Impact of Clustering on Epidemics in Random Networks

Joint work with Marc Lelarge

INRIA-ENS

8 March 2012

Outline

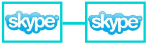
Outline

Game-theoretic contagion model on a given graph $G = (V, E)$, with parameter $q \in (0, 1/2)$:

Two possible choices : ($\leftrightarrow$ susceptible) or ($\leftrightarrow$ infected)

Initially : all use , except one who uses

Possible switch $\rightarrow$, but no switch $\not\rightarrow$

Situation	Payoff (for both users)
	q
	$1 - q > q$
	0

Total payoff
= sum of payoffs from all your neighbors

Switch from to $\Leftrightarrow \frac{|\text{Neighbors using Skype}|}{|\text{Neighbors}|} > q.$

Infinite deterministic graph $G = (V, E)$

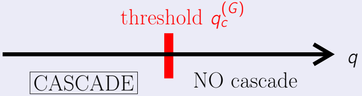
Parameter q varies :

$$\begin{array}{ccc} q \text{ small} & \Rightarrow & \boxed{\text{CASCADE}} \\ q \text{ higher} & \Rightarrow & \text{NO cascade} \end{array}$$

More precisely :

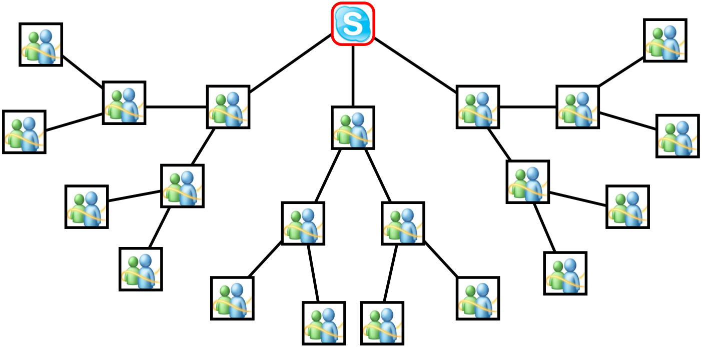
$q_1 \geq q_2$, cascade for $q_1 \Rightarrow$ cascade for q_2

Contagion threshold $q_c^{(G)} := \sup \left\{ q \,\middle|\, \text{CASCADE in } G \text{ for parameter } q \right\}$



Switch from to $\Leftrightarrow \frac{|\text{Neighbors using Skype}|}{|\text{Neighbors}|} > q$

Example : G = d-regular tree



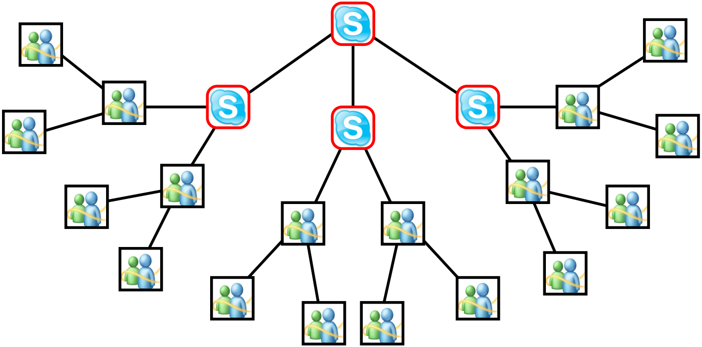
$q \geq 1/d \Rightarrow$ NO cascade

$q < 1/d \Rightarrow$ CASCADE

$$\boxed{\Rightarrow q_c^{(G)} = 1/d}$$

Switch from to $\Leftrightarrow \frac{|\text{Neighbors using Skype}|}{|\text{Neighbors}|} > q$

Example : $G = d$-regular tree



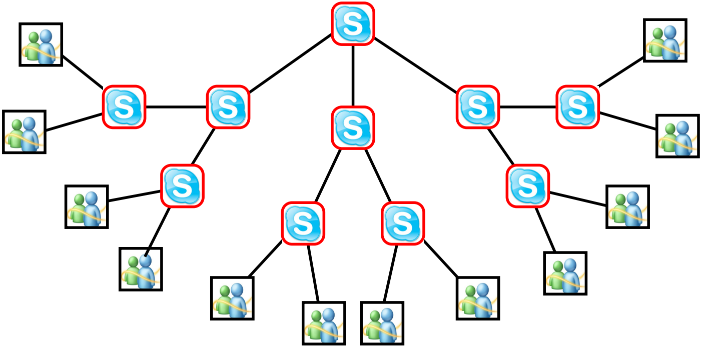
$q \geq 1/d \quad \Rightarrow \quad$ NO cascade

