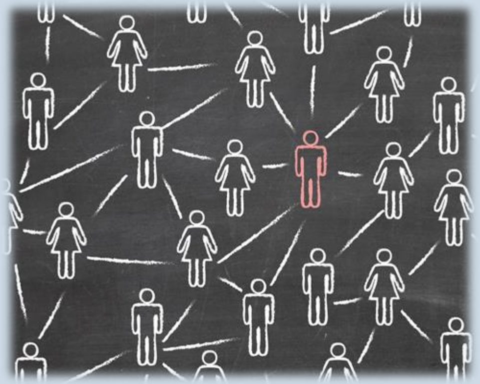
EFFICIENT ALGORITHMS FOR THE SUPPRESSION OF DIFFUSION PROCESSES ON NETWORKS

APPLICATION IN EPIDEMIOLOGY AND MARKETING



Argyris Kalogeratos

Joint work with Kevin Scaman and Nicolas Vayatis

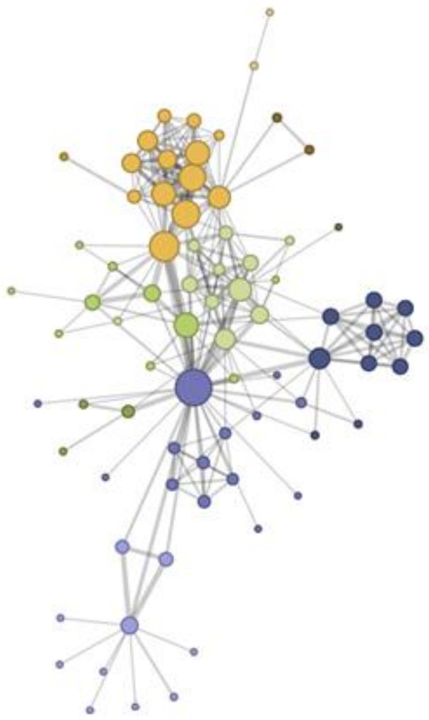
ENS CACHAN · CMLA Centre de Mathématiques et de Leurs Applications · MLMDA MACHINE LEARNING AND MASSIVE DATA ANALYSIS

Big Data et politiques publiques dans les transports : séminaire d'information et d'acculturation

Paris, 15 October 2015

DIFFUSION PROCESSES ON NETWORKS

Basics



DPs arise in systems with interconnected agents (real or electronic network)

- each agent has a *variable state*
- agent behavior depends on, and propagates to, its close environment
- the propagation causes changes in agents' state according to some "*rules*"

Propagating entities: ***from disease epidemics to… digital and social epidemics***

- *Epidemiology*: diseases/viruses
- *Computer systems*: computer viruses, fault cascade, computational errors (e.g. sensor networks)
- *Social and information networks*: information, ideas, rumors, social behaviors…

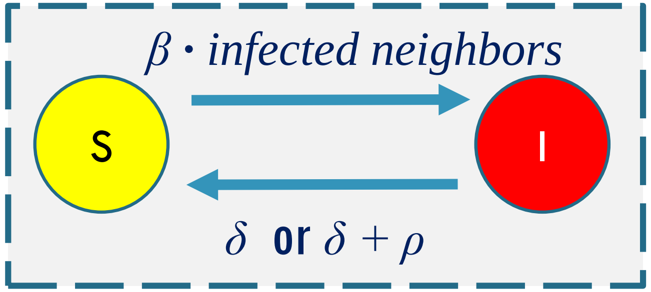
DIFFUSION PROCESSES ON NETWORKS

Diffusion Models – SIS demonstration

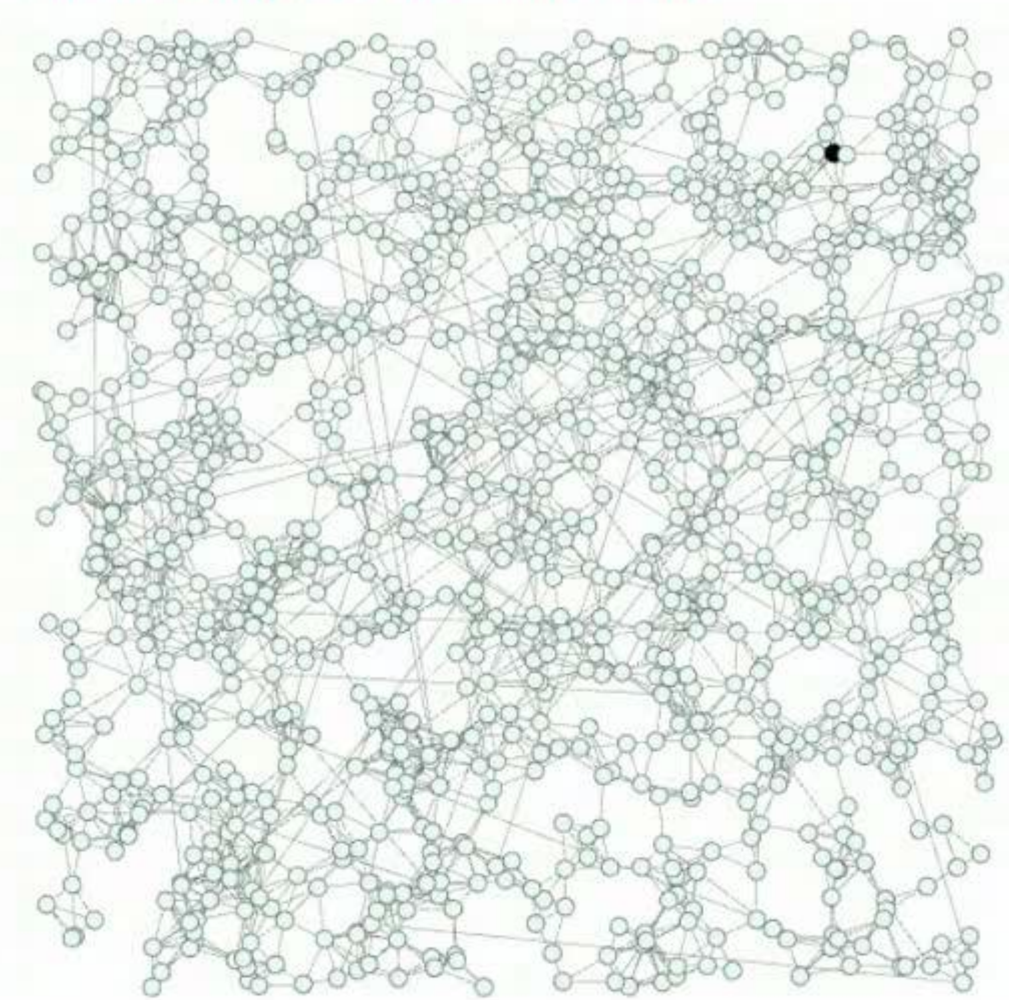
Diffusion model

- a mathematical model that encodes the "*propagation rules*"
- no single model able to describe all possible complex diffusion phenomena
- well-studied: **Compartmental Models** from epidemiology (SIS, SIR, SEIR, …)

