NetdiffuseR

Brief introduction and new capabilities

Aníbal Olivera, George Vega Yon, and Thomas Valente

1. Introduction
    a. Network diffusion of Innovations
    b. netdiffuseR

2. Simulation of diffusion networks
    a. rdiffnet

4. Analyzing rdiffnet
    a. Single-behavior inputs
    b. Multi-behavior inputs

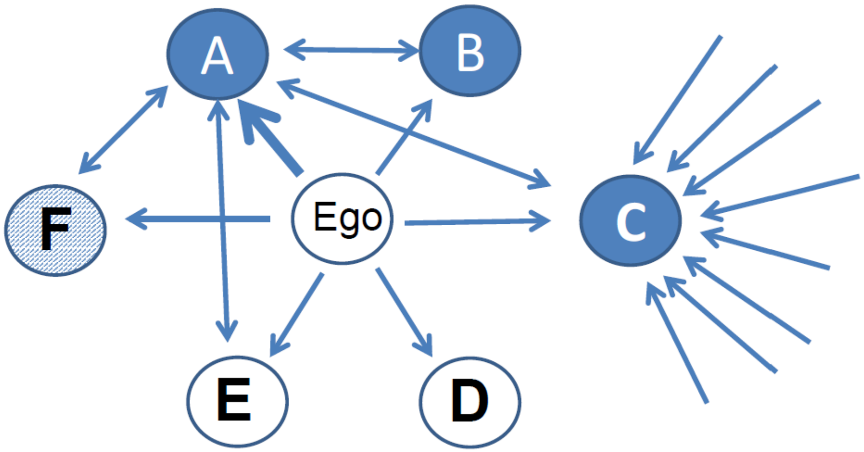
5. Some examples

Introduction

Network diffusion of Innovations

Diffusion networks

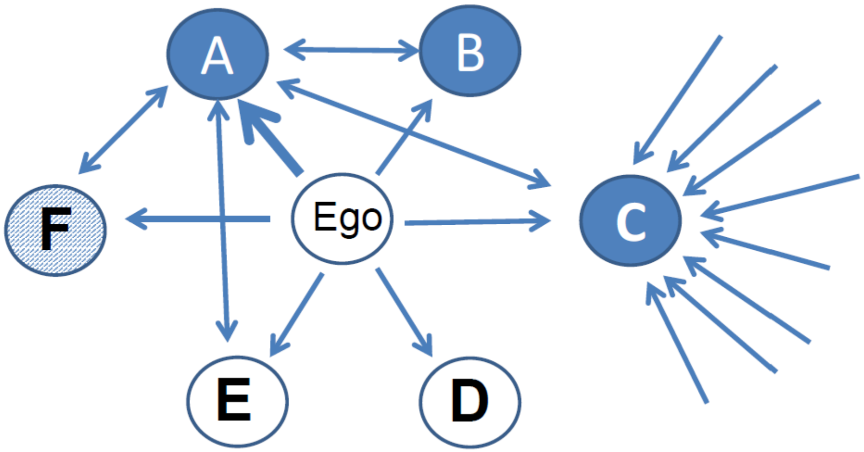
- Considerable research has been conducted to show that certain behaviors are **contagious** and spread through person-to-person contact [1].



Diffusion networks

- Considerable research has been conducted to show that certain behaviors are **contagious** and spread through person-to-person contact [1].

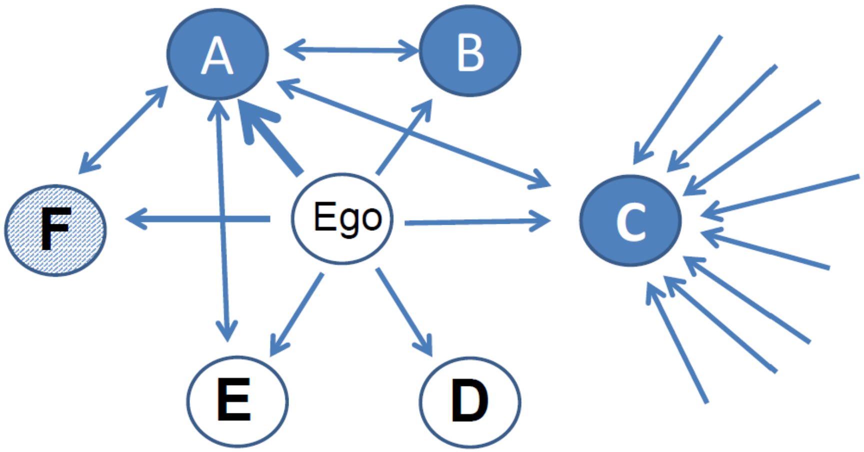
- Healthy habits [2,3],



Diffusion networks

- Considerable research has been conducted to show that certain behaviors are **contagious** and spread through person-to-person contact [1].