$q < 1/d \quad \Rightarrow \quad$ CASCADE

$$\boxed{\Rightarrow \quad q_c^{(G)} = 1/d}$$

Switch from to $\Leftrightarrow \dfrac{|\text{Neighbors using Skype}|}{|\text{Neighbors}|} > q$

Example : G = d-regular tree



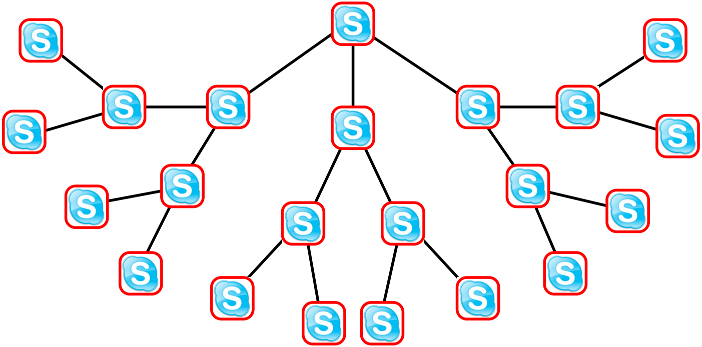
$q \geq 1/d \quad \Rightarrow \quad$ NO cascade

$q < 1/d \quad \Rightarrow \quad$ CASCADE

$$\boxed{\Rightarrow \quad q_c^{(G)} = 1/d}$$

Switch from [MSN] to [Skype] $\Leftrightarrow \dfrac{|\text{Neighbors using Skype}|}{|\text{Neighbors}|} > q$

Example : G = d-regular tree



$q \geq 1/d \quad \Rightarrow \quad$ NO cascade

$q < 1/d \quad \Rightarrow \quad$ CASCADE

$\boxed{\Rightarrow \; q_c^{(G)} = 1/d}$

Define a model of finite *random* graphs (whose size tends to infinity)

- having (asymptotically) the observed properties :
  - *scale-free* networks ⇔ power law degree distribution
    i.e. there exists $\tau > 0$ such that, for all $k \geq 0$, $p_k \propto k^{-\tau}$
    (small number of nodes having a large number of edges)
  - networks with *clustering*
    ("The friends of my friends are my friends", Newman, '03)
- tractable

Epidemic models on finite random graphs :

Final nb of infected nodes negligeable or not / population size ?

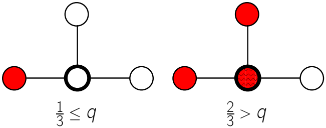
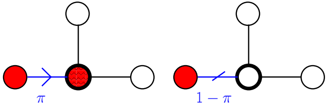
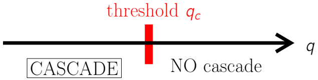
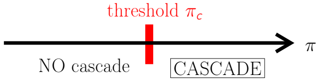
G_n = random graph of size n

S_n = final size of the epidemic in G_n

$$\begin{array}{rcl} \text{CASCADE} & \text{if} & S_n \underset{n\to\infty}{=} \Theta_p(n), \\ \text{NO cascade} & \text{if} & S_n \underset{n\to\infty}{=} o_p(n). \end{array}$$

Epidemic models on finite random graphs :

Final nb of infected nodes negligeable or not / population size ?

	CONTAGION MODEL	DIFFUSION MODEL
Ref.	Morris, Watts	Bond percolation
Para- -meter	A vertex is infected $\Leftrightarrow$ fraction of infected neighbors $> q$ 	π = probability that an edge transmits the epidemic 
Thm		

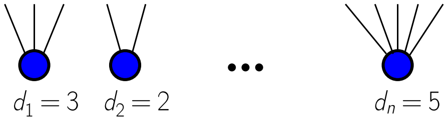
Effect of clustering on these thresholds and on the cascade size

Outline

(i) Start from a uniform graph with given vertex degrees
(ii) Add clustering

(i) Original graph (with given vertex degrees) :

- $n \in \mathbb{N}$, $\boldsymbol{d} = (d_i)_1^n$ sequence of non-negative integers ($s.t.$ $\exists$ a graph with n vertices and degree sequence $\boldsymbol{d}$).
- $G(n, \boldsymbol{d})$ = uniform random graph (among the graphs with n vertices and degree sequence $\boldsymbol{d}$).



