Network Science

Spreading Phenomena

Albert-László Barabási

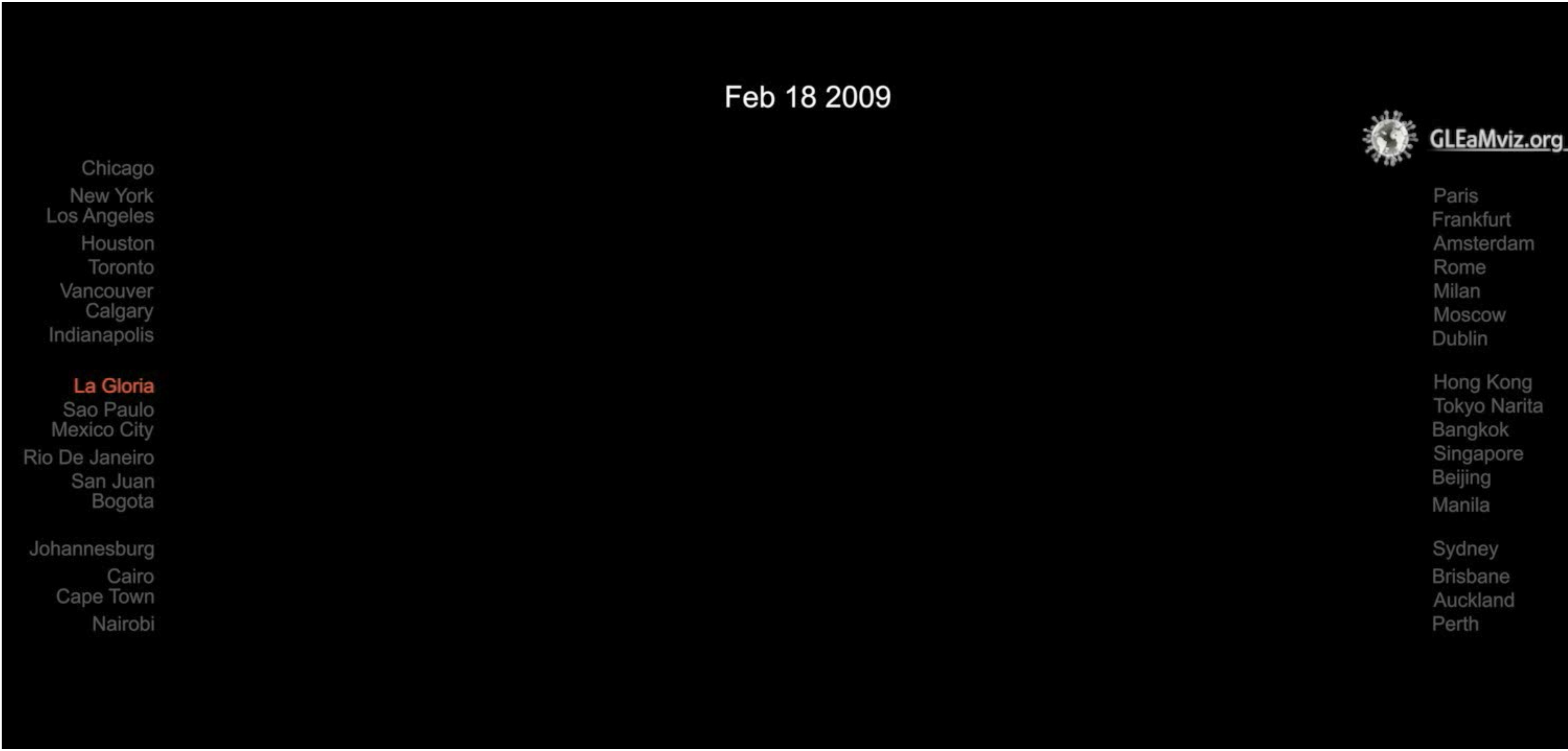
with

Emma K. Towlson and Sean P. Cornelius

www.BarabasiLab.com

Case Study 1: Epidemic Forecast

H1N1



Feb 18 2009
GLEaMviz.org
Chicago
New York
Los Angeles
Houston
Toronto
Vancouver
Calgary
Indianapolis
La Gloria
Sao Paulo
Mexico City
Rio De Janeiro
San Juan
Bogota
Johannesburg
Cairo
Cape Town
Nairobi
Paris
Frankfurt
Amsterdam
Rome
Milan
Moscow
Dublin
Hong Kong
Tokyo Narita
Bangkok
Singapore
Beijing
Manila
Sydney
Brisbane
Auckland
Perth

Section 10.7 Real-Time Forecasts

http://vimeo.com/user3371919

Section 10.7 Epidemic Prediction

- Where did the pathogen originate?
- Where do we expect new cases?
- When will the epidemic arrive at various densely populated areas?
- How many infections are to be expected?
- What can we do to slow its spread?
- How can we eradicate it?