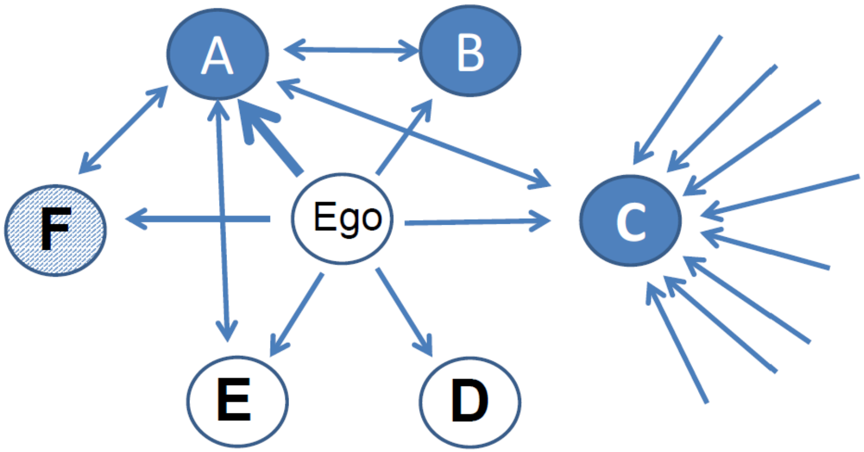
- Healthy habits [2,3],
- Adoption of social norms [4,5],



Diffusion networks

- Considerable research has been conducted to show that certain behaviors are **contagious** and spread through person-to-person contact [1].

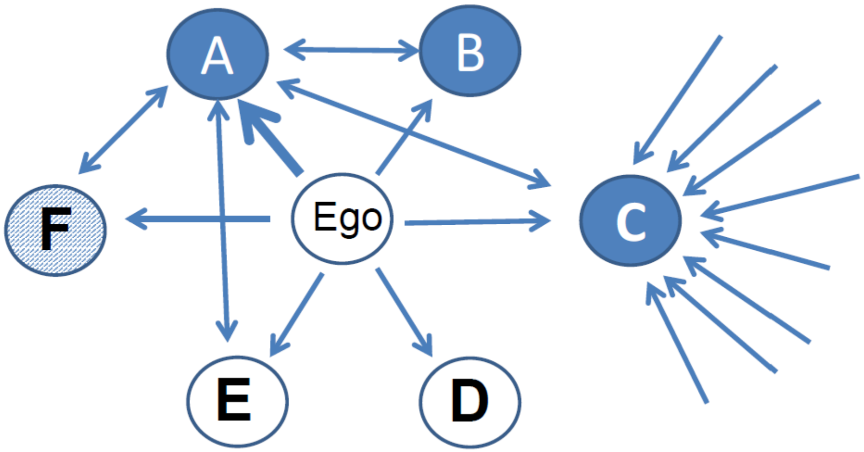
- Healthy habits [2,3],
- Adoption of social norms [4,5],
- Political behavior [6,7],



Diffusion networks

- Considerable research has been conducted to show that certain behaviors are **contagious** and spread through person-to-person contact [1].

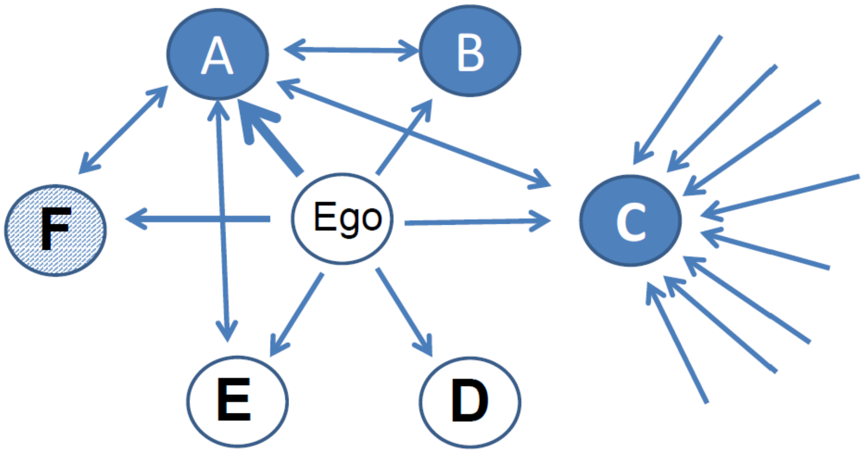
- Healthy habits [2,3],
- Adoption of social norms [4,5],
- Political behavior [6,7],
- Technological innovations [8,9,10],



Diffusion networks

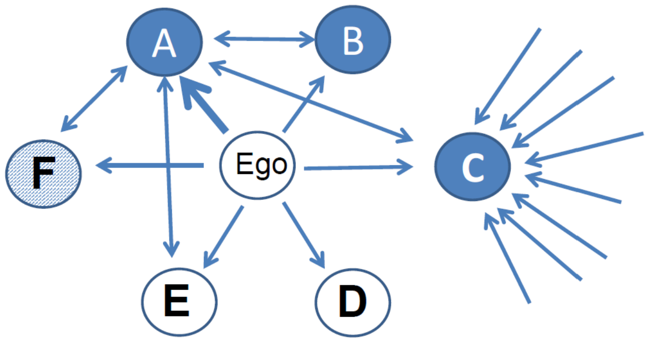
- Considerable research has been conducted to show that certain behaviors are **contagious** and spread through person-to-person contact [1].