Continuous-time SIS model for one node



SIS diffusion process in a contact nework



Watch on youtube: http://www.youtube.com/watch?v=fGSKHxSD-40

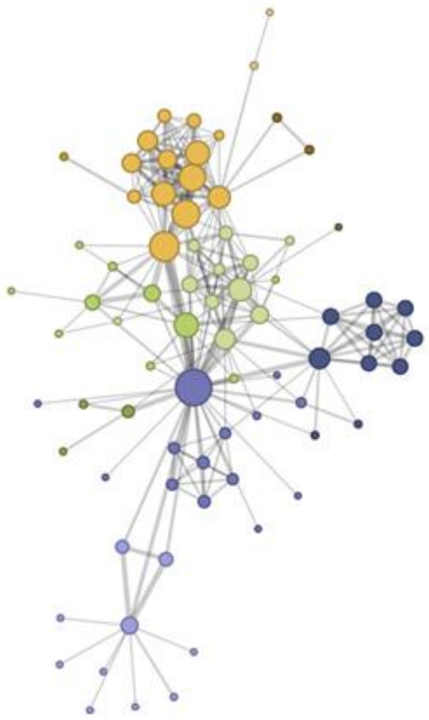
DIFFUSION PROCESSES ON NETWORKS

Directions of research

Depending on the situation, a DP can be desired or undesired

Roughly three directions of research

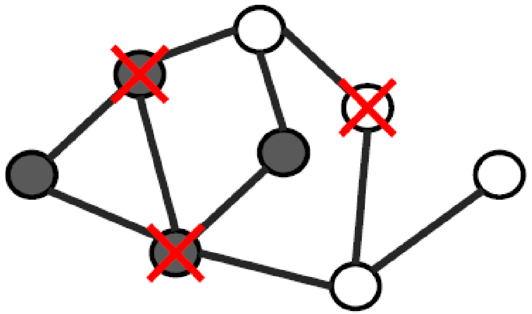
- **Network assessment:** worst case analysis, risk/vulnerability assessment
- **DP engineering:** influence maximization, (viral) marketing
- **DP suppression and control:** containment of viruses, rumors, social behaviors, etc., using ***control actions***.



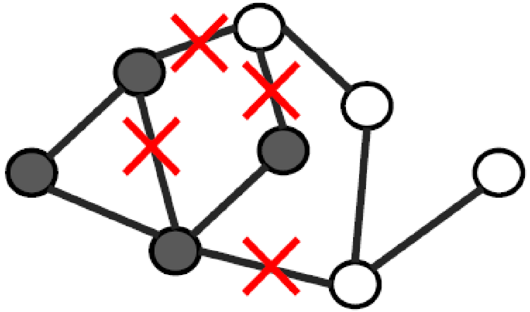
DIFFUSION SUPPRESSION AND CONTROL

Possible control actions

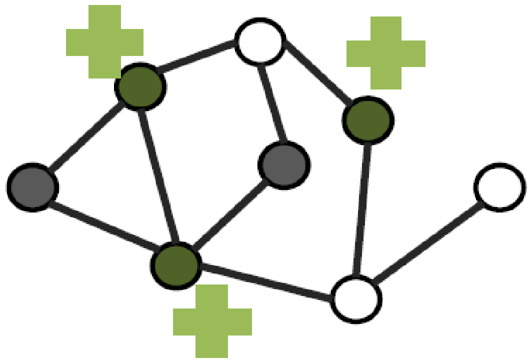
DP suppression and control using ***control actions*** on <u>nodes</u> or <u>edges</u>, such as:



Node deletion



Edge deletion

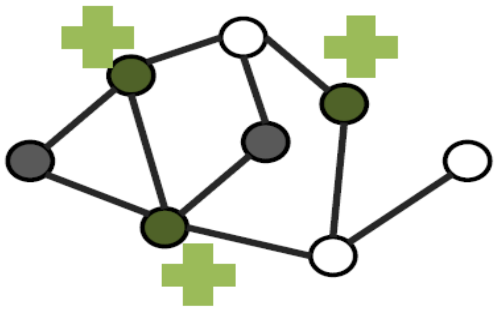


Resources on nodes

DIFFUSION SUPPRESSION AND CONTROL

Healing resources on nodes – more variations

DP suppression and control using ***control actions*** on nodes, such as:



Resources on nodes:

preventive → *vaccines*

preparatory → *antidotes*

Dynamic Resource Allocation

corrective → *treatments* (also → *antidotes*)

