

# Modeling of Diffusion Processes on Networks Spatiotemporal Analysis and Agent-based Models

Jay Lee, Xinyue Ye, and Sagar Naresh Jha

August 10, 2015



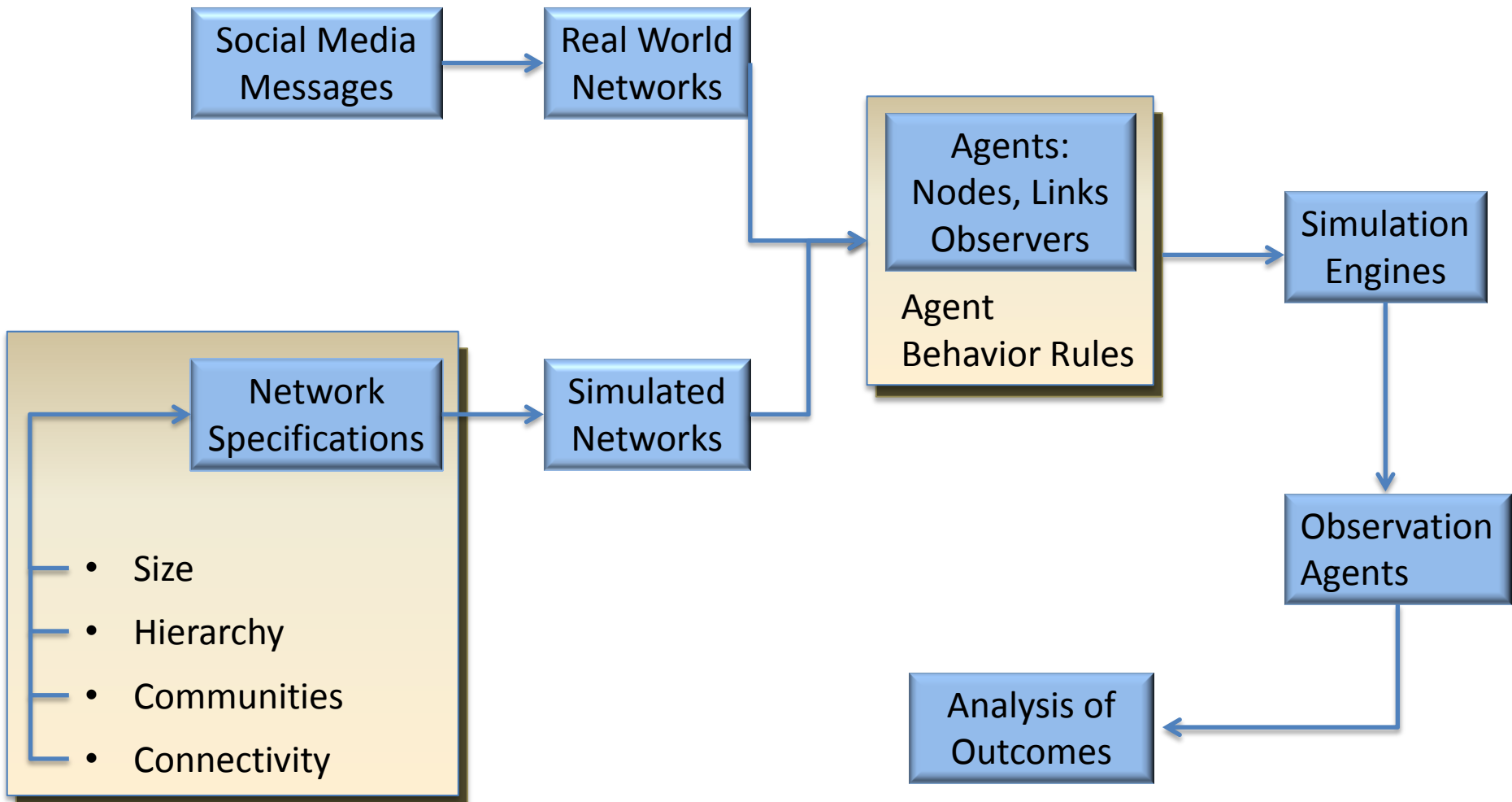
**Spatiotemporal Modeling of Human Dynamics Across Social Media and Social Networks  
Interdisciplinary Behavioral and Social Science Research, National Science Foundation**



## Year 1 Research Objectives

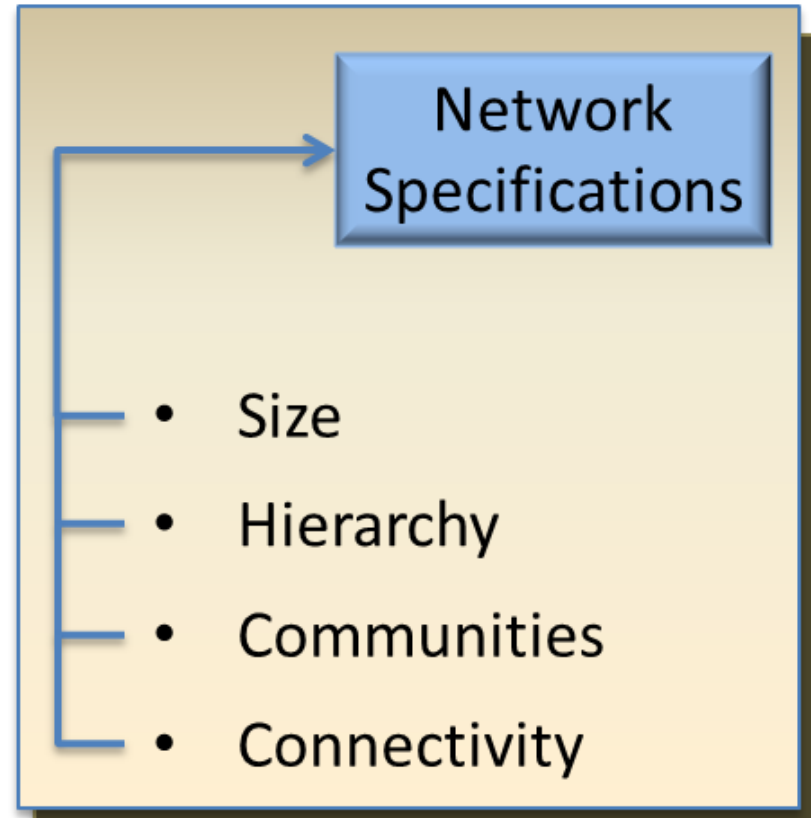
- Lee
  - Investigate and analyze different configurations for meme networks.
  - Categorize network configurations based on tested meme contents.
- Ye
  - Investigate the spatiotemporal effect in the information diffusion process.
  - Characterize the dynamic changes of information landscape through a series of indicators.