- Healthy habits [2,3],
- Adoption of social norms [4,5],
- Political behavior [6,7],
- Technological innovations [8,9,10],
- How urban rumors spread [11],
- etc…



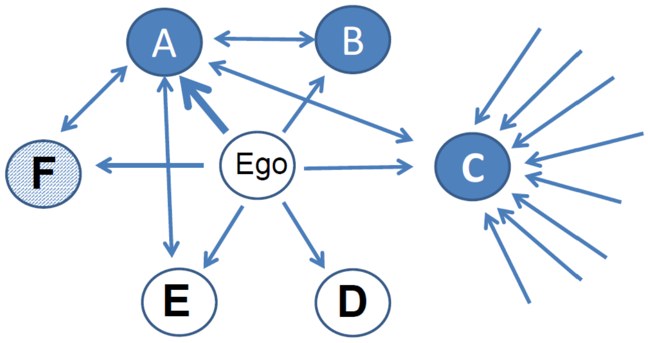
Diffusion networks

- **Contagion:** the transmission of a disease by direct or indirect contact and also an **influence** that spreads rapidly.



Diffusion networks

- **Contagion:** the transmission of a disease by direct or indirect contact and also an **influence** that spreads rapidly.

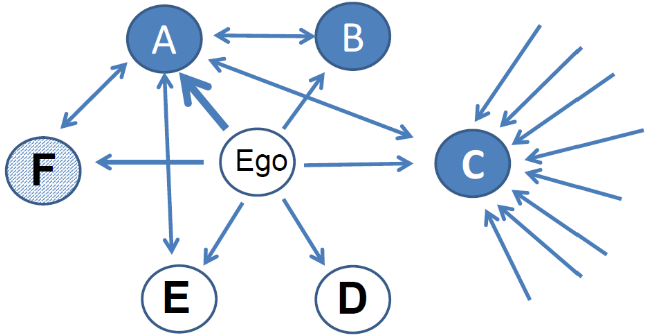


Types of **external** influences:

- Marketing campaign,
- Delivery of flyers to promote a new product
- Etc..

Diffusion networks

- **Contagion:** the transmission of a disease by direct or indirect contact and also an **influence** that spreads rapidly.



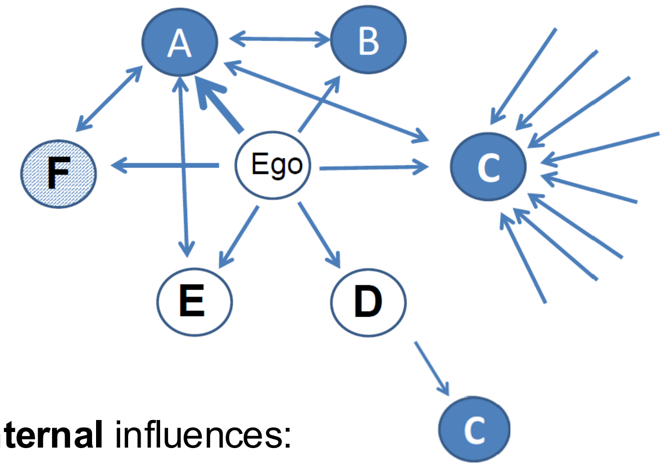
Types of **external** influences:

- Marketing campaign,
- Delivery of flyers to promote a new product
- Etc..

Types of **internal** influences:
- Direct Exposure

Diffusion networks

- **Contagion:** the transmission of a disease by direct or indirect contact and also an **influence** that spreads rapidly.



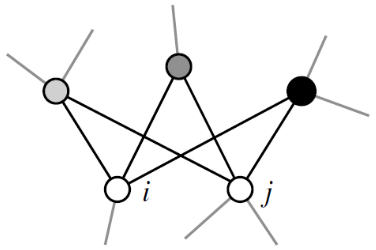
Types of **external** influences:

- Marketing campaign,
- Delivery of flyers to promote a new product
- Etc..

Types of **internal** influences:
- Direct Exposure
- Indirect Exposure

Diffusion networks

- **Contagion:** the transmission of a disease by direct or indirect contact and also an **influence** that spreads rapidly.



(a) Structural equivalence

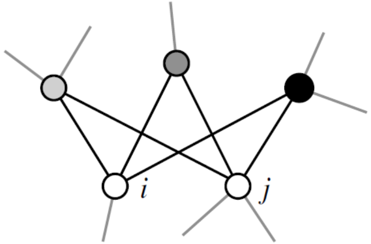
Types of **external** influences:

- Marketing campaign,
- Delivery of flyers to promote a new product
- Etc..

Types of **internal** influences:
- Direct Exposure
- Indirect Exposure
- Structural Equivalence

Diffusion networks

- **Contagion:** the transmission of a disease by direct or indirect contact and also an **influence** that spreads rapidly.



(a) Structural equivalence

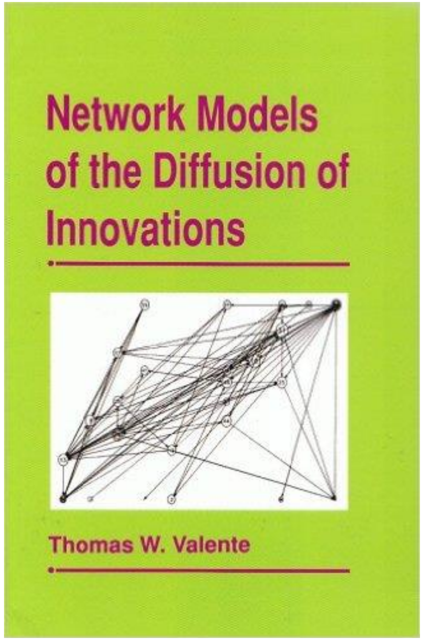
Types of **external** influences:

- Marketing campaign,
- Delivery of flyers to promote a new product
- Etc..