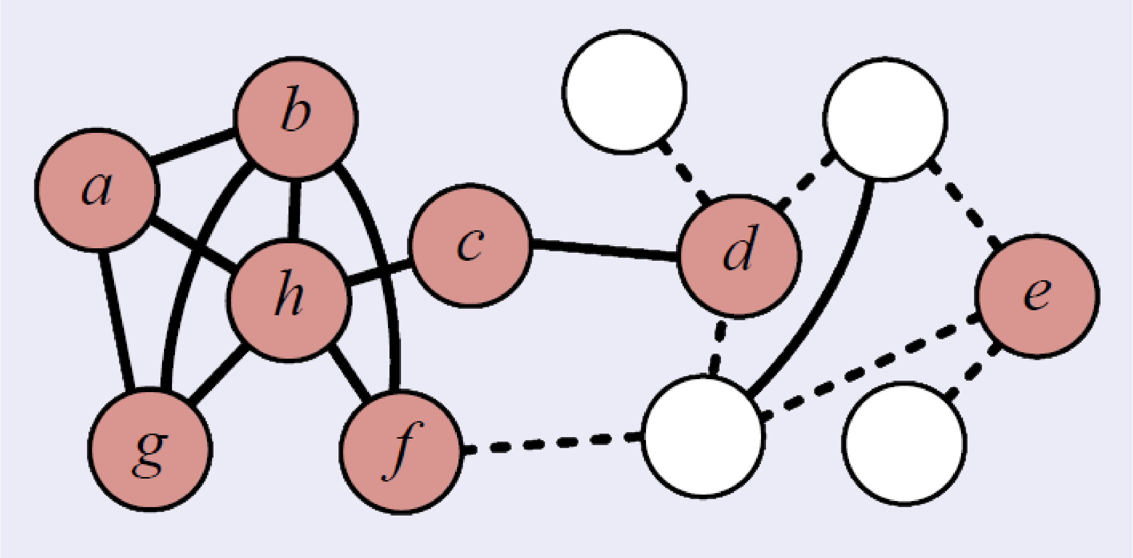
Our setup

- a budget of b_{tot} treatments is available at any moment
- offer additional recovery rate ρ to a receiver infected node

GREEDY DYNAMIC RESOURCE ALLOCATION

Largest Reduction of Infectious Edges (LRIE) Strategy [1, 2]

Toy example



LRIE node ranking:

Priority 1: e

Priority 2: d

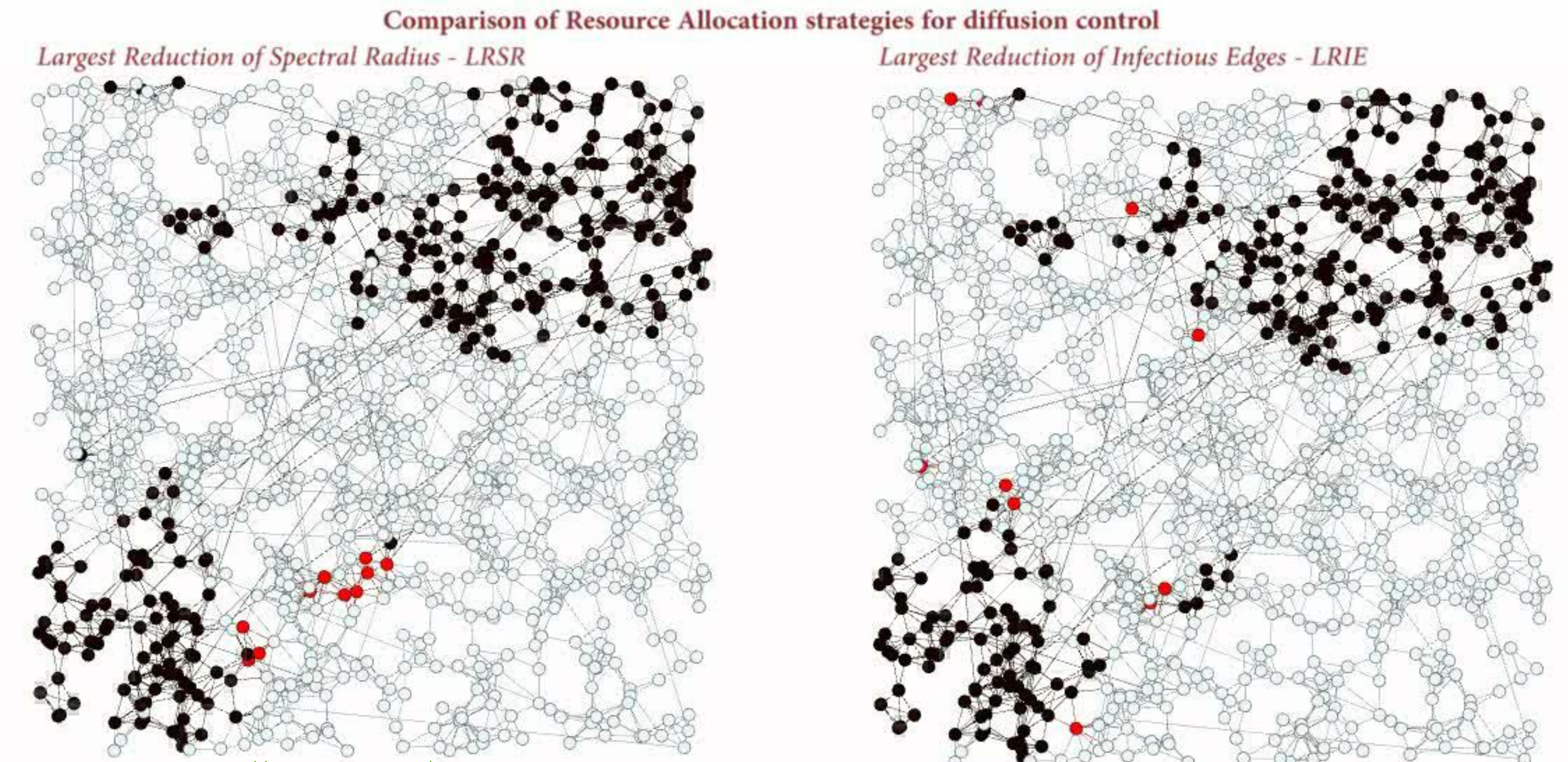
Priority 3: f

…

- The red nodes are infected, the dashed edges are ***infectious***
- Node ***h*** is the most central
- Node **e** and ***d*** are the most viral (can infect others)
- Node **e** is the safest

GREEDY DYNAMIC RESOURCE ALLOCATION

Demonstration



Watch on youtube: http://www.youtube.com/watch?v=xS-0p7h1OeM

PROBLEM SOLVED ?

Question

Is there a way to make an efficient plan by respecting the network properties and follow it persistently throughout the whole process ?

What kind of guarantee could be provided ?