Section 10.7 Peak Time

- **Peak Time**
  Peak time corresponds to the week when most individuals are infected in a particular country. Predicting the peak time helps health officials decide the timing and the quantity of the vaccines or treatments they distribute. The peak time depends on the arrival time of the first infection and the demographic and the mobility characteristics of each country. The observed peak time fell within the prediction interval for 87% of the countries (Figure 10.27). In the remaining cases the difference between the real and the predicted peak was at most two weeks.

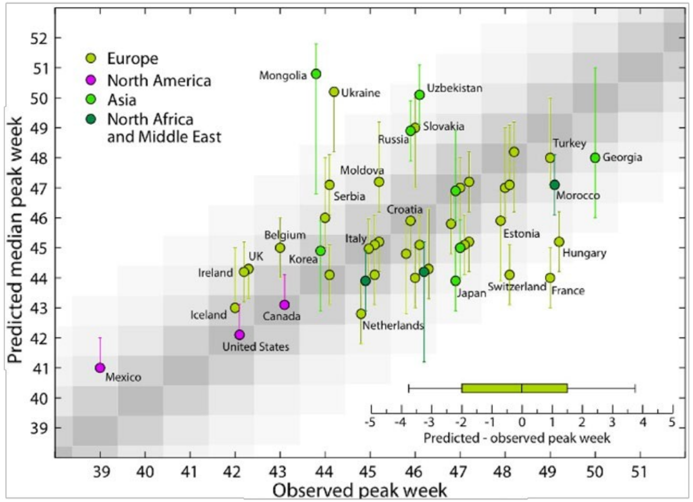


Figure 10.29

Activity Peaks for H1N1

The predicted and the observed activity peaks for the H1N1 virus in several countries. The peak week corresponds to the week when most individuals are infected by the disease, and is measured in weeks after the beginning of the epidemic. The model predictions were obtained by analyzing 2,000 stochastic realizations of the outbreak, generating the error bars in the figure. After [82].

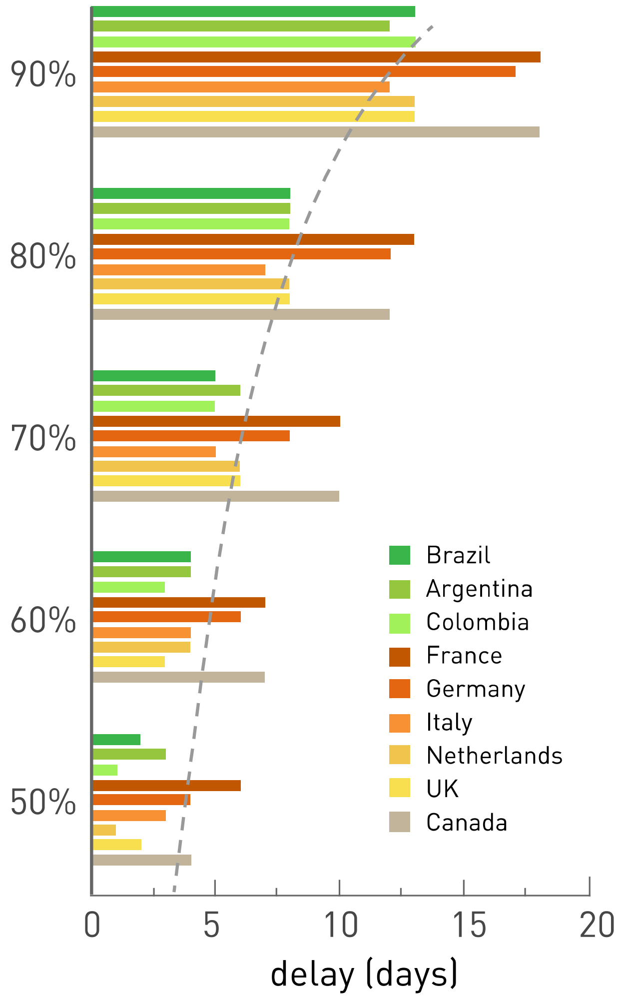
Section 10.7 Peak Time

- **Early Peak**

  GLEAM predicted that the H1N1 epidemic will peak out in November, rather than in January or February, the typical peak time of influenza-like viruses. This unexpected prediction turned out to be correct, confirming the model's predictive power. The early peak time was a consequence of the fact that H1N1 originated in Mexico, rather than South Asia (where many flu viruses come from), hence it took the virus less time to arrive to the northern hemisphere.