Types of **internal** influences:

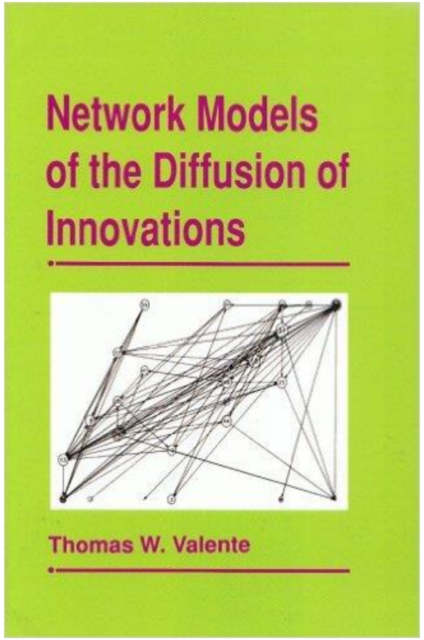
- Direct Exposure
- Indirect Exposure
- Structural Equivalence
- Exposure weighted by tie strength
- Etc..

Diffusion networks



- While a lot of factors have been shown to influence diffusion (Spatial, Economic, Cultural, Biological, etc.), **Social Networks** is a **prominent one**.

Diffusion networks

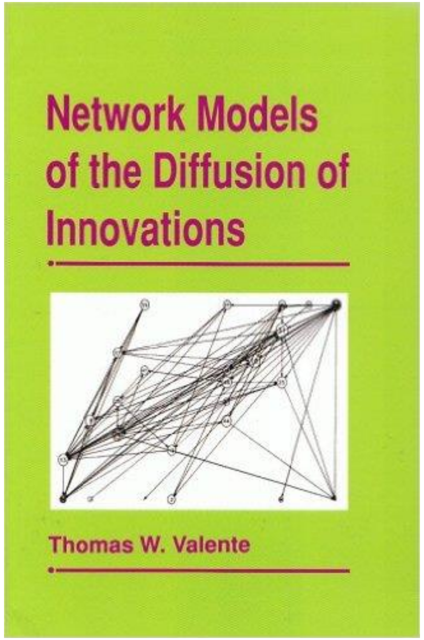


- While a lot of factors have been shown to influence diffusion (Spatial, Economic, Cultural, Biological, etc.), **Social Networks** is a **prominent one**.

- There are many **components** in the diffusion network model including [12]:

  - thresholds of adoption

Diffusion networks

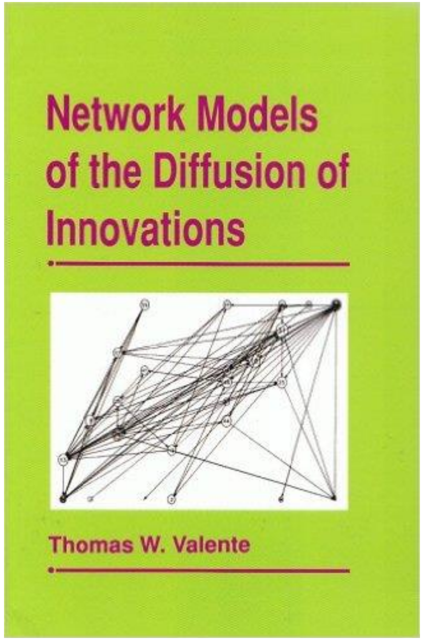


- While a lot of factors have been shown to influence diffusion (Spatial, Economic, Cultural, Biological, etc.), **Social Networks** is a **prominent one**.

- There are many **components** in the diffusion network model including [12]:

  - thresholds of adoption
  - infectiousness → How much *i*'s adoption affects her alters

Diffusion networks

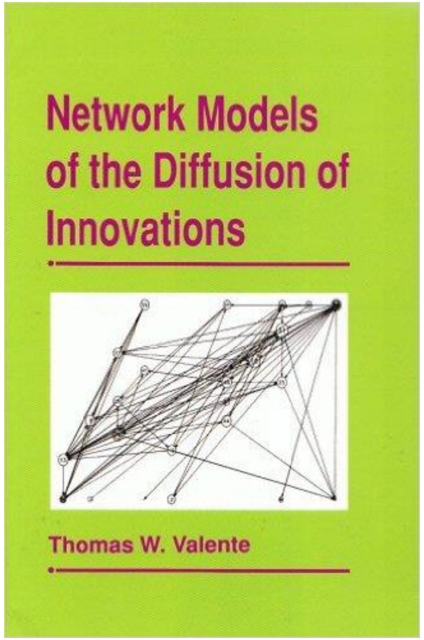


- While a lot of factors have been shown to influence diffusion (Spatial, Economic, Cultural, Biological, etc.), **Social Networks** is a **prominent one**.