# Overall Research Design



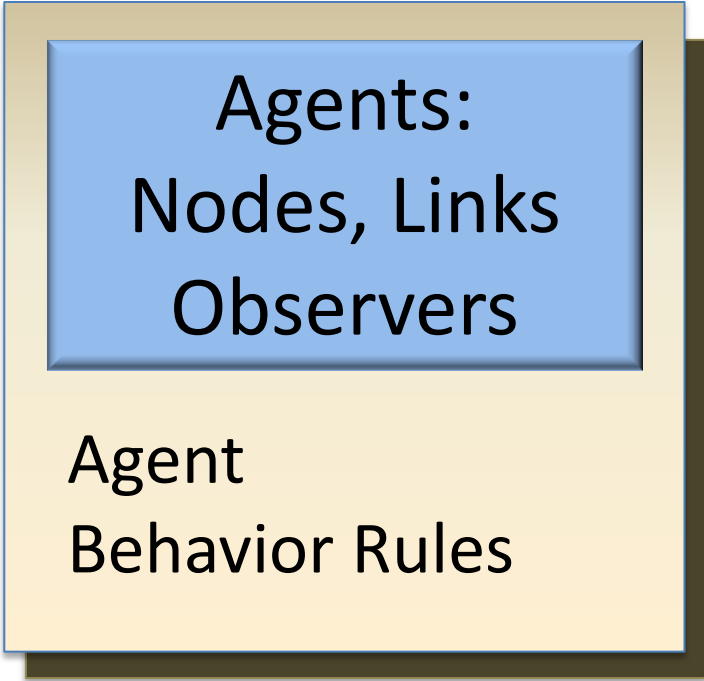
# Network Specifications

- **Size**
  - # nodes
  - # links
- **Hierarchy**
  - # tiers
  - # level 1 opinion leaders
  - # level 2 opinion leaders
  - .....
- **Communities**
  - # communities
- **Connectivity**
  - Connectivity index



# Model Agents

- **Agents**
  - Nodes
    - e.g., Twitter accounts
    - e.g., Opinion leaders
    - e.g., Community interfaces
  - Links
    - e.g., Re-tweet
    - e.g., Passage/Blockage
  - Observer
    - e.g. Statistical summaries



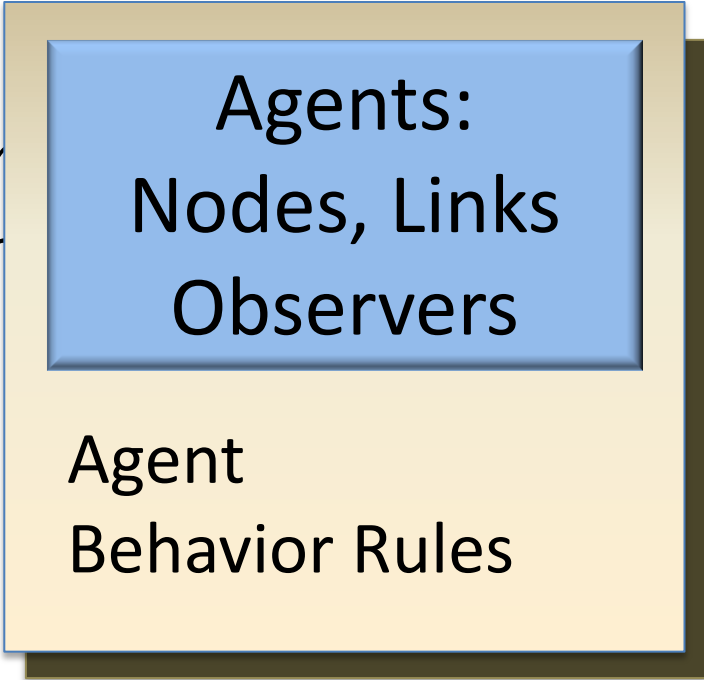
Agents:  
Nodes, Links  
Observers

Agent  
Behavior Rules

# Agent Rules

- **Behavioral rules**

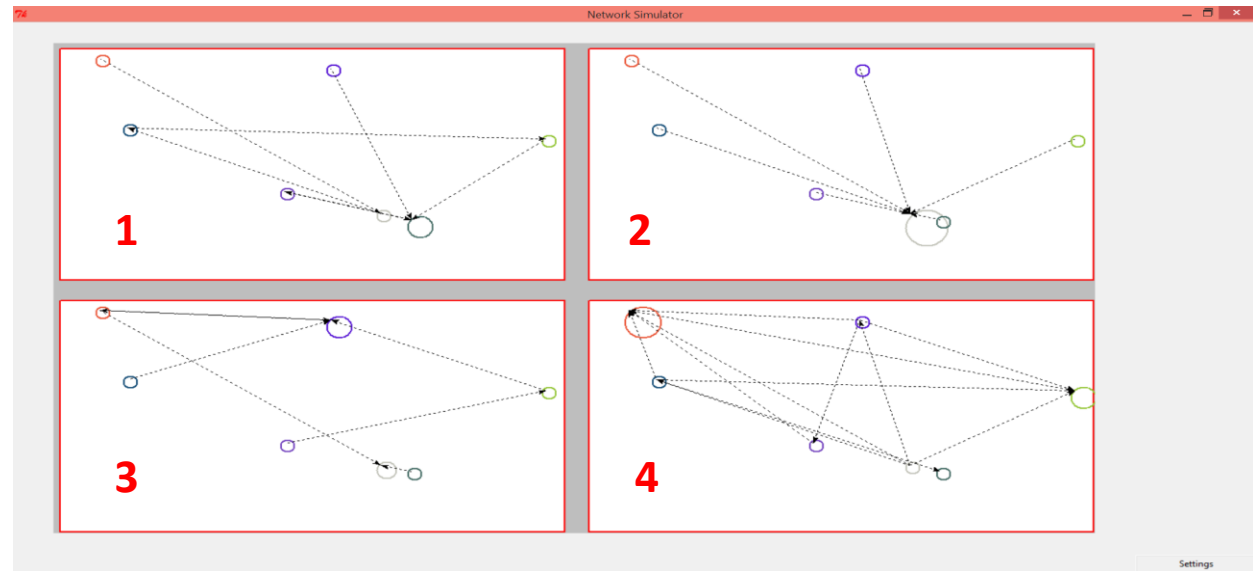
- Tier 1/Tier 2/Tier 3 opinion leader
  - e.g., Probabilities for retweeting
  - e.g., Received messages
- Links
  - e.g., Passage
  - e.g., Blockage
- Communities
  - e.g., Connectivity



Agents:  
Nodes, Links  
Observers

Agent  
Behavior Rules

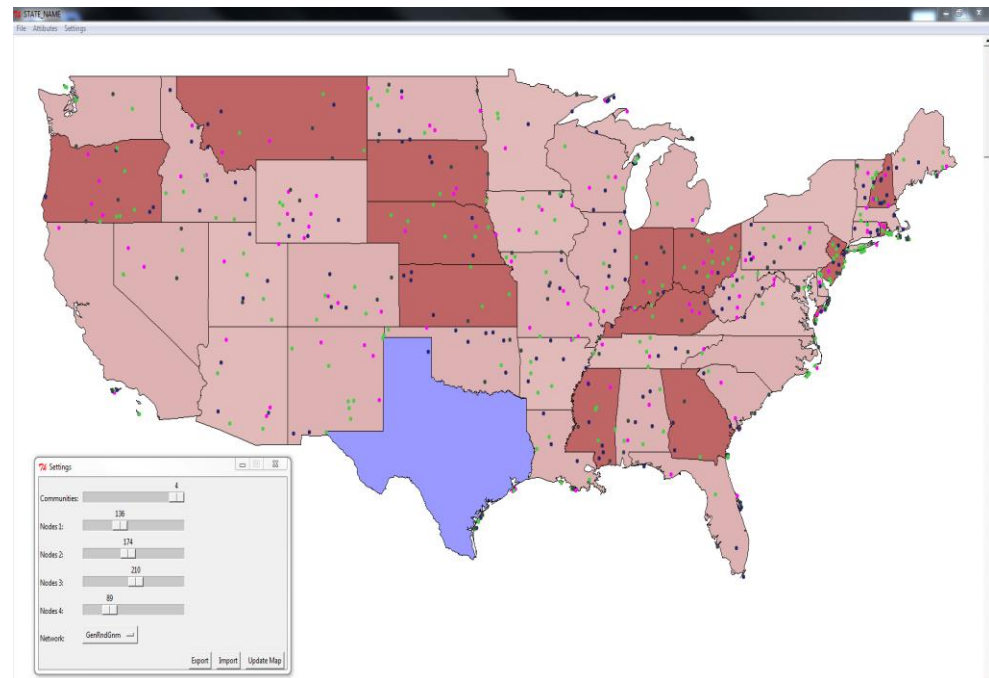
# Network Simulator – V1



- Four examples for different types of networks
  - 1 : Random network
  - 2 : Star topology-based network
  - 3 : Endos – Renyi random network
  - 4 : Directed graph network
  - *Combinations?*