GLOBAL PRIORITY PLANNING

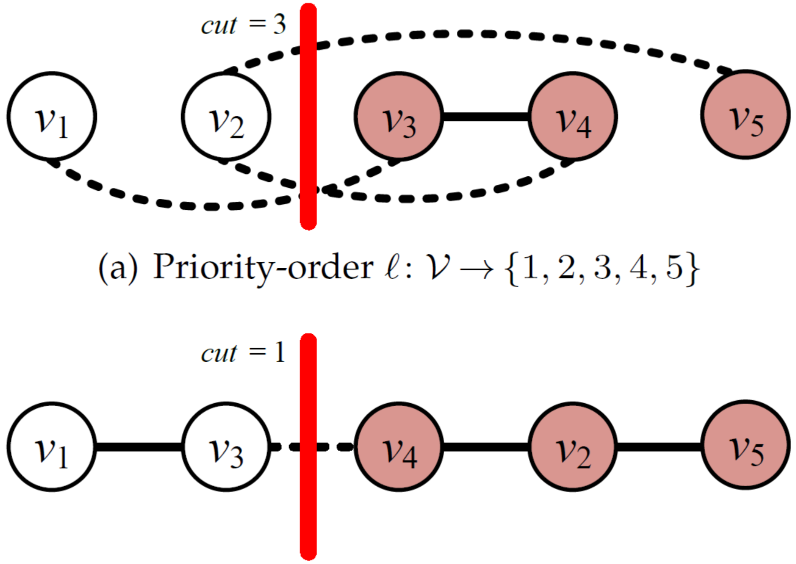
MaxCut Minimization (MCM) Strategy [3, 4]

MCM strategy

- seeks for the priority-order ℓ with the ***minimum maxcut*** $C^*(\ell)$ of edges
- heals the b_{tot} leftmost infected nodes in ℓ
- we proved a strong connection between the *expected extinction time* and *the most difficult step of the plan* which would need a suppression efficiency $\sim \beta\, C^*(\ell)$:

$$b_{tot}\,(\rho + \delta) > \beta\, C^*(\ell)$$

Toy example



(a) Priority-order $\ell: \mathcal{V} \rightarrow \{1, 2, 3, 4, 5\}$

(b) Priority-order $\ell' : \mathcal{V} \rightarrow \{1, 3, 4, 2, 5\}$

- The red nodes are infected, the dashed edges are ***infectious***
- The red vertical line denotes the **front** separating the healthy (left) from the infected part (right) of the network

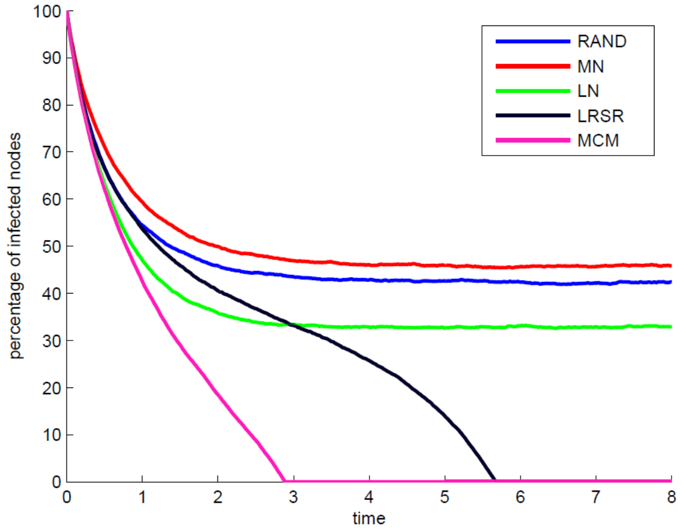
GLOBAL PRIORITY PLANNING

Experiments with simulated diffusions on real network

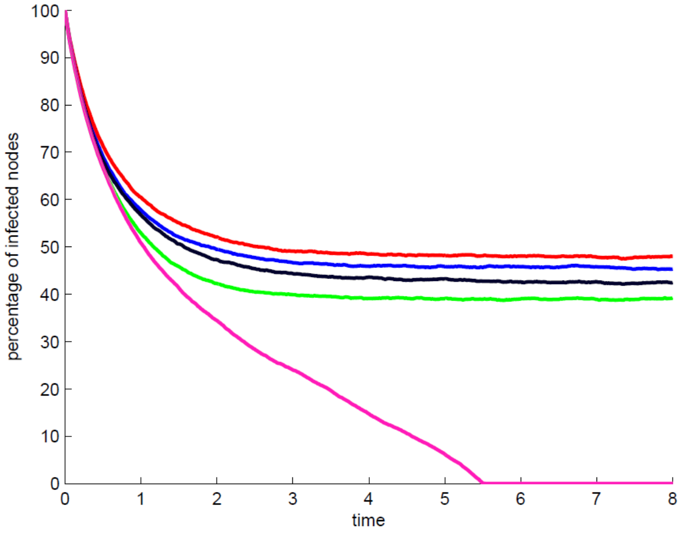
Subset of Twitter network with 81.306 nodes

Strategy	Maxcut	Maxcut % w.r.t. RAND	Expected resource threshold ($\delta = 1$, $\beta = 0.1$, $q = 100$)
RAND	670,000 ± 1000	100.0 %	67,000
MN	628,571	93.8 %	62,957
LN	628,571	93.8 %	62,957
LRSR	349,440	52.2 %	34,944
MCM	**71,956**	**10.7 %**	**7,196**

MCM can remove the contagion with ~5 times less resources than its best competitor !!



