’09. Refs: p81623, p84395...



- Tweeting behavior

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



What is Behavioral Summary?

- Summarizing research trends

Author	Venue	Keyword	Cited	#Paper	Venue	Keyword	#Paper
76 Cheng-xiang Zhai Hui Fang S. Kambhampati	7 SIGIR VLDB TKDE	7 “information retrieval” “data integration” “text classification”	68 p56743 <sup>1</sup> p62995 p76869	32 2003-2007	5 ICML NIPS ...	6 “reinforcement learning” “machine learning”	40 1997-2002
6 Jiawei Han Xifeng Yan	1 SIGMOD	1 p76095 <sup>2</sup>	22 2004-2010	3 ICDM AAAI TKDE	1 “anomaly detection”	25 2005-2013	27 C. Faloutsos J. Pei P. S. Yu X. Lin C. Aggarwal...
12 Ryan White Hang Li Tie-Yan Liu Zhaohui Zheng...	5 SIGIR WWW WSDM CIKM...	3 “web search” “click-through data” “sponsored search”	12 p82630 <sup>3</sup> p116290 p103899 p106191...	32 2006-2013	6 KDD ICDM ICDE TKDE	12 “large graphs” “data streams” “evolving data” “evolving graphs”	70 2006-2013
8 Qiang Yang Duo Shen Sinno Pan...	3 KDD PAKDD AAAI	6 “transfer learning” “data mining” “localization models”	17 2007-2010	8 Qiang Yang Duo Shen Sinno Pan...	3 KDD PAKDD AAAI	6 “transfer learning” “data mining” “localization models”	17 2007-2010

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

- Summarizing “Super Bowl” tweets (see below)

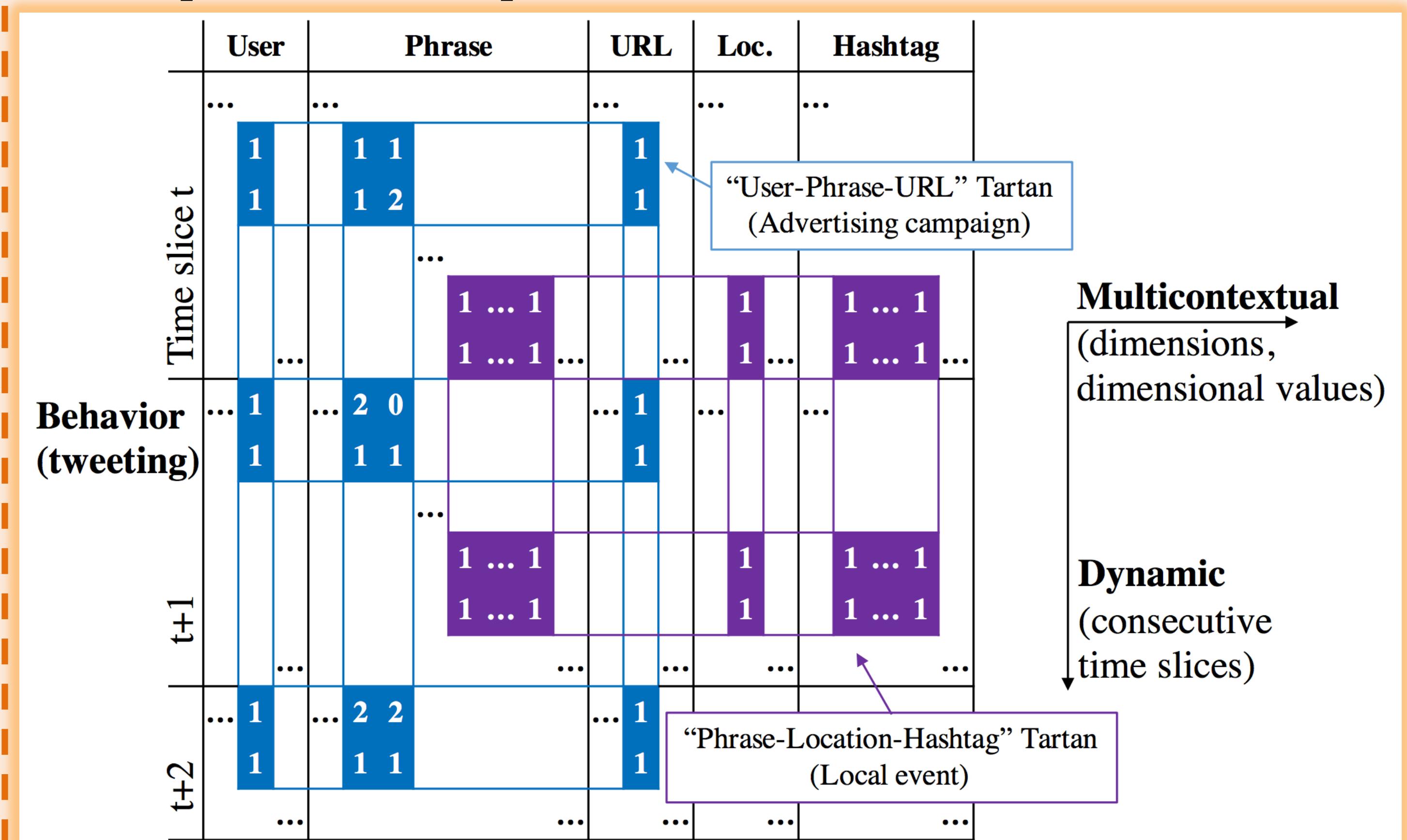
## Q2: How to represent behavioral summary?