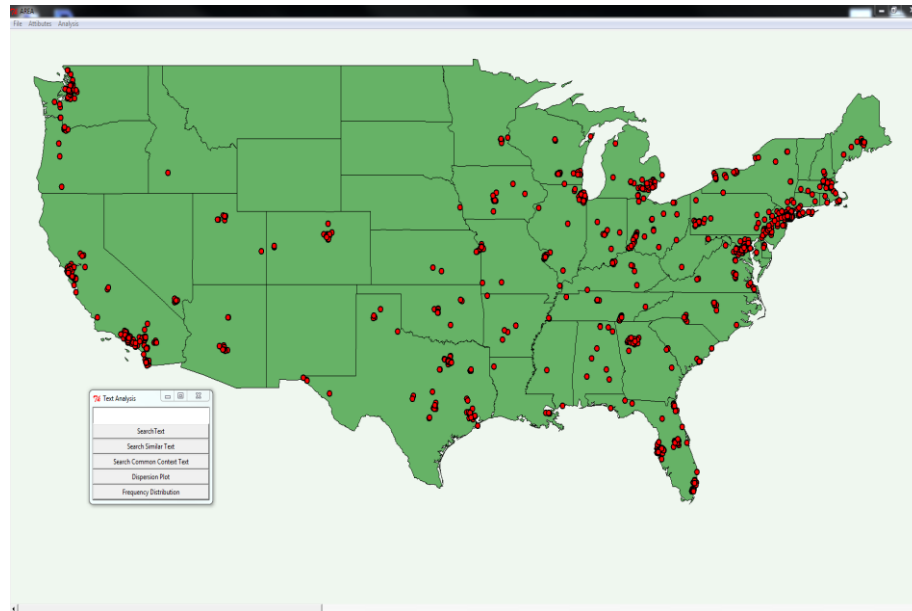
# Network Simulator – V2

- Implemented so far
  - Import and visualize shape files (.shp file) on canvas.
  - Simulate networks on user defined parameters.
  - Import and visualize real world twitter data.
  - Perform text analysis on the twitter data.





# Plot real world twitter data



_keyword	platform	geo_lat	geo_lon	place	country	tweet_in	retweet_c	favourite	user_id	user_scre	user_nam	user_locat	user_geoc	user_follc	user_frier	text
4sq	<a href="f	43.03823	-87.9434	Milwaukee	United Stz	16726631	0	1	2.21E+08	Schroeder	John Schroeder	TRUE	537	642	@fm1021:	
4sq	<a href="f	34.04836	-84.5988	Kennesaw	United States		0	0	13871082	wfowlkes	Will Fowl	Kennesaw	TRUE	1537	1999	Working o
4sq	<a href="f	42.51984	-92.4555	Cedar Fall	United States		0	0	14852278	Naanad	Naanad	Cedar Fall	TRUE	560	730	Time for s
4sq	<a href="f	28.0043	-82.4482	Tampa, FL	United States		0	0	15818129	GaytorKer	Ken Key	ÃœT: 27.9	TRUE	476	884	I'm at Ken
4sq	<a href="f	43.08281	-88.3628	Hartland,	United States		0	0	15674417	charlesrat	charlesrat	Hartland,	TRUE	552	1006	I'm at Hart
4sq	<a href="f	40.26327	-80.1315	Thompson	United States		0	0	2.51E+09	kopitarus	Shelby	The 412 vi	TRUE	158	463	I'm at @Pz
4sq	<a href="f	41.91037	-87.6881	Chicago, IL	United States		0	0	26016096	Chandlers	Chandler	Logan Squ	TRUE	296	406	I'm at @Rz
4sq	<a href="f	33.58447	-111.923	Scottsdale	United States		0	0	15285490	OliviaWilc	Olivia Wil	AZ and CA	TRUE	7820	2209	I'm at Hark
4sq	<a href="f	44.90976	-89.5969	Weston, V	United States		0	0	1.32E+09	cl_hutton	Carrie Hut	Wausau	TRUE	342	829	Monthly G

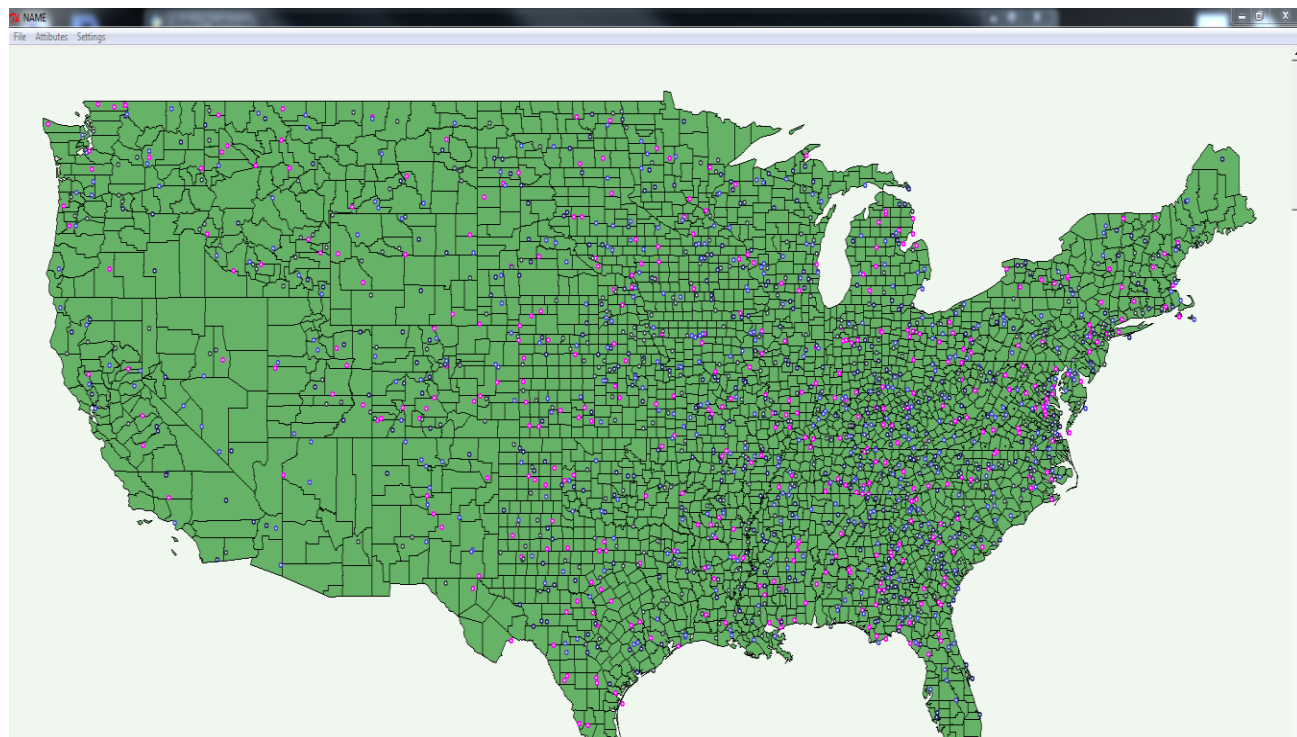
## Classical networks

- Implemented classical networks (**SNAP Networks**)
  - GenRndGnm:
    - Generates an Erdos-Renyi random graph of directed type.
  - GenForestFire:
    - Generates a random Forest Fire, directed graph with given probabilities.
  - GenStar:
    - Generates a graph with star topology. It will have a node connected to all other nodes of the network.
- More implementations are planned