- **The Impact of Vaccination**

  Several countries implemented vaccination campaigns to accelerate the decline of the pandemic. The simulations indicated that these mass vaccination campaigns had only negligible impact on the course of the epidemic. The reason is that the timing of these campaigns was guided by the expectation of a January peak time, prompting the deployment of the vaccines after the November 2009 peak [83], too late to have a strong effect.



The Impact of Travel Reduction

The predicted delay in the arrival time of the H1N1 virus from Mexico to various countries under travel reduction, compared with the reference scenario of no travel reduction. The percentages on the vertical axis show the degree of travel reduction implemented around the world. The largest delay is less than 20 days, observed for a 90% travel restriction. After [77].

Section 10.7 Effective Distance

Given the multiple routes a person can take between any two cities, a pathogen can follow multiple paths on the mobility network. Yet, its spread is dominated by the most probable trajectories predicted by the mobility matrix p_{ij}. This allows us to define the *effective distance* d_{ij} between two connected locations i and j, as

$$d_{ij} = (1 - \log p_{ij}) \geq 0. \tag{10.31}$$

If p_{ij} is small, implying that only a small fraction of individuals that leave from i travel to j, then the effective distance between i and j is large. Note that $d_{ij} \neq d_{ji}$: For a small village i located near a metropolis j we expect d_{ij} to be small, as most travelers from i go to j. Yet, d_{ji} is large as only a small fraction of travelers leaving the metropolis head to the small village. The logarithm in (10.31) accounts for the fact that effective distances are additive, whereas probabilities along multi-step paths are multiplicative.

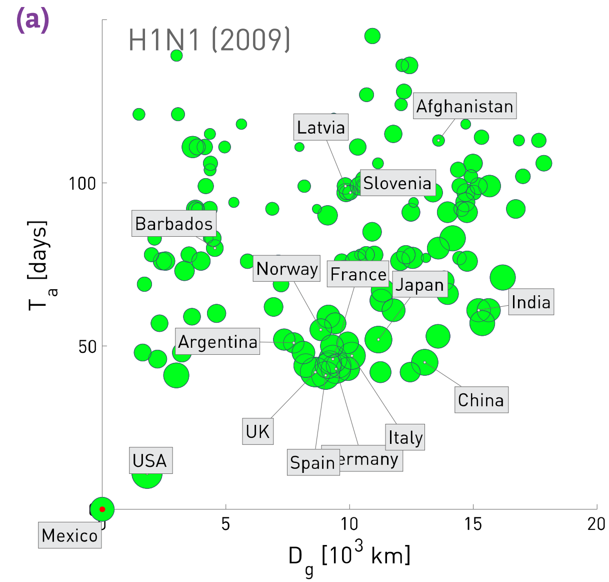
$$T_a = \frac{d_{eff}(P)}{V_{eff}(\beta, R_0, \gamma, \varepsilon)} \tag{10.32}$$

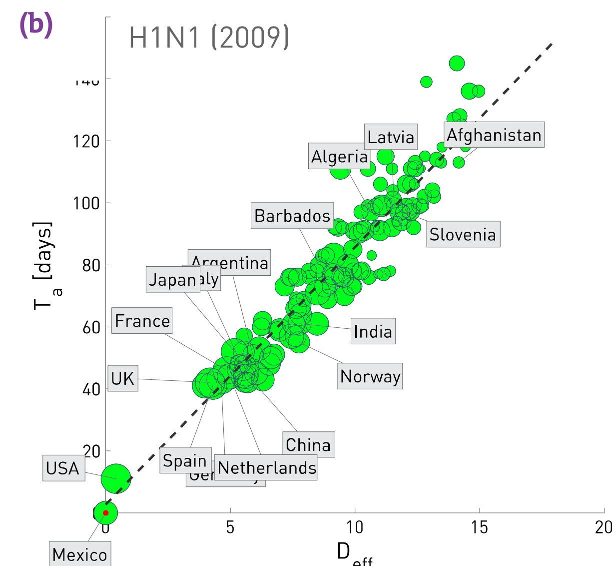
Therefore the arrival time is the ratio of the effective distance d_{eff} and an effective speed V_{eff}. The effective speed is determined only by the epidemiological parameters, whereas the effective distance d_{eff} depends only on the topology of the mobility network encoded in p_{ij}. When confronted with a new outbreak, the disease-specific epidemiological parameters are unknown in the beginning. However, (10.32) predicts that the *relative arrival times are independent of the epidemiological parameters*. For example, for an outbreak that starts at node i, the ratio of the arrival times to nodes j and l is