Ref. : (Lelarge,'11) for the study of contagion and diffusion models on graphs with given vertex degrees

(i) Start from a uniform graph with given vertex degrees
(ii) Add clustering

(i) Original graph (with given vertex degrees) :

- $n \in \mathbb{N}$, $\boldsymbol{d} = (d_i)_1^n$ sequence of non-negative integers ($s.t.$ $\exists$ a graph with n vertices and degree sequence $\boldsymbol{d}$).
- $G(n, \boldsymbol{d})$ = uniform random graph (among the graphs with n vertices and degree sequence $\boldsymbol{d}$).

Condition : Assume there exists a probability distribution $\boldsymbol{p} = (p_r)_{r=0}^{\infty}$ such that :

(i) $\sharp\{i : d_i = r\}/n \to p_r$ as $n \to \infty$, for every $r \geq 0$
(ii) $\lambda := \sum_r rp_r \in (0; \infty)$
(iii) $\sum_i d_i^3 = O(n)$

Ref. : (Lelarge,'11) for the study of contagion and diffusion models on graphs with given vertex degrees

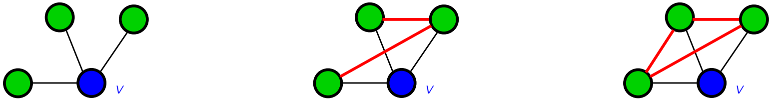
(ii) Clustering Coefficient of $G = (V, E)$:

$C^{(G)} :=$ probability that two vertices share an edge together, knowing that they have a common neighbor

$$\boxed{C^{(G)} = \frac{3 \times \text{nb of triangles}}{\text{nb of connected triples}} = \frac{\sum_v P_v}{\sum_v N_v}}$$

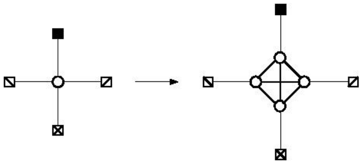
$P_v :=$ nb of pairs of neighbors of v sharing an edge together,
$N_v :=$ nb of pairs of neighbors of v : $N_v = d_v(d_v - 1)/2$.

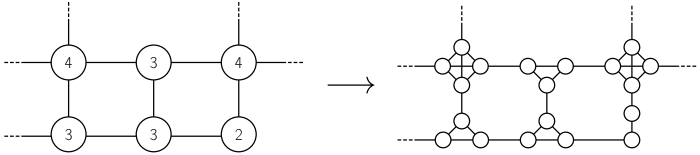
Example : $N_v = 3$



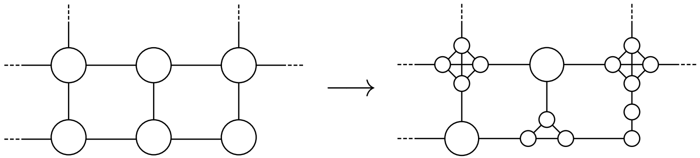
$P_v = 0$ $P_v = 2$ $P_v = 3$

- Idea : Replace a vertex of degree r in $G(n, \boldsymbol{d})$ by a clique of size r :





- Idea : Replace a vertex of degree r in $G(n, \boldsymbol{d})$ by a clique of size r.
- Adding cliques *randomly* : Let $\gamma \in [0, 1]$.
  Each vertex is replaced by a clique with probability γ (independently for all vertices).