## Graph Generators

- GenFull
- GenCircle
- GenGrid
- GenStar
- GenTree
- GenRndGnm
- GenPrefAttach
- GenGeoPrefAttach
- GenForestFire
- GenSmallWorld
- GenBaraHierar
- GenConfModel
- GenConfModel
- GenCopyModel
- GenDegSeq
- GenRewire
- GenRndDegK
- GenRndPowerLaw
- GenRMat
- GenRMatEpinions

# Simulated Network



Nodes only

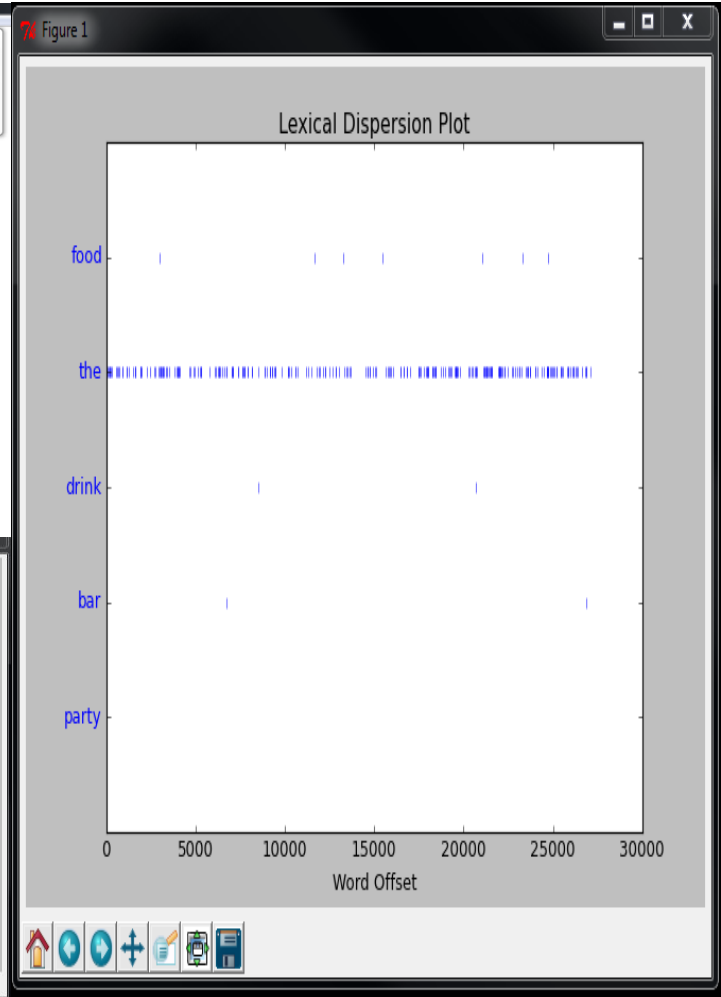
avoids crowded display

Colors represent communities

connectivity depends on connections between communities

# Text Analysis

```
Python 2.7.10 (default, May 23 2015, 09:44:00) [MSC v.1500 64 bit (AMD64)] on win32
Type "copyright", "credits" or "license()" for more information.
>>>
>>>
>>> open shape file!
C:/Users/sjhal/Documents/GitHub/Network_Analysis/Tool_US_Map/Tool/ShapeFiles/states/tweetdata_1.shp
fileLength, number of units: 9130 2270
Search Text
-----
Displaying 16 of 16 matches:
da24U At Sylvia's, the queen of soul food tonight! Corn bread, Cat fish and fr
n Buffalo, NY http://t.co/9W0Qc0lm3s Food Truck Rodeo!!!!!! (@ Southern Twer
tp://t.co/7PUJL1XztG I'm at Jimmy's Food Store in Dallas, TX http://t.co/jzu6
; Family! #Vindaloo (@ Mantra Indian Food in Temecula, CA) http://t.co/UOQk7WY
place brings back memories. awesome food and good times! (@ Blue Moon Cafe in
Night tonight - most amazing Mexican food I've had! @rockinghorsenyc down in..
Love Nino's Mid Town location! Great food and I love their olives. (at @NinosP
sy, TN http://t.co/Yo6n133AKi I'm at Food Court - @smallincolumbia in Columbia,
p://t.co/vpEQz22h3 I'm at Woodman's Food Market - @woodmansFood33 in Carpente
p://t.co/98380JEB3f I'm at Woodman's Food Market - @woodmansFood16 in Madison,
VMYJ78J0 Hanging out w/ my bud. Good food & drinks = good times! (@ Cork F
d & drinks = good times! (@ Cork Food & Drink in Tyler, TX) http://t.c
p://t.co/coUrW2WJ08 I'm at Breakfast food cart (New York, NY) http://t.co/Daaf
//t.co/EZMq0oR36 Out for the second food run of the month. (@ WINGO Foods in
KMDtaQW6TY I'm at Bridgeport Pasty - Food Truck in Chicago, IL http://t.co/9Po
rport) http://t.co/axK9c5eGUC I'm at Food Truck Invasion's Family Night @ Plan
-----
Search Similar Text
-----
vans neighborhood supreme general
-----
Search common context
-----
at_store at_truck breakfast_cart great_and pasty_truck second_run
ny_truck mexican down_at_market good_drinks at_court awesome_and
indian_in_cork_drink soul_corn
-----
Search common context
-----
No common contexts were found
```





## Journal Articles

- Ye, X., Huang, Q., and Li, W. Integrating Big Social Data, Computing, and Modeling for a Synthesized Spatial Social Science, *Cartography and Geographic Information Science* (in press)
- Ye, X., Pomeroy, J., He, C., and Geores, M. The New Data Landscape for Regional and Urban Analysis, *GeoJournal* (in press)
- Ye, X. and Y. Mansury. Behavior Driven Agent Based Models of Spatial Systems, *Annals of Regional Science* (in press)
- Yang, X., Ye, X., Sui, D. We Know Where You Are—In Space and Place: Enriching the Geographical Context of Social Media. *International Journal of Applied Geospatial Research*. 2015.