$$\frac{T_a(j|i)}{T_a(l|i)} = \frac{d_{eff}(j|i)}{d_{eff}(l|i)},$$

i.e. the ratio depends only on the effective distances. Therefore, the relative arrival times of the disease depend only on the topology of the mobility network. As the mobility patterns around the world are unique and model-independent, the predictions of different models converge, independent of the choice of the epidemiological parameters.

Section 10.7 Effective Distance





Effective Distance and Arrival Time

(a) Classical Representation

Arrival times vs. geographic distance from its source (Mexico) for the 2009 H1N1 pandemic. Each circle represents one of the 140 affected countries and the symbol size indicates the total traffic in each country. Arrival times are the date of the first confirmed case in a given country after the beginning of the outbreak on March 17, 2009. In this representation the arrival time and the geographic distance are largely independent of each other (R_0=0.0394).

(b) Effective Distance

Epidemic arrival time T_a vs. effective distance D_{eff} for H1N1, demonstrating the strong correlations between the effective distance (10.31) and the arrival time. After [89].

Section 10.7 IDENTIFYING THE SOurCE OF A PANDEMIC

- In the simplest case at a given moment t we know the nodes that have been infected and the network on which the pathogen spreads. The task is to find the source i [91] (Figure 10.33).

- If we also have the time of infection for each node, we can reconstruct the dynamics of the epidemic, significantly enhancing our ability to detect the source.

- The best strategy is to monitor the hubs, as they have the earliest and the most accurate information about a breakout. For example, for a pathogen spreading on a scale-free network, monitoring the state of 18% of the highest degree nodes can offer a 90% success rate in detecting the source. In contrast, we need to monitor 41% of the nodes if we select them randomly to achieve the same level of accuracy [93].

- In the effective distance representation (Figure 10.31) the infection follows a circular pattern only if we use the right outbreak location. Otherwise the observed pattern is asymmetric. Therefore, we can detect the source by finding the location (node) from which the outbreak pattern shows the highest radial symmetry [89].

Summary

Section 10.8

AT A GLANCE: NETWORK EPIDEMICS

Infection Rate: β

Recovery Rate: μ

Spreading Rate: $\lambda \equiv \frac{\beta}{\mu}$

Reproductive Number: $R_0 = \frac{\beta\langle k\rangle}{\mu}$

SI Model:

$$i(t) = \frac{i_0 \exp(\beta\langle k\rangle t)}{1 - i_0 + i_0 \exp(\beta\langle k\rangle t)}$$

SIS Model:

$$i(t) = (1 - \frac{\mu}{\beta\langle k\rangle})\frac{Ce^{(\beta\langle k\rangle - \mu)t)}}{1 + Ce^{(\beta\langle k\rangle - \mu)t)}}$$

Characteristic time:

SI: $\tau = \frac{\langle k\rangle}{\beta(\langle k^2\rangle - \langle k\rangle)}.$

SIS: $\tau = \frac{\langle k\rangle}{\beta(\langle k^2\rangle - \mu\langle k\rangle)}.$

SIR: $\tau = \frac{\langle k\rangle}{\beta\langle k^2\rangle - (\mu + \beta)\langle k\rangle}.$

Epidemic Threshold:

SIS: $\lambda_c = \frac{\langle k\rangle}{\langle k^2\rangle}$

SIR: $\lambda_c = \frac{1}{\frac{\langle k^2\rangle}{\langle k\rangle} - 1}$

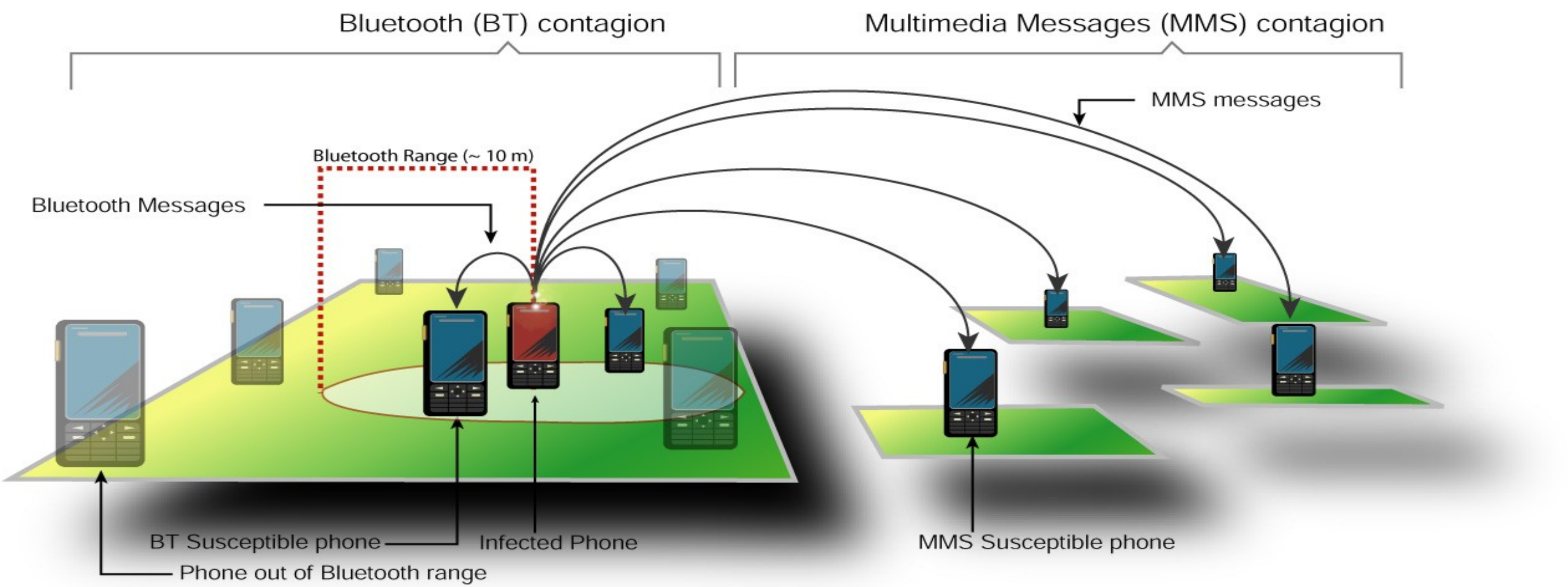
Immunization Threshold:

$$g_c = 1 - \frac{\mu}{\beta}\frac{\langle k\rangle}{\langle k^2\rangle}$$

Case Study 2: Mobile Phone Viruses

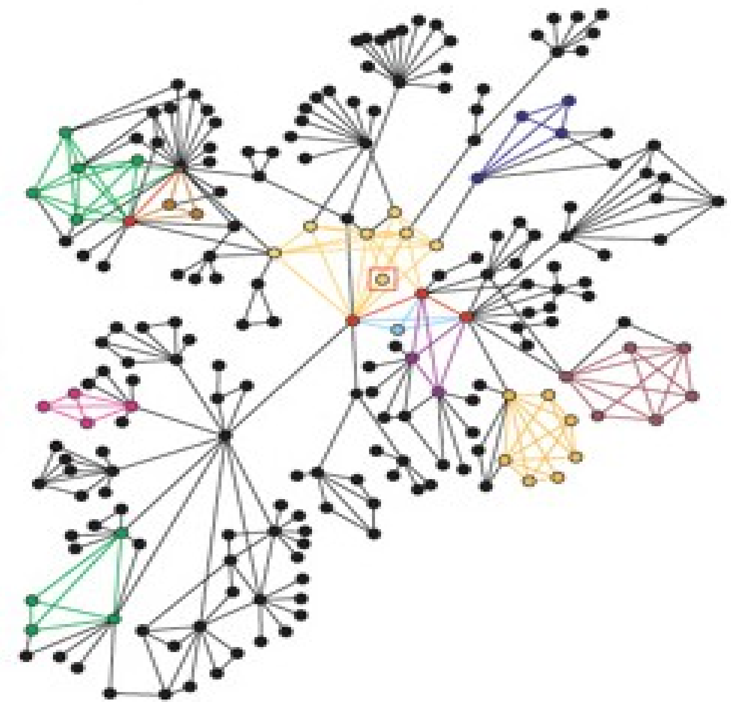
How do Mobile viruses spread?