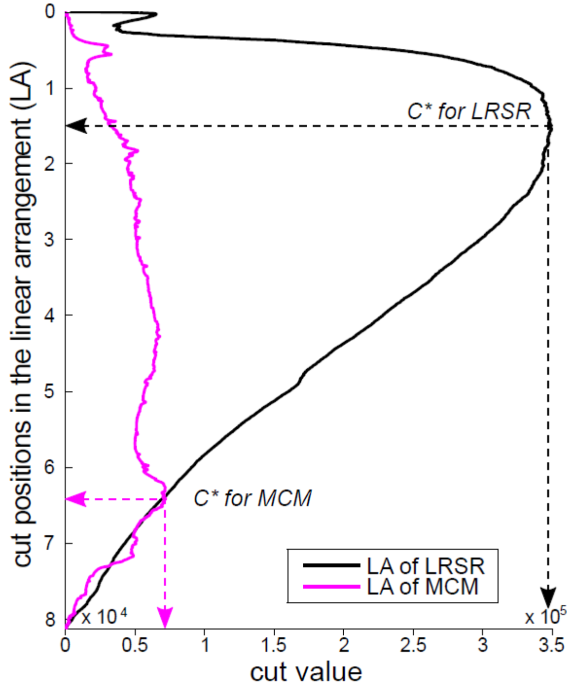
(a) high resource budget: $r = 20,000$



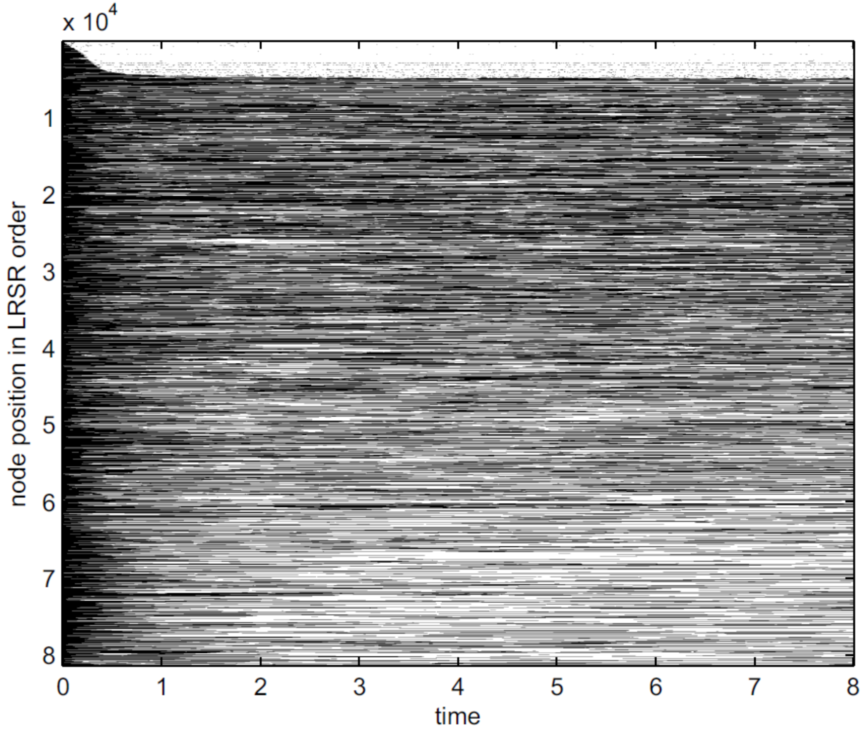
(b) low resource budget: $r = 12,000$

GLOBAL PRIORITY PLANNING

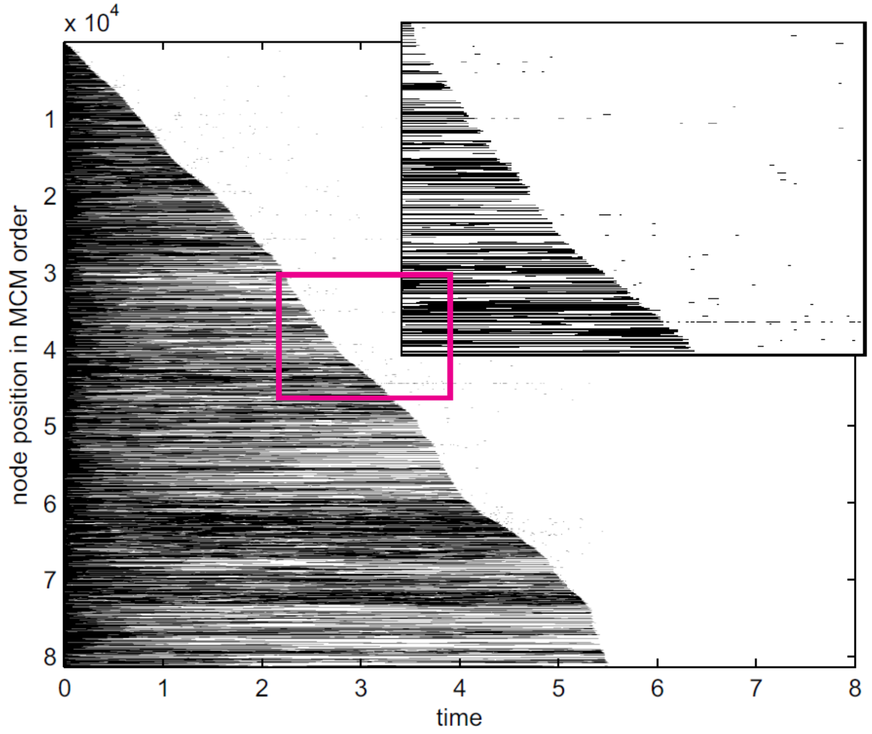
Experiments on a subset of Twitter graph



(c) cuts and maxcuts



(d) network state under LRSR



(e) network state under MCM

APPLICATIONS

Potential playground for diffusion control

- **Google Flu Trends**
- **Monitoring Ebola Outbreak**
- **Foursquare check-ins**
- **Social behaviors: the case of *obesity***