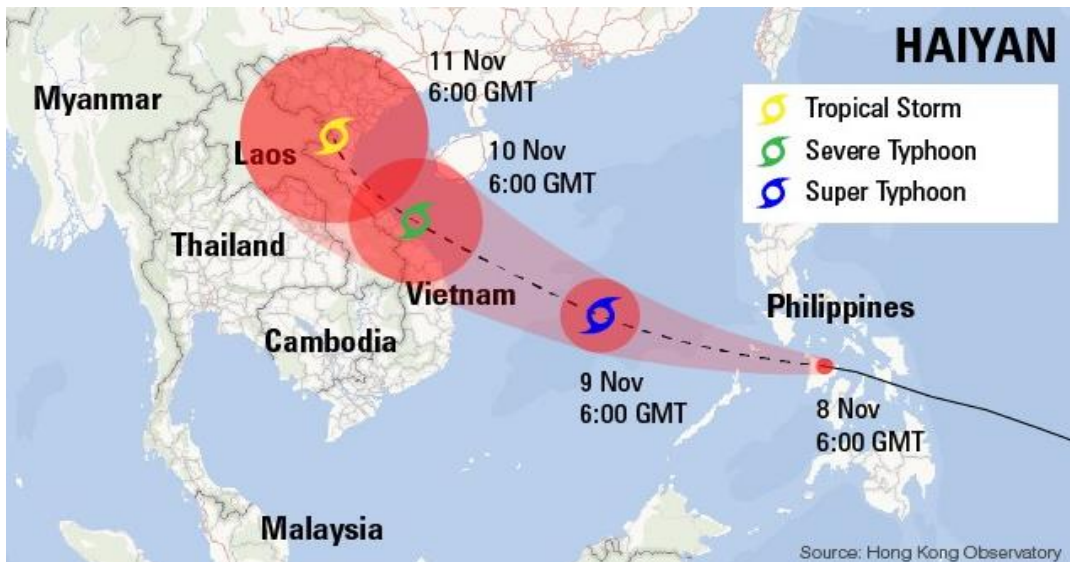


## Book Chapters and Proceeding Papers

- Ye, X. Spatial econometrics, The Wiley-AAG International Encyclopedia of Geography.
- Lai, C.H., Ye, X., She, B. A mixed mode view of the evolutionary dynamics of humanitarian organizing networks: Integration of bona fide networks, media multiplexity, and affordances. Proceeding of the 65th annual conference of International Communication Association, Puerto Rico. 2015.
- Lai, C.H, Ye, X., She, B., & Tao, C.C. A longitudinal investigation of technological affordances for humanitarian organizing. Proceeding of the 65th annual conference of International Communication Association, Puerto Rico. 2015.
- Ye, X., Li, S., Lee, J. Exploring Housing Market Dynamics using Big Data, Proceeding of The Fourth International Workshop on Regional, Urban, and Spatial Economics, Beijing, June, 2015.
- Ye, X., Li, S., Lee, J. Analyzing Ebola Fear in China, 23rd International Conference on Geoinformatics, 2015.

# Ye, X.\* , Lai, S., She, B. (2015) A Space-Time framework of Analyzing Network Structure and Community Evolution: A Case Study of Typhoon Haiyan on Twitter

Typhoon Haiyan, known in the Philippines as **Typhoon Yolanda**, was one of the strongest tropical cyclones ever recorded, which devastated portions of Southeast Asia, particularly the Philippines, on November 8, 2013.<sup>[1]</sup> It is the deadliest Philippine typhoon on record,<sup>[2]</sup> killing at least 6,300 people in that country alone.<sup>[3]</sup> Haiyan is also the strongest storm recorded at landfall, and unofficially the strongest typhoon ever recorded in terms of wind speed.<sup>[4]</sup> As of January 2014, bodies were still being found.<sup>[5]</sup>





- The voluntary and humanitarian organizations involved in Typhoon Haiyan in the Philippines
- The evolution of organized relief efforts and the role of technology and inter-organizational relationships in the process of relief action
- Longitudinal evidence about the process and outcomes of organizational networks



## Mapping Social Connection

- Process raw data
- Eight two-week
- two binary matrices: “mention” network (70 organizations and 15,636 **contacts**) and “shared **hashtag**” network (70 organizations and 5,983 keywords). **Element: contact or hashtag.**
- four square one-mode valued matrices (70X70 and 15,636X15,636; 70X70 and 5,983X5,983)
- In total, 48 matrices were created ( $2*8+4*8 = 48$ ).  
[96 matrices were created ( $4*8+8*8 = 96$ )]



## Type and Geography of 70 organizations

- Seven categories: business, government agency, academic, NPO, NGO, IGO, and coalition.
- Similarity in geographic region (1: Philippines, n=17; 2: Asia and Pacific Region, n=1; 3: Europe, n=16; 4: North America, n=36)

# Data preparation

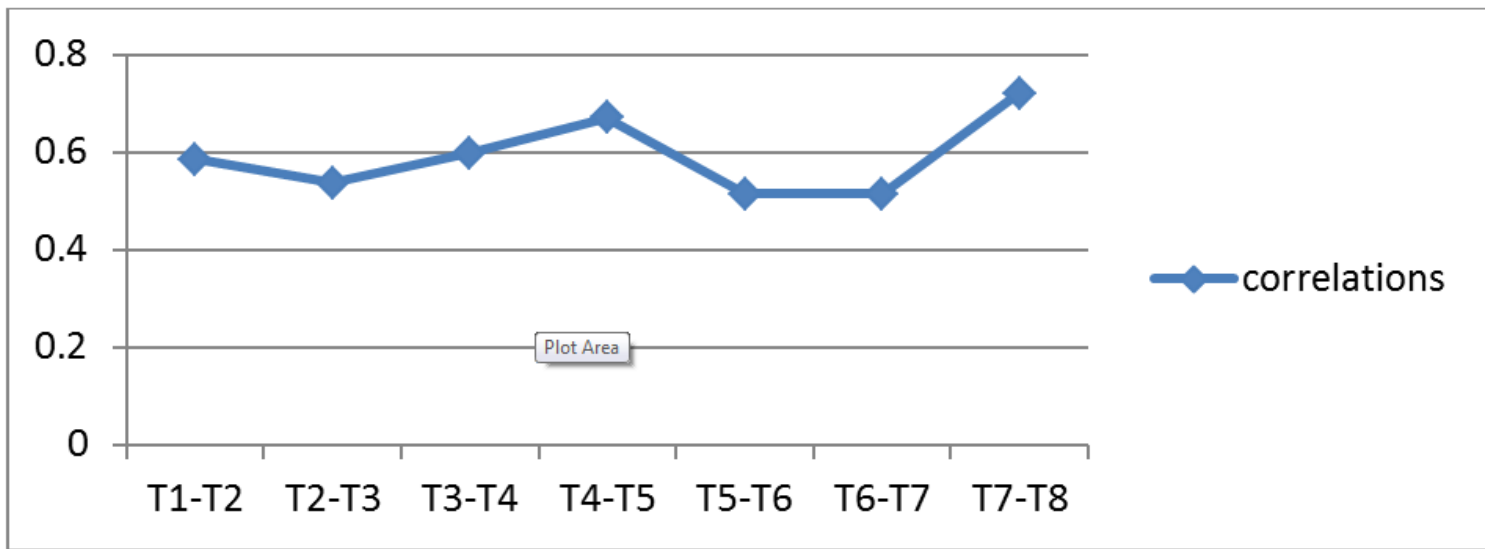
- Data
  - 70 organizations
  - Four-month period: October 12, 2013 to January 31, 2014: starting before the typhoon to three months after
  - 44,280 tweets
- Tweet content:
  - Raw text
  - Date
  - ID
- Example

*“RT @VisitorNews: Make a difference in the lives of those suffering from the effects of #TyphoonHaiyan  
<http://t.co/Wfw8hHctaW> via @Adventistâ”*