## Representation of Behavior

Tensor fails...

- Require one value each dimension!
- Empty values?
- Set values?
- Timestamps?

Our Proposed Novel Representations



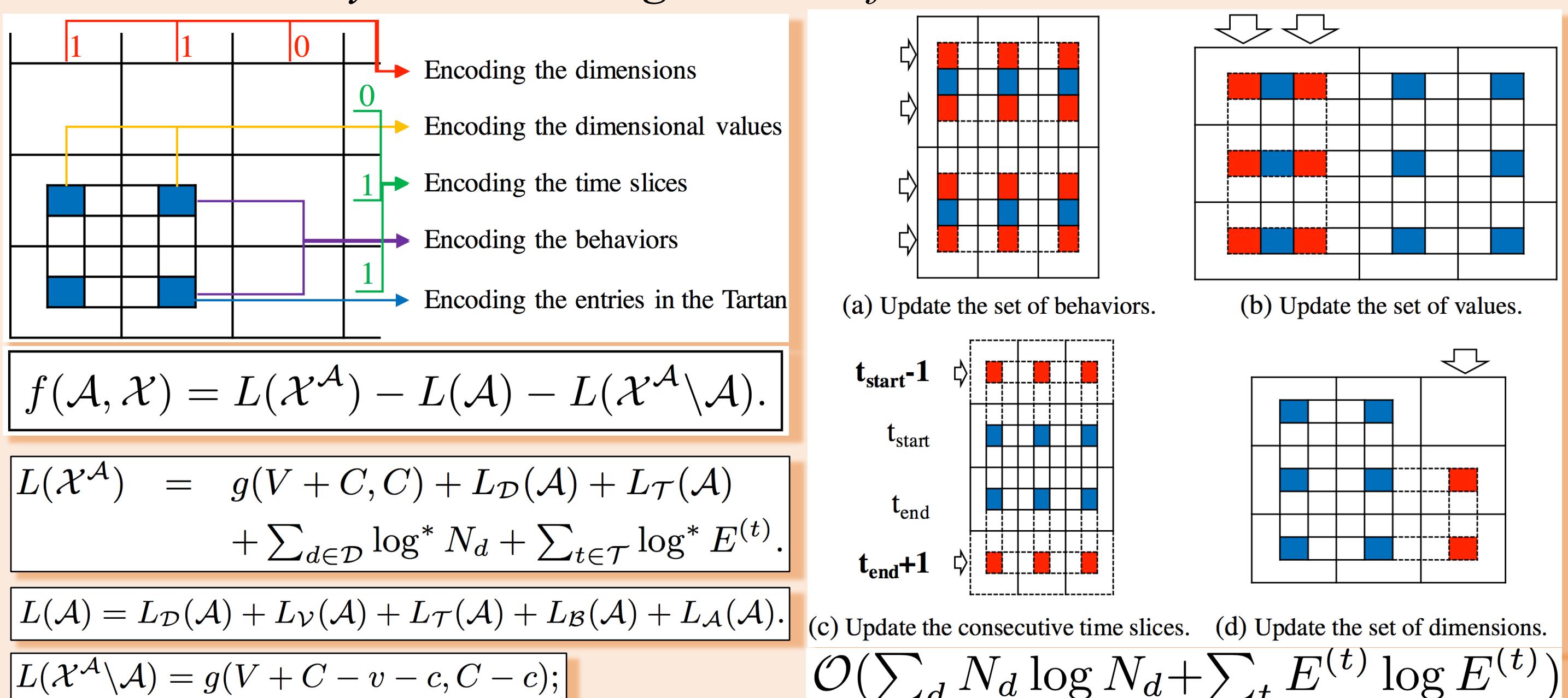
A1: “Two-Level Matrix” for Behavior.