- $\tilde{G}(n, \boldsymbol{d}, \gamma)$ = resulting random graph (with additional cliques)
  Similar model : (Trapman,'07), (Gleeson,'09)
- **Particular cases :**
  - $\gamma = 0 \Rightarrow \tilde{G}(n, \boldsymbol{d}, \gamma) = G(n, \boldsymbol{d})$,
  - $\gamma = 1 \Rightarrow$ all vertices in $G(n, \boldsymbol{d})$ have been replaced by cliques.
- New asymptotic degree distribution $\tilde{\boldsymbol{p}} = (\tilde{p}_k)_{k \geq 0}$
- Asymptotic clustering coefficient $C > 0$

Outline

Contagion model with parameter q :

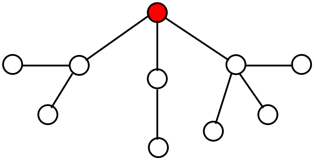
- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

$$\boxed{\text{proportion of its infected neighbors} > q}$$

Heuristically...

The random graph $G(n, \boldsymbol{d})$ converges locally to a random tree such that :

$$\mathbb{P}(r-1 \text{ children}) = rp_r/\lambda$$



$q = \frac{1}{4}$

Infected nodes = those with degree $< 1/q$

Infinite tree (of infected nodes)

$$\iff \sum_{r<1/q} (r-1)\frac{rp_r}{\lambda} > 1$$

Contagion model with parameter q :

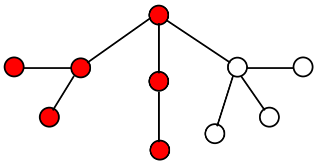
- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

$$\boxed{\text{proportion of its infected neighbors} > q}$$

Heuristically...

The random graph $G(n, \boldsymbol{d})$ converges locally to a random tree such that :

$$\mathbb{P}(r-1 \text{ children}) = rp_r/\lambda$$



$q = \frac{1}{4}$

Infected nodes = those with degree $< 1/q$

Infinite tree (of infected nodes)

$$\iff \sum_{r<1/q} (r-1)\frac{rp_r}{\lambda} > 1$$

Contagion model with parameter q :

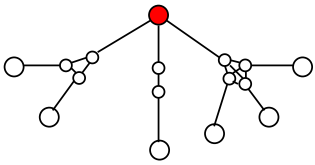
- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

$$\boxed{\text{proportion of its infected neighbors} > q}$$

Heuristically...

The random graph $G(n, \boldsymbol{d})$ converges locally to a random tree such that :

$$\mathbb{P}(r-1 \text{ children}) = rp_r/\lambda$$



$q = \frac{1}{4}$

Infected nodes = those with degree $< 1/q$

Infinite tree (of infected nodes)

$$\iff \sum_{r<1/q} (r-1)\frac{rp_r}{\lambda} > 1$$

Contagion model with parameter q :

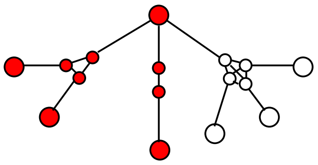
- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

$$\boxed{\text{proportion of its infected neighbors} > q}$$

Heuristically...

The random graph $G(n, \boldsymbol{d})$ converges locally to a random tree such that :

$$\mathbb{P}(r-1 \text{ children}) = rp_r/\lambda$$



$q = \frac{1}{4}$

Infected nodes = those with degree $< 1/q$

Infinite tree (of infected nodes)

$$\iff \sum_{r<1/q} (r-1)\frac{rp_r}{\lambda} > 1$$

$$q_c := q_c(\boldsymbol{p}) = \sup\left\{q' : \sum_{r<1/q'}(r-1)\frac{rp_r}{\lambda} > 1\right\}$$

Fixed q, $\mathcal{P}^{(n)}$ = set of pivotal players in $\tilde{G}(n, \boldsymbol{d}, \gamma)$:

- G_0 = induced subgraph with vertices of degree $< 1/q$
- Pivotal players = vertices in the largest connected component of G_0

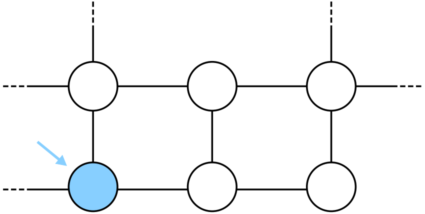
Theorem (CONTAGION THRESHOLD)

- $q < q_c$: $|\mathcal{P}^{(n)}| = \Theta_p(n)$
  Each pivotal player can trigger a global cascade.
- $q > q_c$: the size of the epidemic generated by a vertex u (chosen uniformly at random) is negligeable : $o_p(n)$.