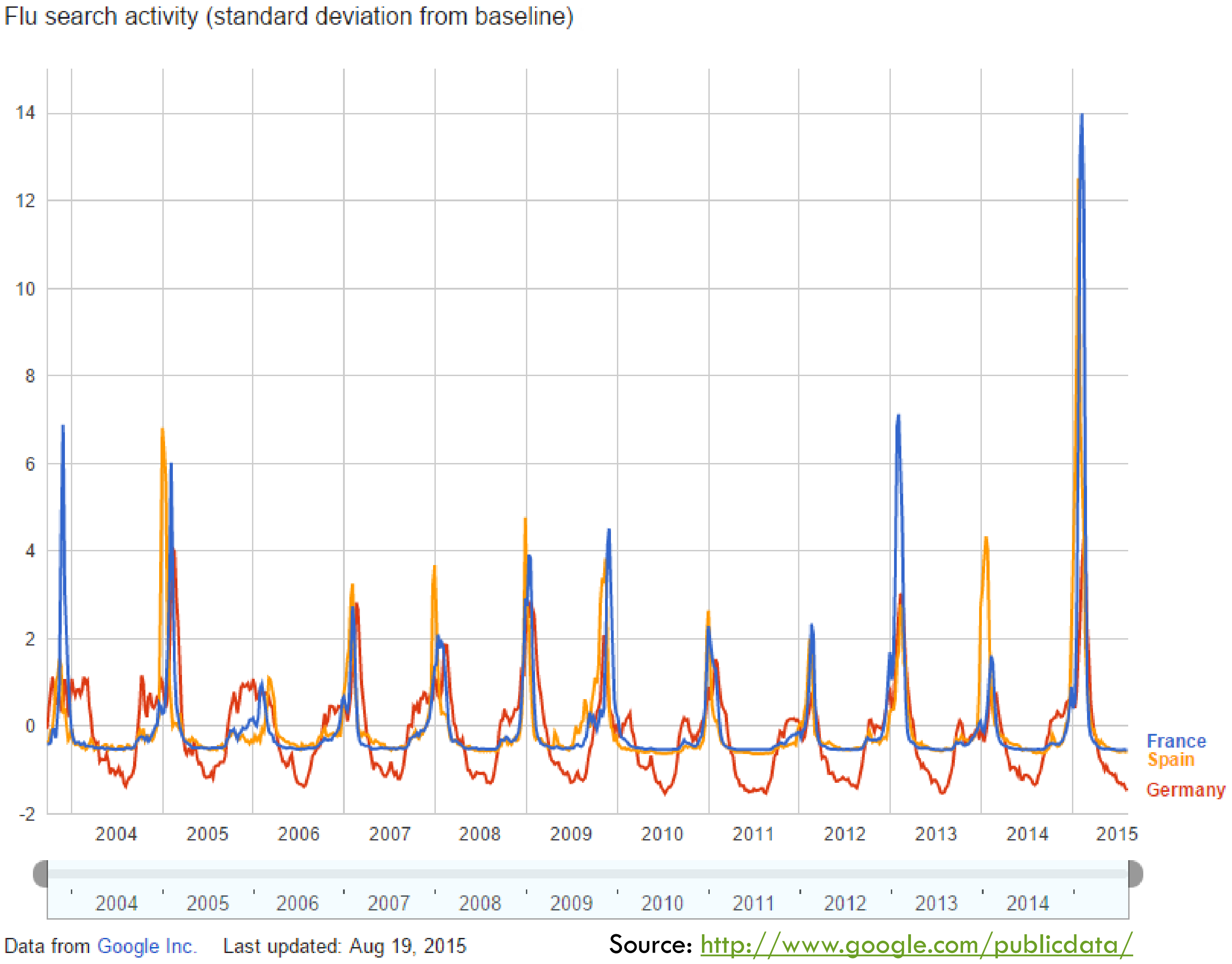
APPLICATIONS
Epidemics

Google Flu Trends (2008-2015)

- *claiming* to predict the flu trend with 97% accuracy

The figure shows:

- google search activity about flu
- comparison between France, Spain, and Germany
- period: 12/2004 – 08/2015



Source: http://www.google.com/publicdata/

APPLICATIONS

Epidemics

Google Flu Trends (2008-2015)

- *claiming* to predict the flu trend with 97% accuracy

The video shows:

- worldwide google search activity about flu
- period: 12/2002 – 08/2015

Source: http://www.google.com/publicdata/

APPLICATIONS

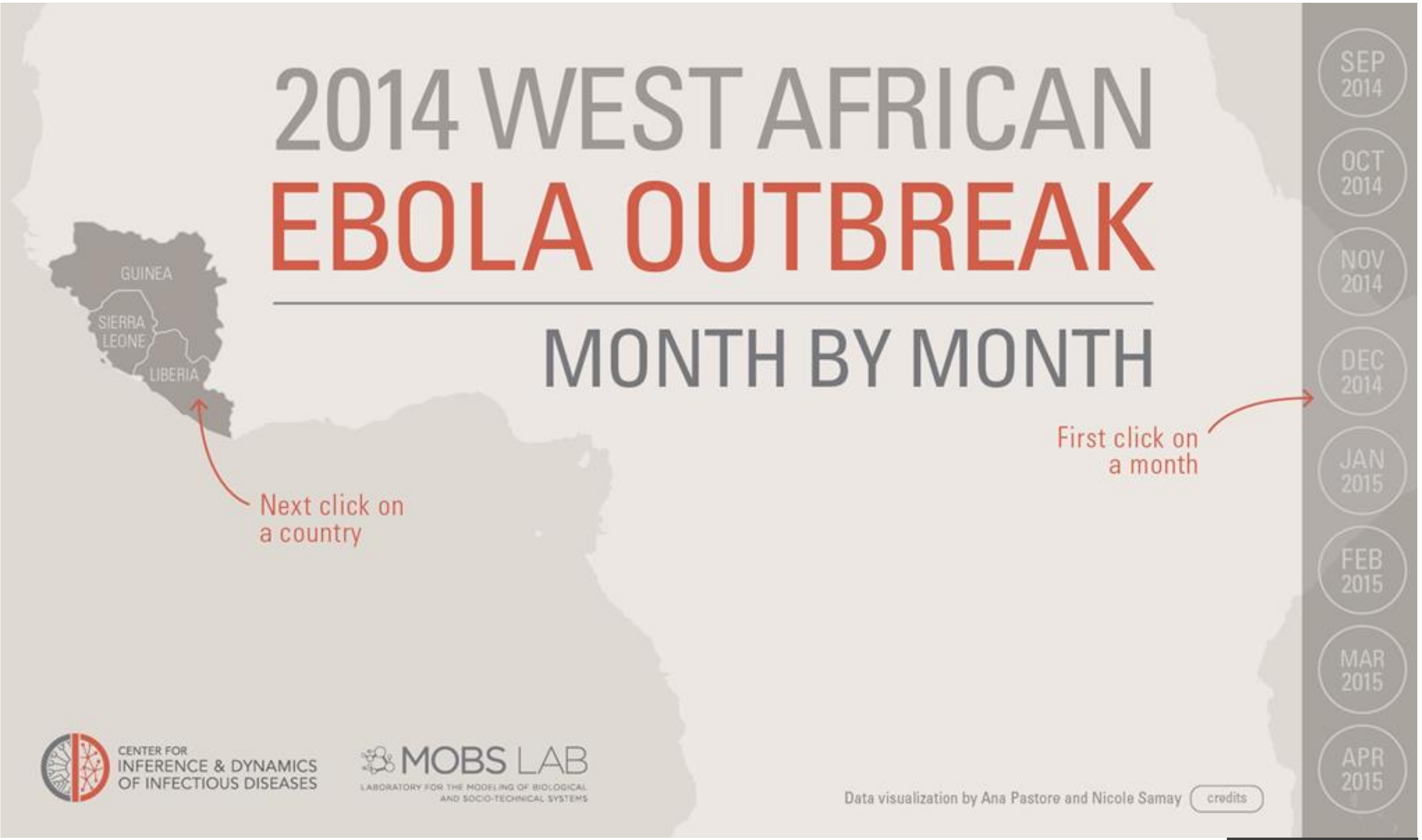
Epidemics - transportation

Monitoring Ebola Outbreak (2014-2015)