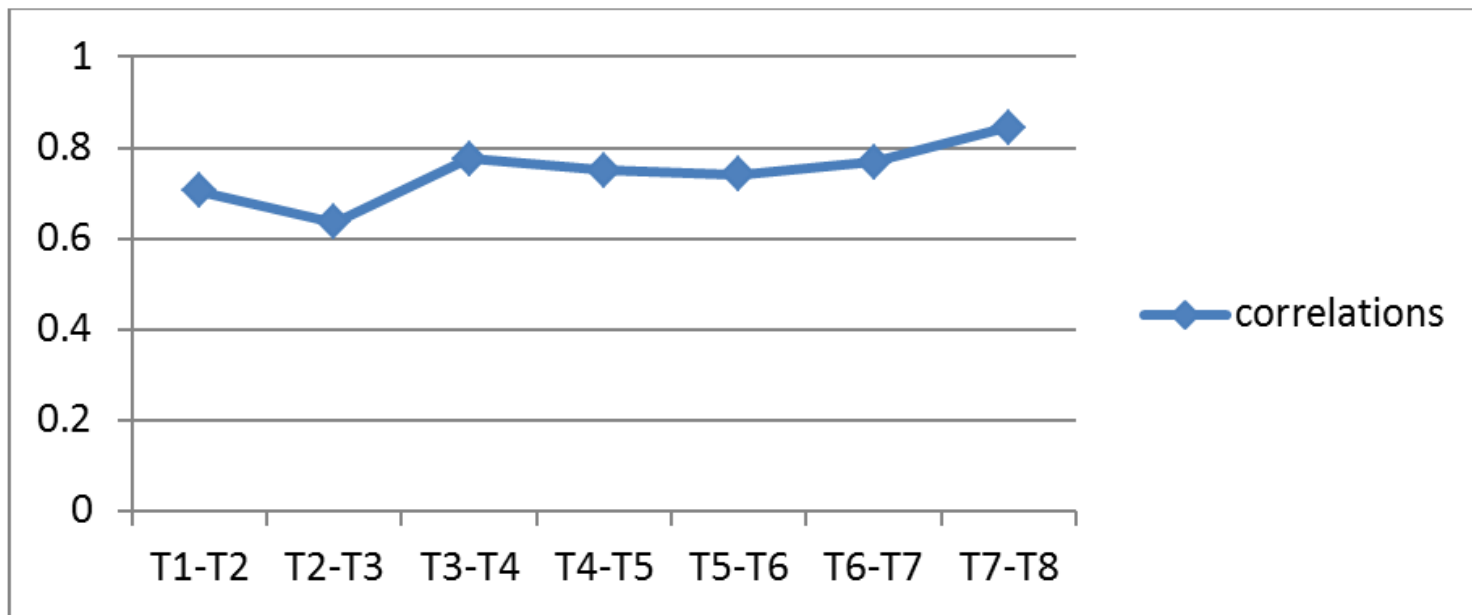
*2013/12/3 20:33:17*

*407970817214255000*

(a) Mentioned contacts (all of the seven correlations were significant at  $p < .001$ )



(b) Shared hashtag contacts (all of the correlations were significant at  $p < .001$ )



# MR-QAP Analysis of T8 (Typhoon Haiyan occurred at T3)

	Model 1 Post-Haiyan Network (mentioned contacts)	Model 2 Post-Haiyan Network (shared hashtag contacts)
Intercept	.000 (-.001) ***	.000 (.004) ***
Organizational Type Similarity	-.015(-.031)	.010 (.048)
Geographic Proximity	-.008(-.016)	-.019 (-.096)
T1	.184 (0.237) ***	.176(.237) ***
T2	-.025 (-.030)	.073(.089) **
T3	.052 (.024) *	-.048 (-.040) *
T4	-.023 (-.020)	.119 (.097) **
T5	.196 (.192) ***	.062 (.068) *
T6	.001 (.002)	.205 (.314) ***
T7	.488 (.383) ***	.412(.437) ***
Adjusted R <sup>2</sup>	.580	.778
Significance level	.000	.000
Number of Observation	4830	4830



## Year 2 Research Objectives

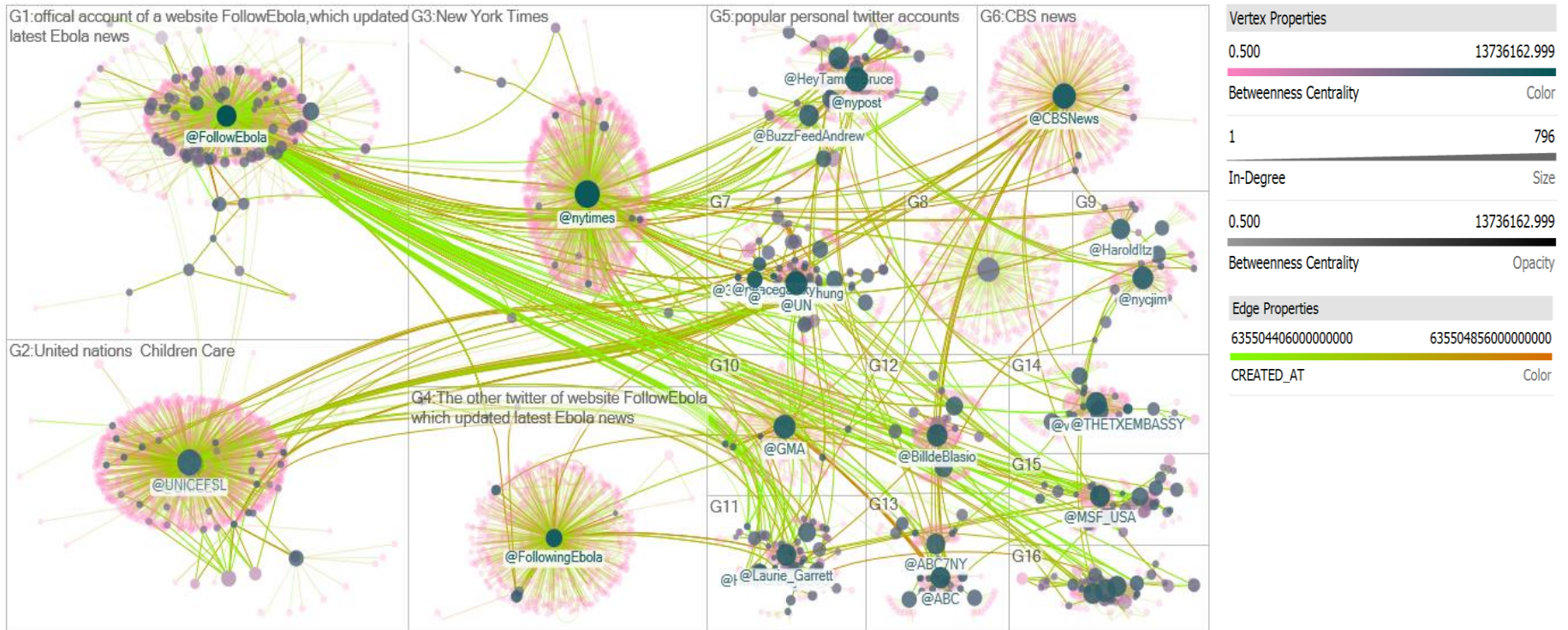
- Lee
  - Development of an agent-based model to model meme diffusion processes.
  - Build a software tool to run the simulation of meme diffusion processes over different networks.
- Ye
  - Apply proposed data model and methods to analyzing the patterns and trends of memes across networks.
  - Develop modeling approaches to examine the driving forces.



# Influence Spread in social media network

- **Haoran Sun** and **Ruoming Jin**
  - August 10, 2015

# Begins with the spread of Ebola news on twitter



Data is provided by HDMA at Santiago State University.



# Propagation of American Express Promotions

The American Express allow their card members sync their twitter accounts with their American Express credit card.

**Amex Offers you easy ways to save on shopping, dining and more when you tweet.**

 Sign in with Twitter

Sign in to Twitter and connect your Card to get offers today!