Contagion model with parameter q :

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

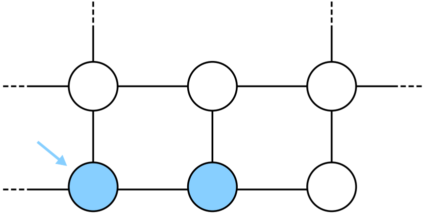
$$\boxed{\text{proportion of its infected neighbors} > q = \frac{1}{4}}$$

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Contagion model with parameter q :

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

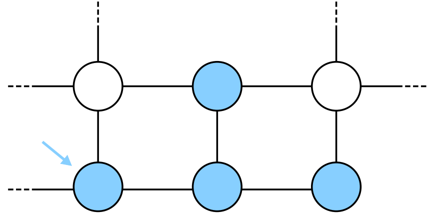
$$\boxed{\text{proportion of its infected neighbors} > q = \frac{1}{4}}$$

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Contagion model with parameter q :

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

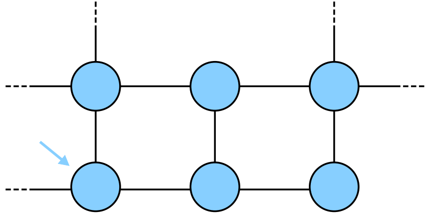
$$\boxed{\text{proportion of its infected neighbors} > q = \frac{1}{4}}$$

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Contagion model with parameter q :

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

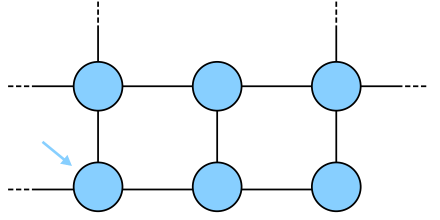
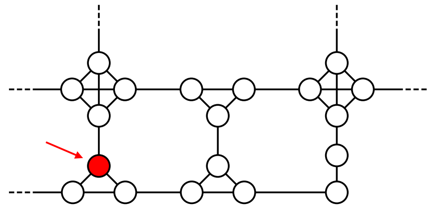
$$\boxed{\text{proportion of its infected neighbors} > q = \frac{1}{4}}$$

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Contagion model with parameter q :

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

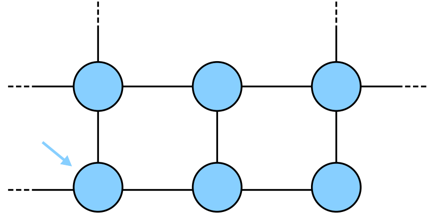
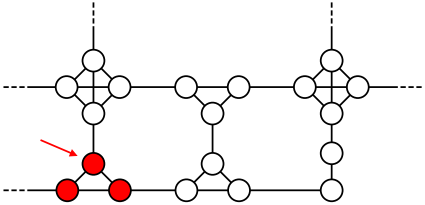
$$\boxed{\text{proportion of its infected neighbors} > q = \frac{1}{4}}$$

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Contagion model with parameter q :

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

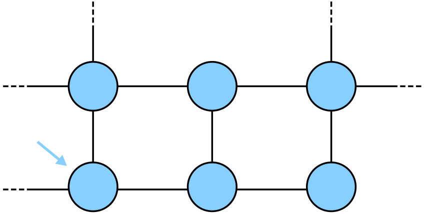
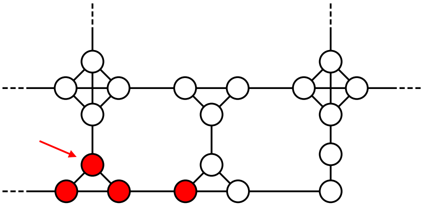
$$\boxed{\text{proportion of its infected neighbors} > q = \frac{1}{4}}$$

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Contagion model with parameter q :