- using medical records and Ebola news mentions for predicting the spread
- developed by **MOBS Lab,** Northeastern University, Boston

The figure shows:

- monthly reports for the Ebola outbreak
- period: 09/014 – 04/2015



Source: http://www.mobs-lab.org/ebola.html

APPLICATIONS

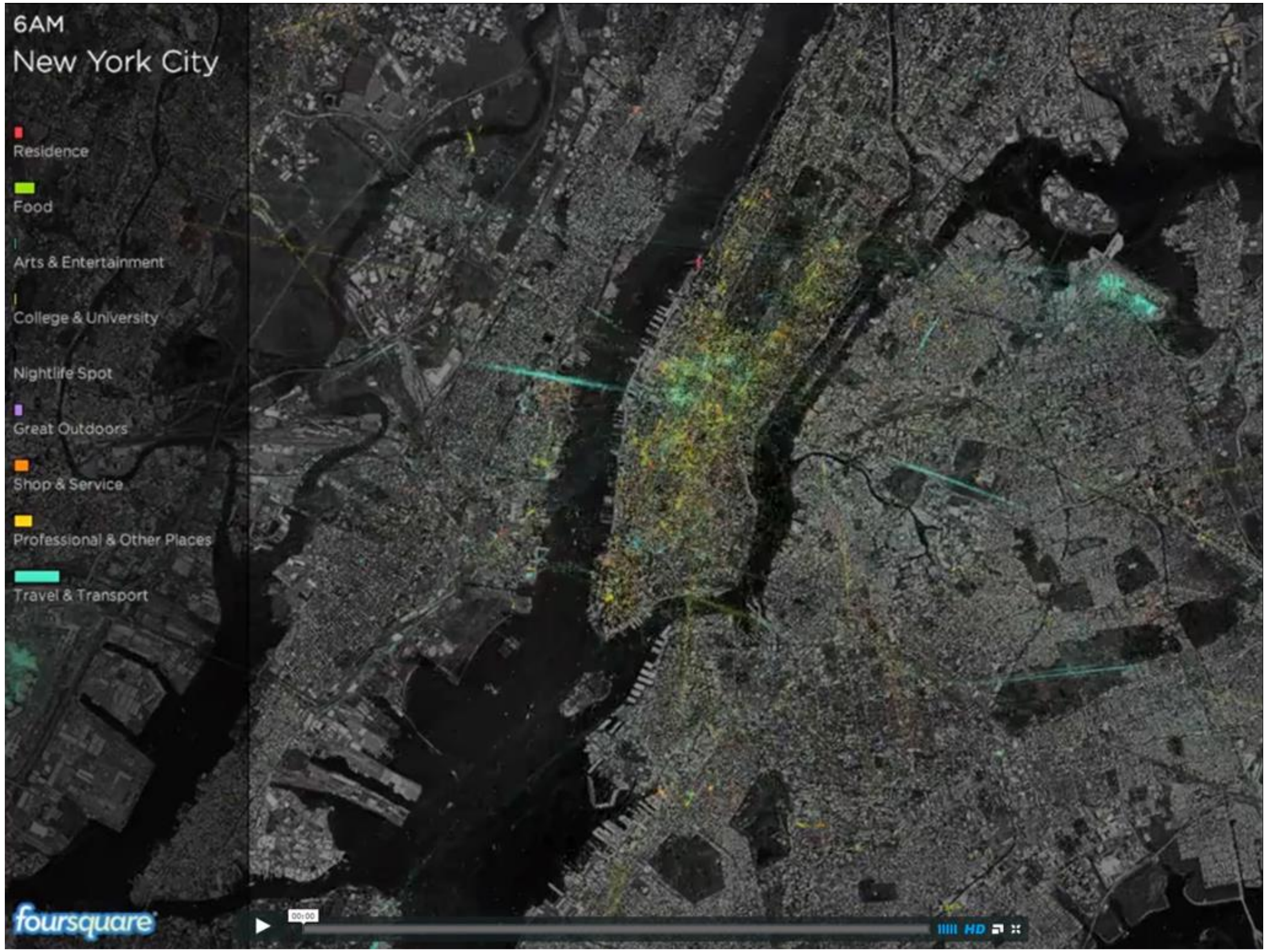
Transportation - urban mobility

Foursquare check-ins

- check-in of users to places, e.g. restaurants, cafes, museums, etc.

The video shows:

- the underlying flow of people indicated by their check-ins
- period: 1 day in 2014