Bluetooth and MMS viruses

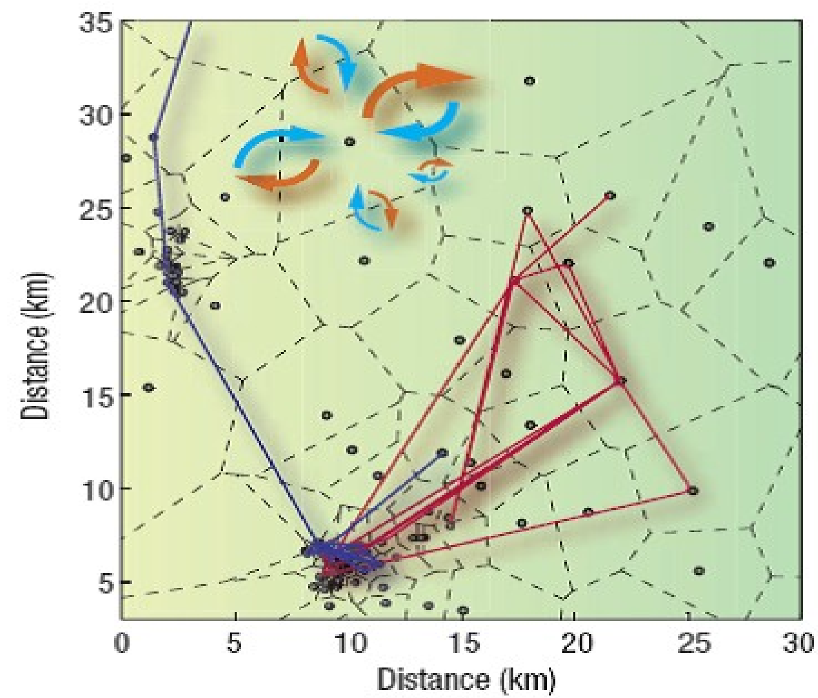


Wang, Gonzalez, Barabasi, *Science*, 2009

MMS and Bluetooth Viruses



Social Network (MMS virus)

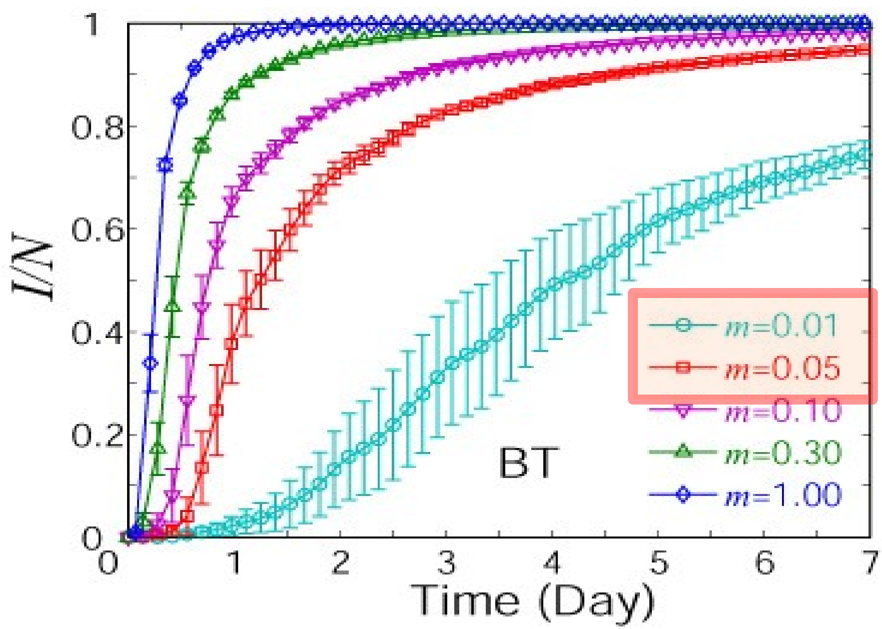


Human Mobility (Bluetooth virus)