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

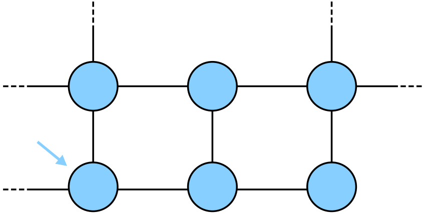
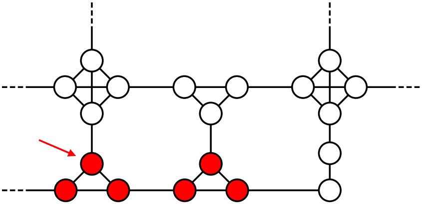
$$\boxed{\text{proportion of its infected neighbors} > q = \frac{1}{4}}$$

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Contagion model with parameter q :

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

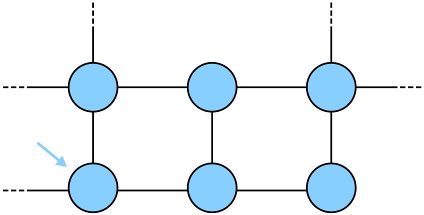
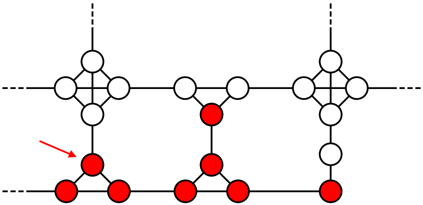
$$\boxed{\text{proportion of its infected neighbors} > q = \frac{1}{4}}$$

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Contagion model with parameter q :

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

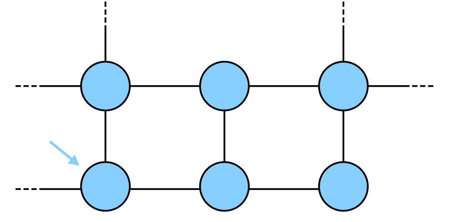
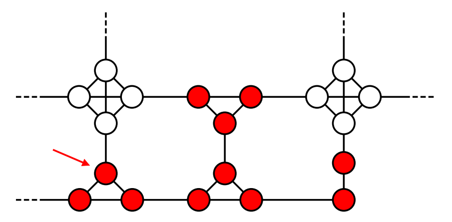
$$\boxed{\text{proportion of its infected neighbors} > q = \frac{1}{4}}$$

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Contagion model with parameter q :

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

$$\boxed{\text{proportion of its infected neighbors} > q = \frac{1}{4}}$$

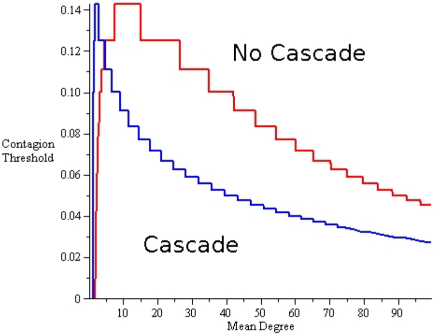
INITIAL GRAPH	GRAPH WITH CLIQUES
	

$\Longrightarrow$ Clustering decreases the cascade size.

Contagion Threshold (q_c) vs Mean Degree

Graphs with the SAME asymptotic degree distribution :

$$\tilde{p}_k \propto k^{-\tau} e^{-k/50}$$

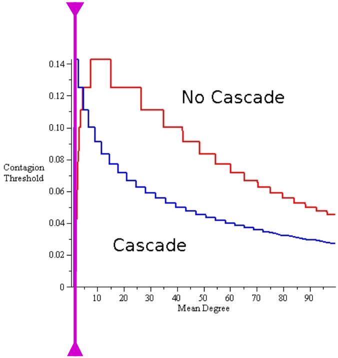


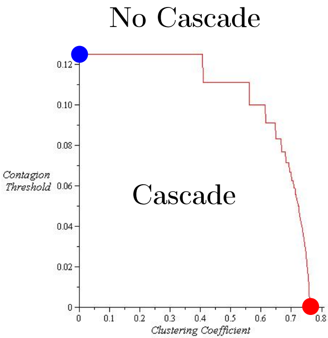
— Graph with clustering (cliques)
— Graph with no clustering