Once the twitter account @Americanexpress posted a hashtag such as #AmexStaples, then those cardmembers can tweet the exactly same hash tag to get the promotions when they shopped with their Amex-express card.

# The @AmericanExpress posts promotions



American Express @AmericanExpress · Aug 5

Tweet #AmexOrvis, get \$20 back 1x on purchs totaling \$100+ at Orvis w/cnctd Amex Card! RegLtd Exp 9/19/15 Terms:[amex.co/1lbh9dw](https://amex.co/1lbh9dw)

ORVIS

**Orvis - Spend \$100+, Get \$20 Back**

Valid in-store and online. Valid at any participating location in the US. Not valid at outlet locations. Tweet #AmexOrvis to get started!

[sync.americanexpress.com](https://sync.americanexpress.com)

User @3.141592653...  
got the promotion as his  
retweet contains the hash  
tag #AmexOrvis



3.141592653589793238 retweeted



American Express @AmericanExpress · Aug 5

Tweet #AmexOrvis, get \$20 back 1x on purchs totaling \$100+ at Orvis w/cnctd Amex Card! RegLtd Exp 9/19/15 Terms:[amex.co/1lbh9dw](https://amex.co/1lbh9dw)

ORVIS

**Orvis - Spend \$100+, Get \$20 Back**

Valid in-store and online. Valid at any participating location in the US. Not valid at outlet locations. Tweet #AmexOrvis to get started!

[sync.americanexpress.com](https://sync.americanexpress.com)

- Once a user was confirmed to get the promotion, another account, @AmexOffers will mention the card member in a tweet. If it failed, it will tell the reason.



Amex Offers @AmexOffers · Jul 20

@scott\_fia\_li Thx for enrolling in [#AmexBestBuy](#) offer. Spend w/connected Card & receive credit. Terms: [amex.co/1UApDni](https://amex.co/1UApDni)

← ↻ 8 ★ 5 📧 ⋮

[View summary](#)

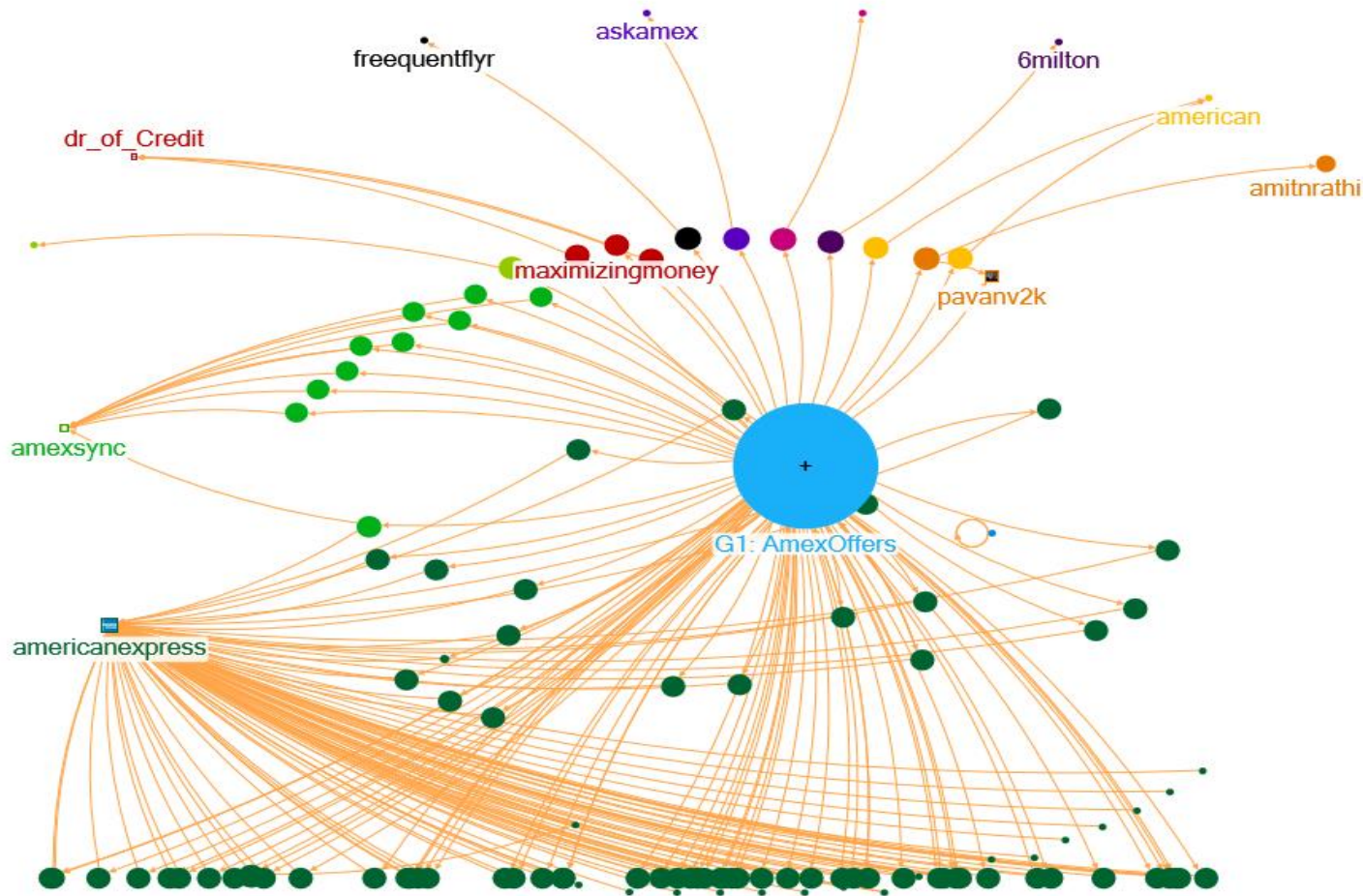


[Amex Offers](#) @AmexOffers · Jul 23

@948\_bc\_bao Sorry, max # of Card Members have enrolled in [#AmexHP](#) offer. Visit [amex.co/faves](https://amex.co/faves) for more offers

← ↻ 1 ★ 2 📧 ⋮

# Visualization of the whole process



## What is behind the process

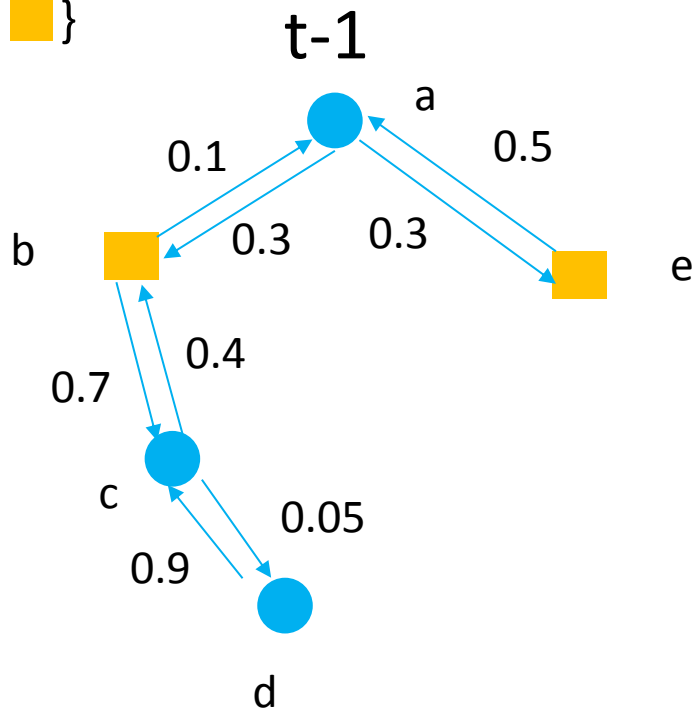
If we consider any user who tweets the hash tag as an “activated” vertex in the social media network, then we can interpret this process as a **stochastic diffusion model**.