- There are many **components** in the diffusion network model including [12]:

  - thresholds of adoption
  - infectiousness → How much *i*'s adoption affects her alters
  - susceptibility → How much *i*'s alters' adoption affects her

Diffusion networks

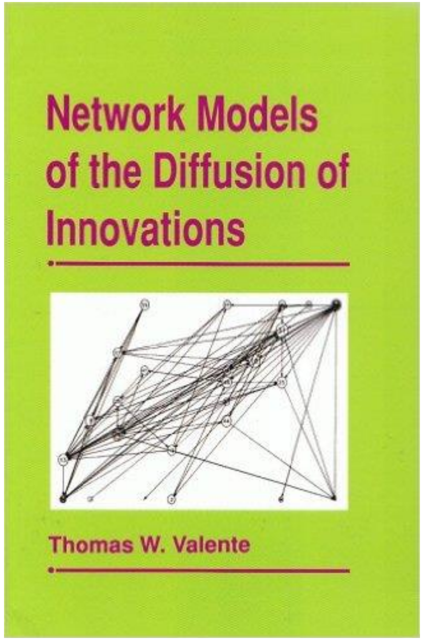


- While a lot of factors have been shown to influence diffusion (Spatial, Economic, Cultural, Biological, etc.), **Social Networks** is a **prominent one**.

- There are many **components** in the diffusion network model including [12]:

  - thresholds of adoption
  - infectiousness → How much *i*'s adoption affects her alters
  - susceptibility → How much *i*'s alters' adoption affects her
  - hazard rates, → likelihood nodes will adopt at that time *t*

Diffusion networks



- While a lot of factors have been shown to influence diffusion (Spatial, Economic, Cultural, Biological, etc.), **Social Networks** is a **prominent one**.

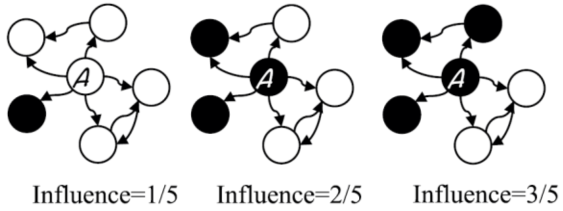
- There are many **components** in the diffusion network model including [12]:

  - thresholds of adoption
  - infectiousness → How much *i*'s adoption affects her alters
  - susceptibility → How much *i*'s alters' adoption affects her
  - hazard rates, → likelihood nodes will adopt at that time *t*
  - diffusion rates (bass model),
  - etc..

Thresholds

- One of the cannonical concepts is the threshold of adoption τ_i :

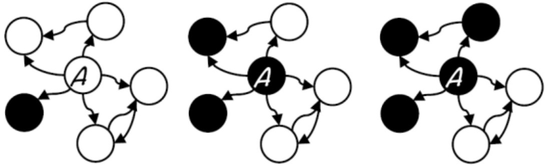
"Required proportion or number of neighbors that leads you to adopt a particular behavior or innovation"



Thresholds

- One of the cannonical concepts is the threshold of adoption τ_i :

“Required proportion or number of neighbors that leads you to adopt a particular behavior or innovation”



Influence=1/5 Influence=2/5 Influence=3/5

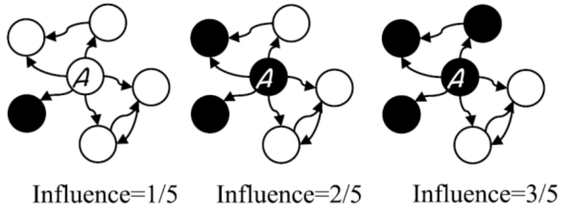
- To calculate the exposition:

$$E_i \equiv \frac{\sum_{j \neq i} \mathbf{X}_{ij} a_j}{\sum_{j \neq i} \mathbf{X}_{ij}}$$

Thresholds

- One of the cannonical concepts is the threshold of adoption τ_i :

"Required proportion or number of neighbors that leads you to adopt a particular behavior or innovation"



- To calculate the exposition:

$$E_i \equiv \frac{\sum_{j \neq i} \mathbf{X}_{ij} a_j}{\sum_{j \neq i} \mathbf{X}_{ij}}$$

- So we can check if the node *i* adopts or not:

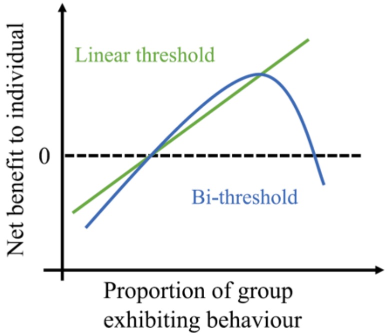
$$a_i = \begin{cases} 1 & \text{if } \tau_i \leq E_i \\ 0 & \text{Otherwise} \end{cases}$$

Adoption and disadoption

- In marketing, there is a high interest in simulating the **competition** between innovations.

Adoption and disadoption

- In marketing, there is a high interest in simulating the **competition** between innovations.
- Also, high interest in testing theoretical models for **disadoption** [13,14]

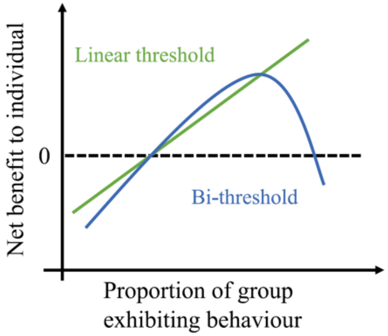


$$P(adopt_i) = \mathrm{logit}^{-1}(\beta_{adopt} exposure_i + \beta^{adopt} x)$$

$$P(disadopt_i) = \mathrm{logit}^{-1}(\beta_{disadopt} exposure_i + \beta^{disadopt} x)$$