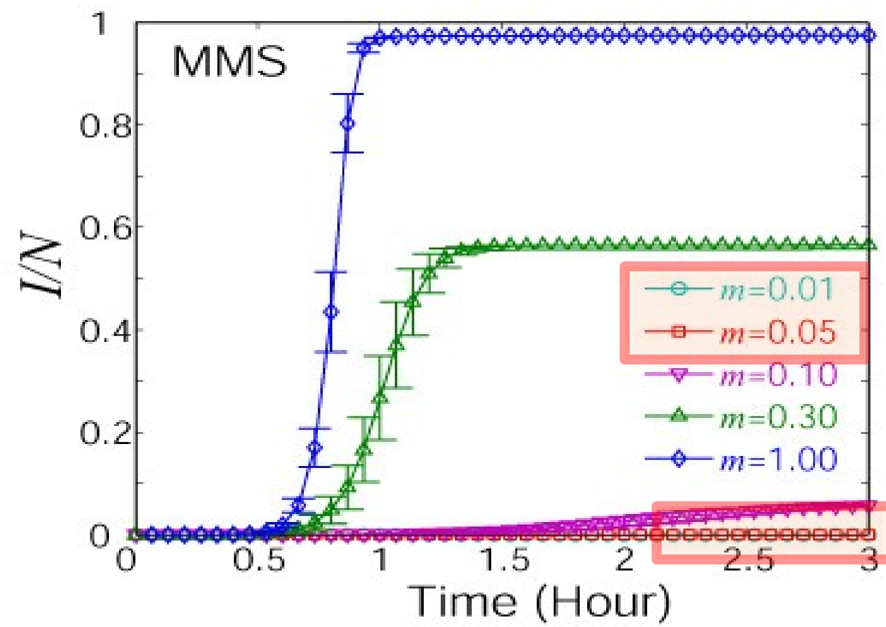
Onella et al, PNAS (2007); Palla et al, *Nature* (2007).
González, Hidalgo and A-L.B., *Nature* 453, *779* (2008)

Spreading Patterns of Bluetooth and MMS viruses

Bluetooth Virus



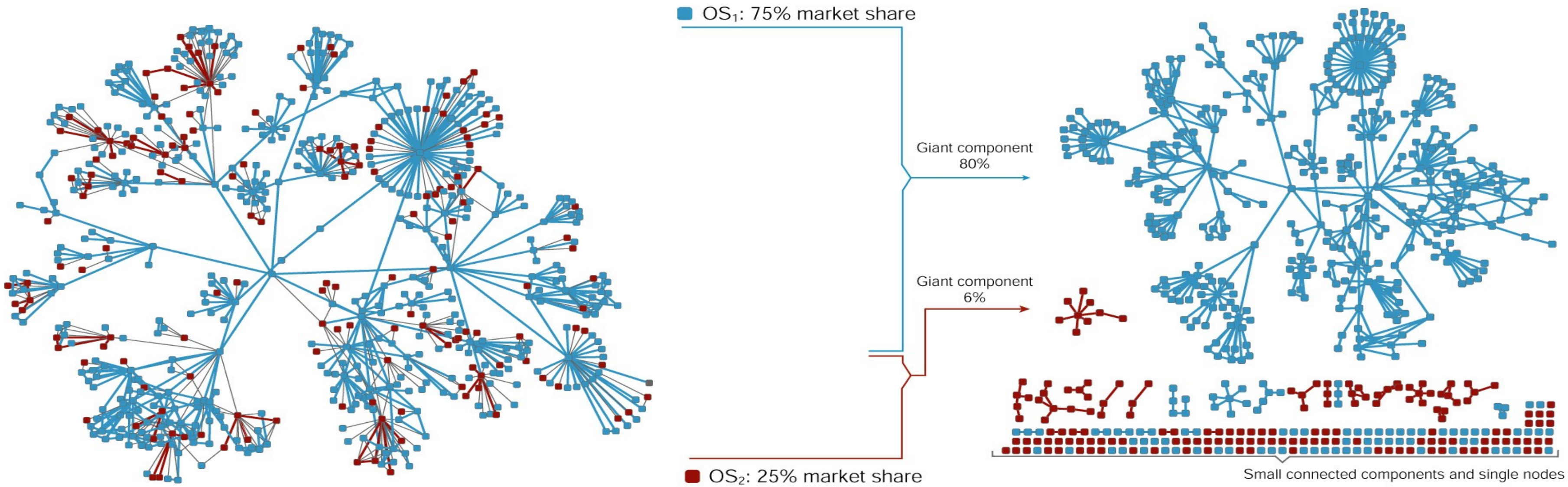
MMS Virus



m: market share of the OS and/or handset the virus can infect.