A ***stochastic diffusion model*** for a social graph  $G=(V,E)$  specifies the randomized progress of generating active sets  $S_t$  for all  $t \geq 1$  given the initial seed set  $S_0$ . If all  $t \geq 1$ ,  $S_{t-1} \subseteq S_t$ , nodes only go from inactivity to activity,

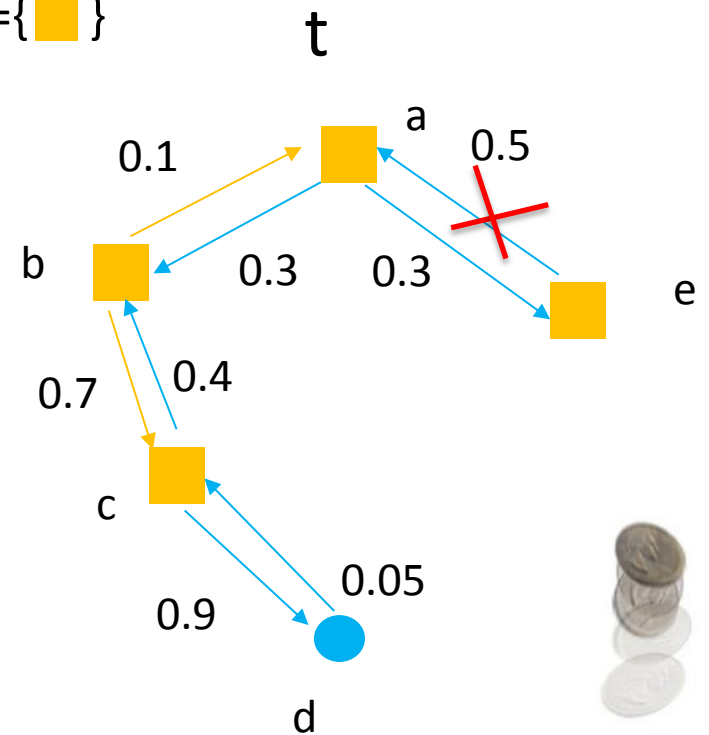
we call it **progressive process**.

# Example—Independent Cascade Model

$S_0 = \{ \text{■} \}$



$S_1 = \{ \text{■} \}$





## What is **influence spread**? Definition

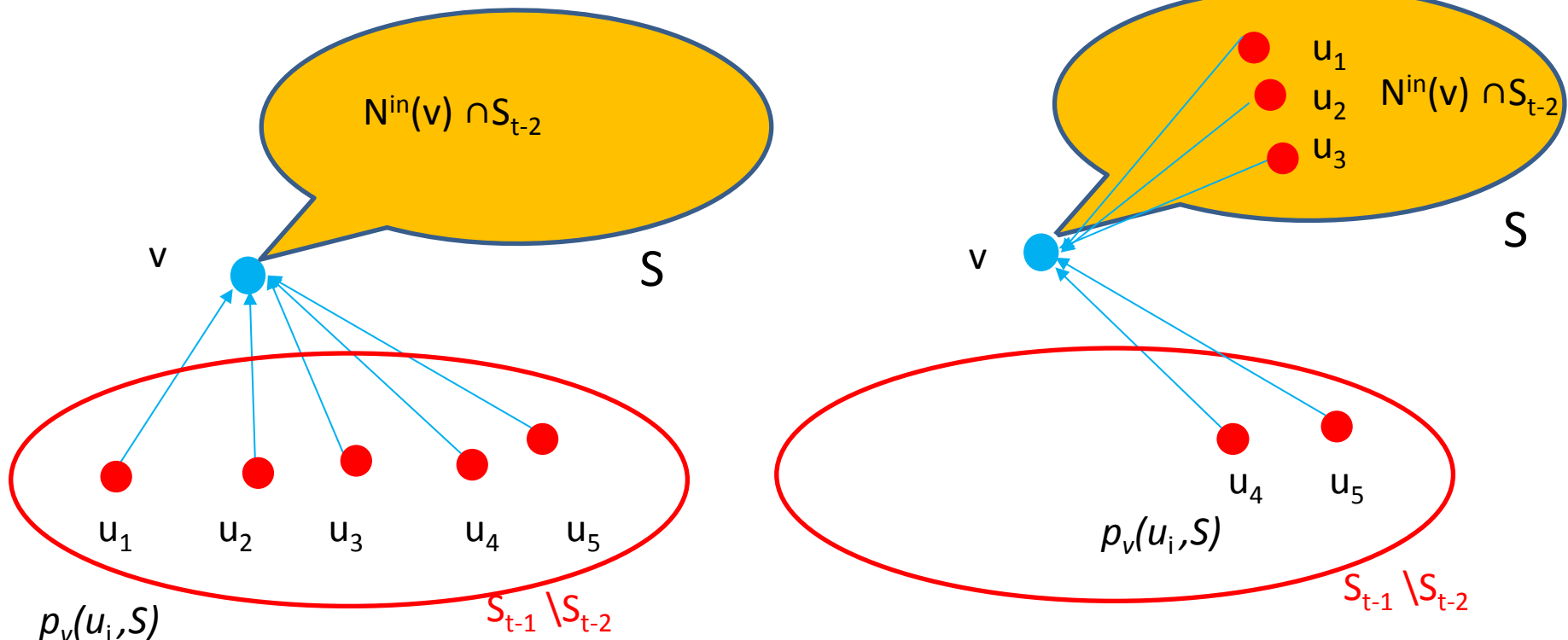
- is the *final stable activated* set of nodes in stochastic process of diffusion model, with seed set  $S_0$
- is the expectation of this set(# of nodes), called influence spread

## To interpret it with General Cascade Model :

If inactive node  $v$  at **time  $t-1$**  has *just activated* in-neighbors sets  $\{u\} = N^{\text{in}}(v) \cap (S_{t-1} \setminus S_{t-2})$ , nodes in  $\{u\}$  will activate  $v$  one by one

If part of the nodes set  $u$ ,  $\{u_1, u_2, \dots, u_{i-1}\}$  failed to activate  $v$ . They will form a set  $S = (N^{\text{in}}(v) \cap S_{t-2}) \cup \{u_1, u_2, \dots, u_{i-1}\}$

The node  $v$  has a probability  $p_v(u_i, S) \in [0, 1]$  to be activated for each in-neighbor  $u_i$







## Why this model

- Once a credit card member add the promotion, he doesn't have to cancel that promotion. This event satisfies the definition of progressive process.
- The promotions appears randomly and face to all the card holders. So every card member have a potential to make one or two purchase.



## To do list

- Build more sophisticated math models to describe the promotion activity online , not just the American Express promotion. And prove its correctness.
- Develop a good algorithm to maximize the influence spread, which means to find a better way of advertisement.
- Define the entropy in the propagation process.
- Incorporate the development into the open source tool in Kent State.



# Thank you for your attention!

- Comments/suggestions are welcome:
  - Jay Lee, [jlee@kent.edu](mailto:jlee@kent.edu)
  - Xinyue Ye, [xye5@kent.edu](mailto:xye5@kent.edu)
  - Sagar Jha, [sagar.jhaa@gmail.com](mailto:sagar.jhaa@gmail.com)
  - Haoran Sun, [hsun9@kent.edu](mailto:hsun9@kent.edu)