Adoption and disadoption

- In marketing, there is a high interest in simulating the **competition** between innovations.
- Also, high interest in testing theoretical models for **disadoption** [13,14]



$$P(adopt_i) = \text{logit}^{-1}(\beta_{adopt} exposure_i + \beta^{adopt} x)$$

$$P(disadopt_i) = \text{logit}^{-1}(\beta_{disadopt} exposure_i + \beta^{disadopt} x)$$

- So, we need a way to simulate the diffusion of **multiple behaviors**.

netdiffuseR

netdiffuseR

netdiffuseR is an R package that:

- Is designed for **Simulating**, Analyzing and Visualizing network diffusion data (in general).

netdiffuseR

netdiffuseR is an R package that:

- Is designed for **Simulating**, Analyzing and Visualizing network diffusion data (in general).

- Depends on some pretty popular packages:
  - RcppArmadillo: So it’s **fast**,
  - Matrix: So it’s **big**,
  - statnet and igraph: So it’s **not from scratch**

netdiffuseR

netdiffuseR is an R package that:

- Is designed for **Simulating**, Analyzing and Visualizing network diffusion data (in general).
- Depends on some pretty popular packages:
  - RcppArmadillo: So it's **fast**,
  - Matrix: So it's **big**,
  - statnet and igraph: So it's **not from scratch**
- Can handle big graphs, e.g., an adjacency matrix with more than 4 billion elements.

netdiffuseR

netdiffuseR is an R package that:

- Is designed for **Simulating**, Analyzing and Visualizing network diffusion data (in general).

- Depends on some pretty popular packages:
  - RcppArmadillo: So it's **fast**,
  - Matrix: So it's **big**,
  - statnet and igraph: So it's **not from scratch**

- Can handle big graphs, e.g., an adjacency matrix with more than 4 billion elements.

- +6,000 downloads since its first version, 2016.

netdiffuseR

netdiffuseR is an R package that:

- Is designed for **Simulating**, Analyzing and Visualizing network diffusion data (in general).

- Depends on some pretty popular packages:
  - RcppArmadillo: So it's **fast**,
  - Matrix: So it's **big**,
  - statnet and igraph: So it's **not from scratch**

- Can handle big graphs, e.g., an adjacency matrix with more than 4 billion elements.

- +6,000 downloads since its first version, 2016.

- A lot of features to make it easy to work with several kinds of networks, so has **high compatibility** with others net packages.

Simulation of diffusion networks

netdiffuseR

Let's focus on **simulating** data

→ rdiffnet()

`rdiffnet` *Random diffnet network*

Description

Simulates a diffusion network by creating a random dynamic network and adoption threshold levels.

Usage

```
rdiffnet_multiple(R, statistic, ..., ncpus = 1L, cl = NULL)

rdiffnet(
  n,
  t,
  seed.nodes = "random",
  seed.p.adopt = 0.05,
  seed.graph = "scale-free",
  rgraph.args = list(),
  rewire = TRUE,
  rewire.args = list(),
  threshold.dist = runif(n),
  exposure.args = list(),
  name = "A diffusion network",
  behavior = "Random contagion",
  stop.no.diff = TRUE
)
```

netdiffuseR

Let's focus on **simulating** data

→ rdiffnet()

- The simulation algorithm is as follows:
  - If required, a baseline graph is created,
  - Set of initial adopters and threshold distribution are established,
  - The set of t networks is created (if required), and
  - Simulation starts at t=2, assigning adopters based on exposures and thresholds:
    - For each *i* ∈ N, if its exposure at t−1 is greater than its threshold, then adopts, otherwise continue without change.
    - next *i*

netdiffuseR

For example:

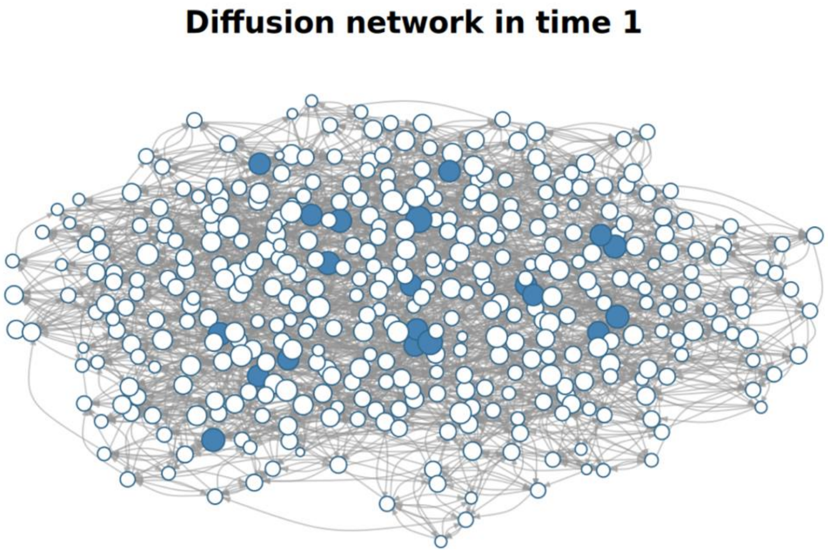
```r
set.seed(12315)
x <- rdiffnet(
  400, t = 6, rgraph.args = list(k=6, p=.3),
  seed.graph = "small-world",
  seed.nodes = "central", rewire = FALSE, threshold.dist = 1/4
  )
```