Effect of Clustering on the Contagion Threshold

Asymptotic degree distribution :
$\tilde{p}_k \propto k^{-\tau} e^{-k/50}$

Mean degree $\tilde{\lambda} \approx 1.65$



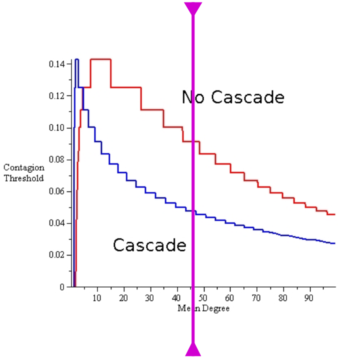


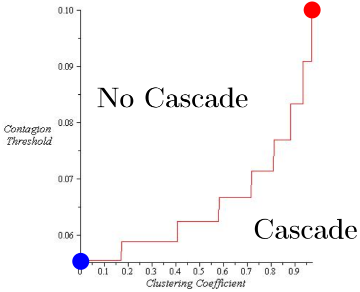
— Graph with maximal clustering coefficient
— Graph with no clustering

Effect of Clustering on the Contagion Threshold

Asymptotic degree distribution : $\tilde{p}_k \propto k^{-\tau} e^{-k/50}$

Mean degree $\tilde{\lambda} \approx 46$



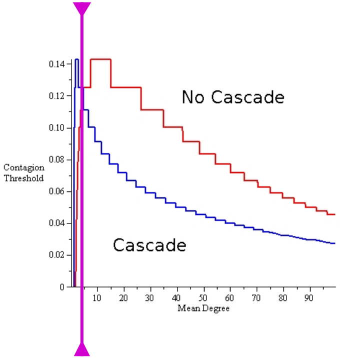


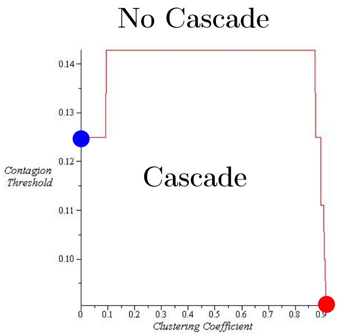
— Graph with maximal clustering coefficient
— Graph with no clustering

Effect of Clustering on the Contagion Threshold

Asymptotic degree distribution :
$\tilde{p}_k \propto k^{-\tau} e^{-k/50}$

Mean degree $\tilde{\lambda} \approx 3.22$





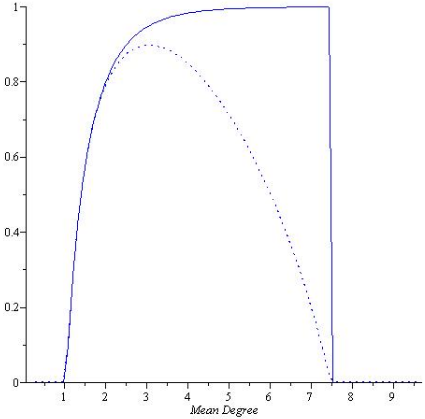
— Graph with maximal clustering coefficient
— Graph with no clustering

Effect of Clustering on the Cascade Size

Fraction of infected neighbors needed to become infected :

$$q = 0.15 \text{ (fixed)}$$

- $\tilde{p}_r = \frac{0.2r+0.8}{0.2\lambda+0.8} \frac{e^{-\lambda}\lambda^r}{r!}$
- $C = 0$



$\cdots$ Pivotal players in the graph with no clustering

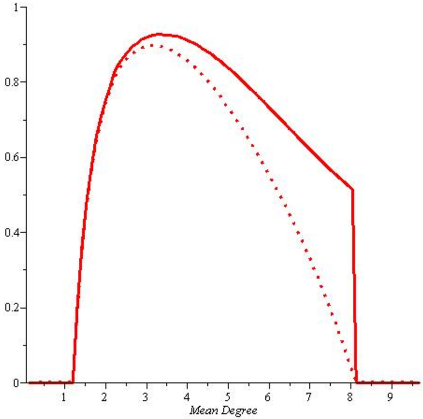
— Cascade size in the graph with no clustering

Effect of Clustering on the Cascade Size

Fraction of infected neighbors needed to become infected :

$q = 0.15$ (fixed)

- $\tilde{p}_r = \frac{0.2r+0.8}{0.2\lambda+0.8}\frac{e^{-\lambda}\lambda^r}{r!}$
- $C = \frac{0.2\lambda}{0.2\lambda+1.2} > 0$