A2: “Tartan” for Behavioral Summary.

Q3: How to catch the Tartans?

## CatchTartan: MDL Principle-driven Algorithm

How many bits encoding a Tartan from the matrix can save?



Contact Meng Jiang (mjiang89@gmail.com): www.meng-jiang.com

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 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.	“my prediction”	user (3,325) phrase 226 hashtag (0) URL (0)	3,397 tweets	Tartan #1: (1 dim) 16:30-17:30
17:00	17:30:51 RT @LMAOTWITPICTS: Make Your Prediction. Retweet For 49ers http://t.co/KKksEist 17:31:01 RT @LMAOTWITPICTS: Make Your Prediction. Retweet For 49ers http://t.co/KKksEist 17:31:16 RT @LMAOTWITPICTS: Make Your Prediction. Retweet For 49ers http://t.co/KKksEist 17:31:19 RT @LMAOTWITPICTS: Make Your Prediction. Retweet For 49ers http://t.co/KKksEist	“make your prediction”	user (196) phrase RT @user URL 1 1	196 tweets	Tartan #2: (3 dims) 17:00-18:00
17:30	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”, “1st Qtr”	user (213) phrase RT @user URL 21 3 (0)	215 tweets	Tartan #3: (2 dims) 18:30-19:30
18:00	20:20:01 RT @ExtraGrumpyCat: No Superbowl halftime show will ever surpass this. http://t.co/0VSy7Cv6 20:20:02 RT @WolfpackAlan: No Superbowl halftime show will ever surpass this. http://t.co/6BilloPXs 20:20:04 RT @ExtraGrumpyCat: No Superbowl halftime show will ever surpass this. http://t.co/0VSy7Cv6 20:20:05 RT @WolfpackAlan: No Superbowl halftime show will ever surpass this. http://t.co/6BilloPXs	halftime show”	user (617) phrase RT @user URL 11 4 4	617 tweets	Tartan #4: (3 dims) 20:00-21:00
18:30	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 20:24:32 (Manhattan, NY) No one can ever top that performance by Beyoncé EVER. #Beyonce #superbowl #halftimestow	“beyonce”, #beyonce, #superbowl, #DestinysChild	location 2 phrase 55 hashtag 17 URL (0)	166 tweets	Tartan #5: (3 dims) 20:00-21:00
19:00	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 69 hashtag 11 URL (0)	653 tweets	Tartan #6: (2 dims) 21:00-22:00
19:30	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 248 hashtag (0) URL (0)	1,950 tweets	Tartan #7: (1 dim) 22:00-23:00
20:00					
20:30					
21:00					
21:30					
22:00					