- 400 nodes
- 6 time steps
- ‘Small-world’ network
- (k=6) Each node is initially connected to 6 neighbors
- (p=.3) probability of rewiring
- Seed nodes selected as those with higher degree centrality
- Static network
- Threshold uniform to all nodes

netdiffuseR

For example:

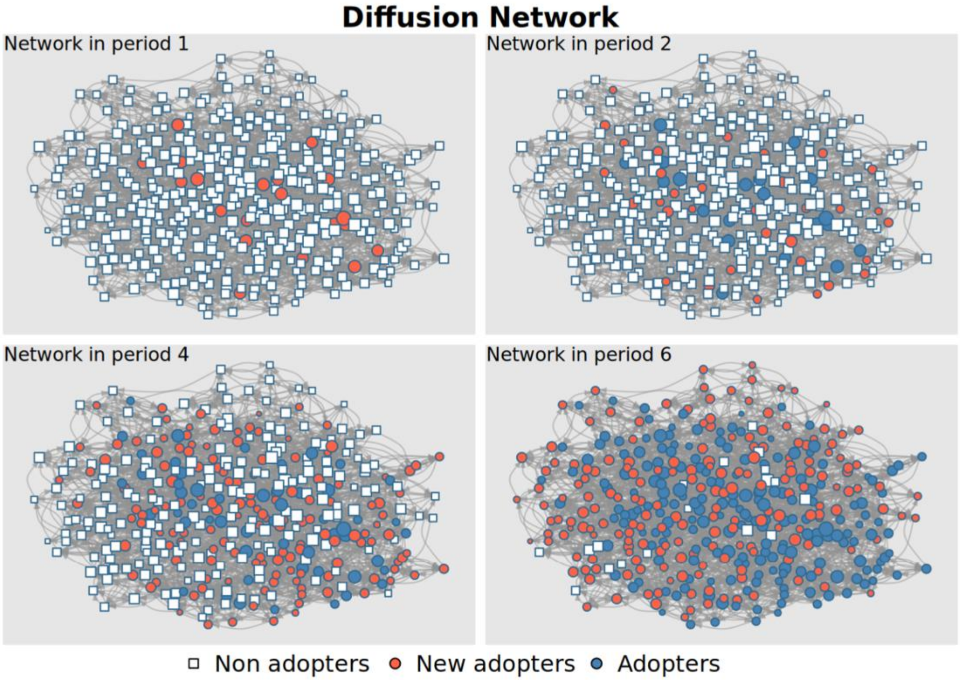
```
plot(x)
```



netdiffuseR

For example:

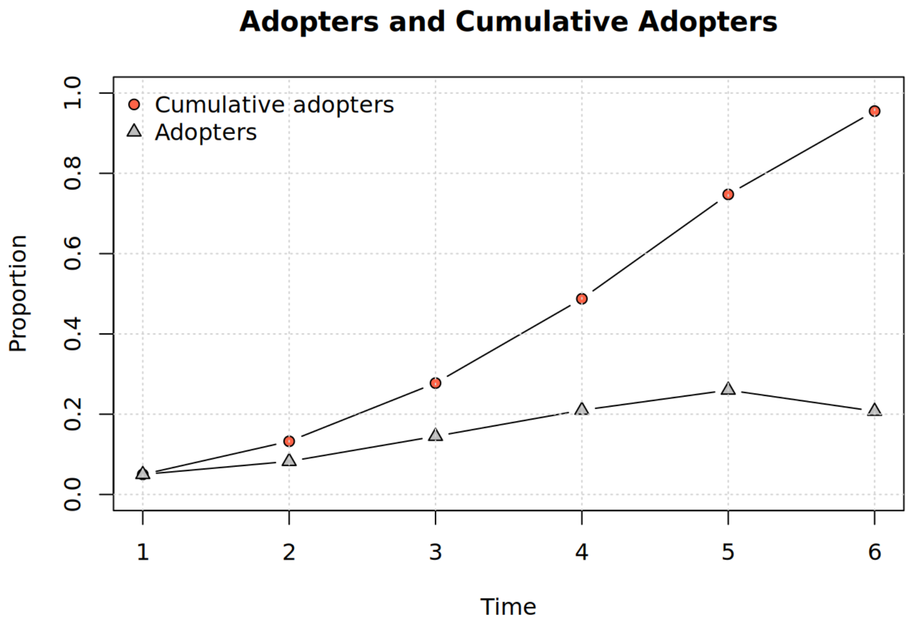
```
plot_diffnet(x)
```



netdiffuseR

For example:

```
plot_adopters(x)
```



Analyzing rdiffnet

Single-behavior rdiffnet

rdiffnet(graph, seed.p.nodes, seed.nodes, behavior, threshold.dist)

Step 0.0: seeting *n* and *t* if not provided

- this depends on graph

Step 1.0: seeting seed nodes

- this depends on seed.p.nodes and seed.nodes

Step 2.0: seeting threshold for each node

⮡ `rdiffnet_make_threshold( threshold.dist, n )`

Step 3.0: running the simulation

⮡ `exposure( graph, exposure.args = list(...) )`

⮡ allow us to compute time of adoption (toa)