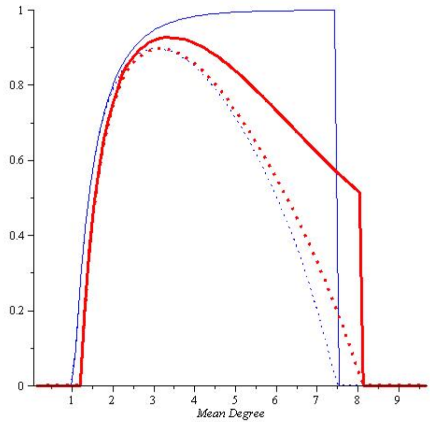
$\cdots$ Pivotal players in the graph with positive clustering

— Cascade size in the graph with positive clustering

Effect of Clustering on the Cascade Size

Fraction of infected neighbors needed to become infected :

$q = 0.15$ (fixed)



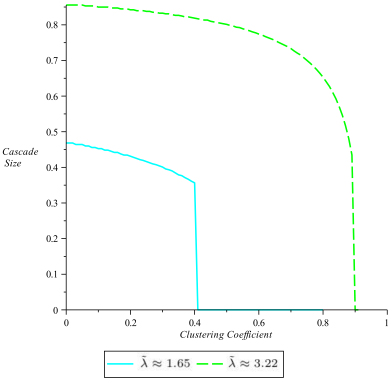
$\cdots$ Pivotal players in the graph with no clustering

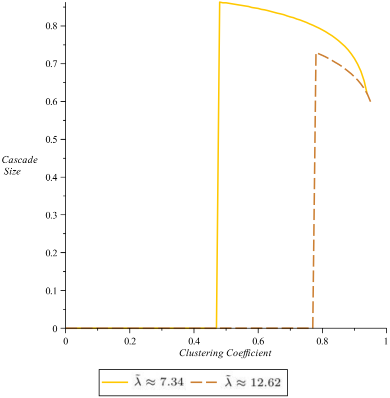
— Cascade size in the graph with no clustering

$\cdots$ Pivotal players in the graph with positive clustering

— Cascade size in the graph with positive clustering

Effect of Clustering on the Cascade Size





Asymptotic degree distribution : $\tilde{p}_k \propto k^{-\tau} e^{-k/50}$

Conclusion

- Model of random graphs with a given degree distribution, and a tunable clustering coefficient
- Effect of clustering on the contagion model :
  - Clustering decreases the contagion threshold for low values of the mean degree, while the opposite happens in the high values regime
  - Clustering decreases the cascade size (when a cascade is possible)
- For the following questions, see our paper on arXiv :1202.4974 :
  - Effect of clustering on the diffusion model (bond percolation)
  - Positive proportion of the population initially infected

Conclusion

- Model of random graphs with a given degree distribution, and a tunable clustering coefficient
- Effect of clustering on the contagion model :
  - Clustering decreases the contagion threshold for low values of the mean degree, while the opposite happens in the high values regime
  - Clustering decreases the cascade size (when a cascade is possible)
- For the following questions, see our paper on arXiv :1202.4974 :
  - Effect of clustering on the diffusion model (bond percolation)
  - Positive proportion of the population initially infected

Thanks for your attention !

If you liked the presentation, I am looking for a post-doc position for September 2012...

References

T. Britton, M. Deijfen, A. N. Lagerås, and M. Lindholm.
Epidemics on random graphs with tunable clustering.
J. Appl. Probab., 45(3) :743–756, 2008.

A. Hackett, S. Melnik, and J. P. Gleeson.
Cascades on a class of clustered random networks.
Physical Review E, 83, 2011.

M. Lelarge.
Diffusion and cascading behavior in random networks.
under revision for Games and Economic Behavior, arxiv/0805.3155, 2009.

M. E. J. Newman.
Properties of highly clustered networks.
Phys. Rev. E, 68(2) :026121, Aug 2003.

P. Trapman.
On analytical approaches to epidemics on networks.
Theoretical Population Biology, 71(2) :160–173, 2007.

D. J. Watts and S. H. Strogatz.
Collective dynamics of 'small-world' networks.
Nature, 393(6684) :440–442, June 1998.