Source: http://foursquare.com/infographics/pulse

APPLICATIONS

Social behaviors - obesity

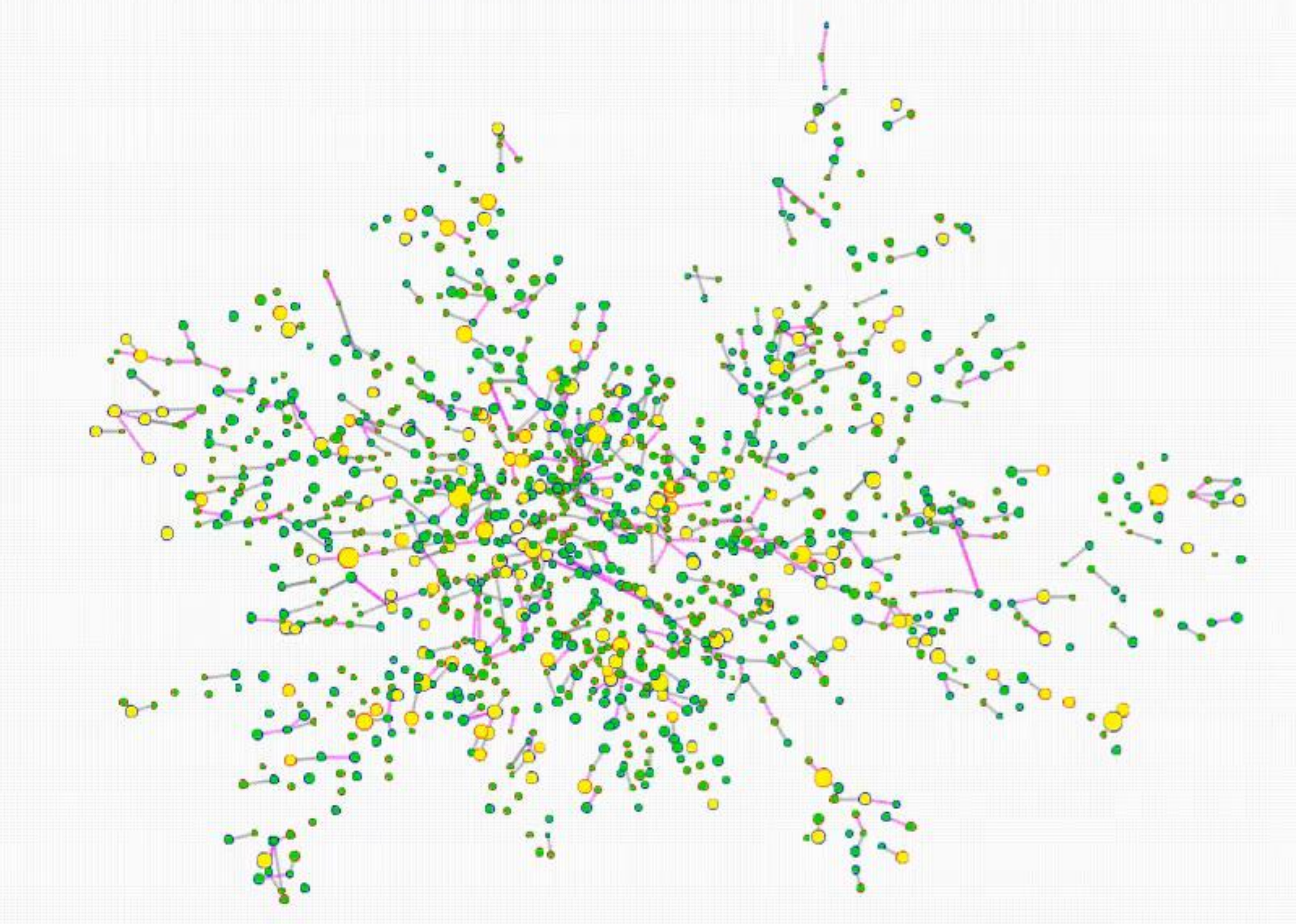
Social behaviors are diffusive

The case of ***obesity***:

- Records for 2.200 people of a community for over 32 years

The video shows:

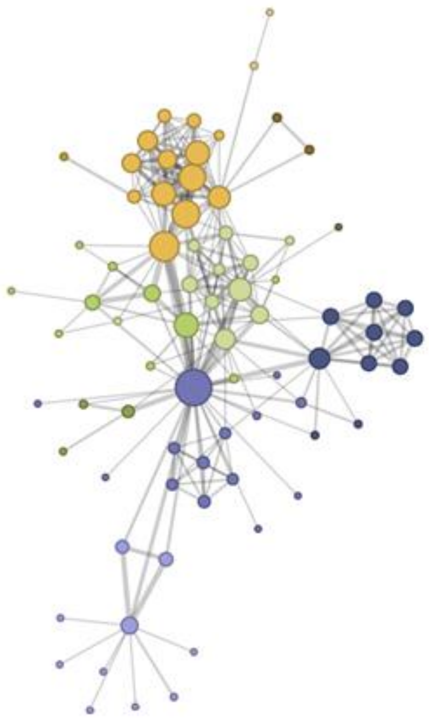
- the evolution of BMI index of the population and how obesity (larger yellow circles) diffuses and creates clusters
- study: 2007, data: 1971-2003



Source: The Spread of Obesity in a Large Social Network over 32 Years, 2007 [5].

CONCLUSION

- Diffusion processes and control
- We presented two efficient methods for dynamic resource allocation
  - Computational approaches that can be applied in multiple network resolutions
  - They can be used for epidemic control, MCM also as an assessment tool
    a. The **MaxCut** assesses the quality of a plan
    b. The **Minimum MaxCut** assess the resource needs of a network
- DPs are super-significant in the new socio-economic context
  - the potential playground for large-scale DP control was illustrated
  - some applications interesting for the public affairs and private sector were highlighted



REFERENCES

[1] ***A Greedy Approach for Dynamic Control of Diffusion Processes in Networks***, K. Scaman, A. Kalogeratos, and N. Vayatis, to appear in the IEEE International Conference on Tools with Artificial Intelligence (ICTAI), November 9-12, 2015 -- [pdf][suppl. material] .::||::. [software].

[2] ***Dynamic Treatment Allocation for Epidemic Control in Arbitrary Networks***, K. Scaman, A. Kalogeratos, N. Vayatis, to appear in WSDM 2014 Diffusion in Networks and Cascade Analytics (DiffNet) Workshop, February 2014 -- [workshop link][pdf][slides][poster].

[3] ***What Makes a Good Plan? An Efficient Planning Approach to Control Diffusion Processes in Networks***, K. Scaman, A. Kalogeratos, and N. Vayatis, CORR arXiv:1407.4760, 18 pages, July 17, 2014 -- [arXiv link][pdf].

[4] ***Suppressing epidemics on Networks with Treatments of Limited Efficiency***, K. Scaman, A. Kalogeratos, and N. Vayatis, *paper under journal review*, 2015.

[5] ***The Spread of Obesity in a Large Social Network over 32 Years*****,** N. A. Christakis, J. H. Fowler, N. Engl. J. Med. 357;4, 2007.

RESOURCES

[1] Google Flu Trends, *http://en.wikipedia.org/wiki/Google_Flu_Trends*.

[2] Foursquare, *http://www.fastcodesign.com/3018574*.

[3] MOBS Lab, Northeastern University, Boston, MA, US, *http://www.mobs-lab.org*.

QUESTIONS

Thank you!