Step 4.0: creating diffnet object

⮡ `new_diffnet( graph, ..., toa, ... )`

Single-behavior inputs

		single
seed.p.adopt	numeric	0.1 c(0.1)
	character	✗
	list	✗
seed.nodes	numeric	c(2,4,6)
	character	“random” c(“random”)
	list	✗
behavior	numeric	✗
	character	“tabacco” c(“tabacco”)
	list	✗
threshold	numeric	0.33 rep(0.33, 100)
	function	function() runif(1)
	matrix	✗
	list	✗

Multiple-behavior rdiffnet

rdiffnet(graph, seed.p.nodes, seed.nodes, behavior, threshold.dist)

Step 0.0: seeting *n* and *t* if not provided

- this depends on graph

Step 1.0: validate the arguments

```
rdiffnet_validate_args(seed.p.adopt, seed.nodes, behavior)
```

- this allow us to seeting seed nodes

Step 2.0: seeting threshold for each node

```
rdiffnet_make_threshold( threshold.dist, n, num_of_behaviors)
```

Step 3.0: running the simulation

```
exposure( graph, exposure.args = list(...) )
```

```
disadopt( expo, cumadopt, t)
```

- to compute time of adoption (toa) for each behavior

Step 4.0: creating diffnet object

```
new_diffnet( graph, ..., toa, ... )
```

Multiple-behavior rdiffnet

Step 3.0: running the simulation

exposure(graph, exposure.args = list(...))

disadopt(expo, cumadopt, t)

- to compute time of adoption (toa) for each behavior

Step 4.0: creating diffnet object

new_diffnet(graph, ..., toa, ...)

Step 5.0: splitting behaviors (optional)

split_behaviors(rdiffnet_multiple_obj)

Multiple-behavior inputs

		single		multiple
seed.p.adopt	numeric	0.1 c(0.1)	numeric	
	character		character	
	list		list	list(0.1, 0.05)
seed.nodes	numeric	c(2,4,6)	numeric	c(2,4,6)
	character	“random” c(“random”)	character	“random” c(“random”) c(“random”, “central”)
	list		list	list(“random”, “central”) list(c(1,3,5), c(2,4,6))

Multiple-behavior inputs

		single	multiple
behavior	numeric		
	character	"tabacco" c("tabacco")	"tabacco" c("tabacco") c("tabacco", "alcohol")
	list		list("tabacco", "alcohol")
threshold	numeric	0.33 rep(0.33, 100)	0.33 rep(0.33, 100)
	function	function() runif(1)	function() runif(1)
	matrix		matrix(runif(100), n_nodes, n_behavior)
	list		list(0.33, 0.66) list(runif(100), runif(100)) list(function() runif(1), function() runif(1))

Some examples

To Rstudio

Thank you!

Reference

[1] Guilbeault, D., Becker, J., Centola, D. (2018): Complex Contagions: A Decade in Review. Complex Spreading Phenomena in Social Systems. Computational Social Sciences. Springer.
[2] Christakis, N., Fowler J. (2008): The collective dynamics of smoking in a large social network. N Engl J Med 358(21): 2249–2258
[3] Zhang, J., Brackbill, D., Yang, S., Becker, J., Herbert, N., Centola, D. (2016): Support or competition? How online social networks increase physical activity: a randomized controlled trial. Prev Med Rep 4:453–458
[4] Campbell, E., Salathe, M. (2013): Complex social contagion makes networks more vulnerable to disease outbreaks. Sci Rep 3:1905
[5] Salathe, M., Bonhoeffer, S. (2008): The effect of opinion clustering on disease outbreaks. J R Soc Interface 5(29):1505–1508
[6] Barash, V., Kelly, J. (2012): Salience vs committment: dynamics of political hashtags in Russian Twitter. Berkman Center for Internet and Society, Research Publication No. 2012–9.
[7] Fink, C., Schmidt, A., Barash, V., Cameron, C., Macy, M. (2016): Complex contagions and the diffusion of popular Twitter hashtags in Nigeria. Soc Networks 6: 1

Reference

[8] Karsai, M., Iniguez, G., Kaski, K., Kertesz, J. (2014): Complex contagion process in spreading of online innovation. J R Soc Interface 11(101): 20140694

[9] Ugander, J., Backstrom, L., Marlow, C., Kleinberg, J. (2012): Structural diversity in social contagion. PNAS 109(16): 5962–5966

[10] Oster, E., Thornton, R. (2012): Determinants of technology adoption: peer effects in menstrual cup take-up. J Eur Econ Assoc 10(6):1263–1293[11] Akerlof, G.A. (1997): Social Distance and Social Decisions. Econometrica 65(5), 1005–1027.

[11] Mønsted, B., Sapiezynski, P., Ferrara, E., Lehmann, S. (2017): Evidence of complex contagion of information in social media: an experiment using Twitter bots. PLoS One 12(9): e0184148.

[12] Valente, T.W. (1995) Networks Models of the Diffusion of Innovations. Hampton Press, Inc., Cresskill, NJ, 171.

[13] Alipour, F., Dokshin, F., Enough but not too many: A bi-threshold model for behavioral diffusion, PNAS Nexus, Volume 3, Issue 10, October 2024, pg 428.

[14] Donald R. Lehmann & Jeffrey R. Parker, 2017. "Disadoption," AMS Review, Springer; Academy of Marketing Science, vol. 7(1), pages 36-51, June.