SmartPhones together *m*=0.05 (5%) of the whole mobile market
Largest OS: Symbian, ~70% of all SmartPhones: m_{max}~0.03

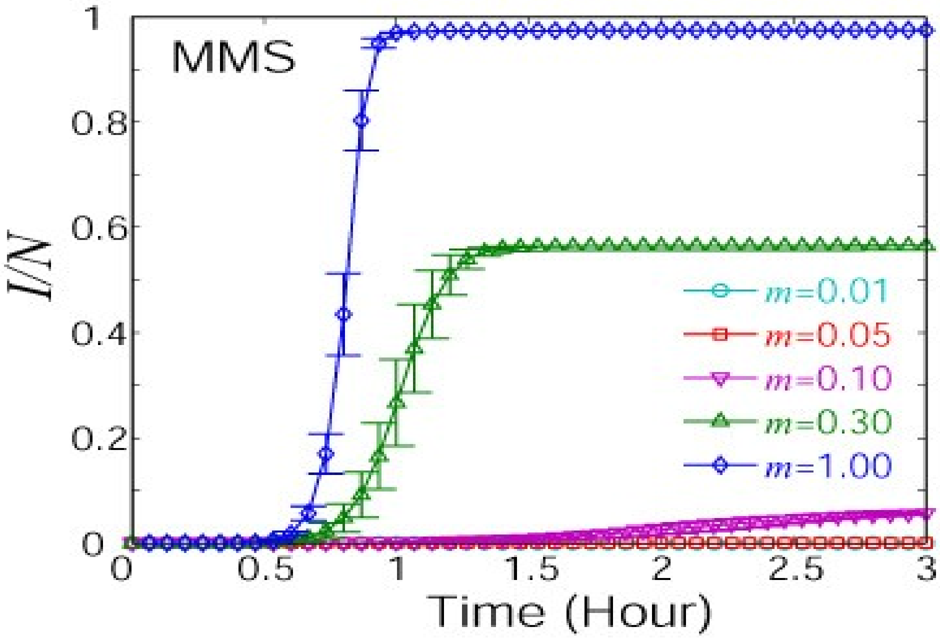
The impact of market share (m) on MMS viruses

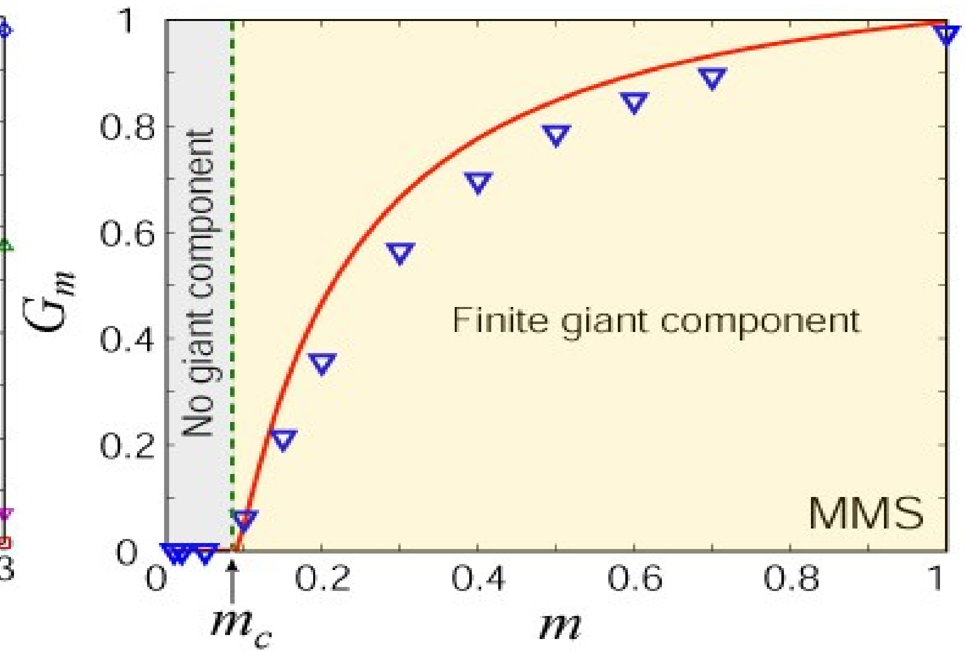


Market share induced fragmentation of the call network.

Wang, Gonzalez, Barabasi, *Science*, 2009

Percolation phase transition limits the spread of MMS viruses





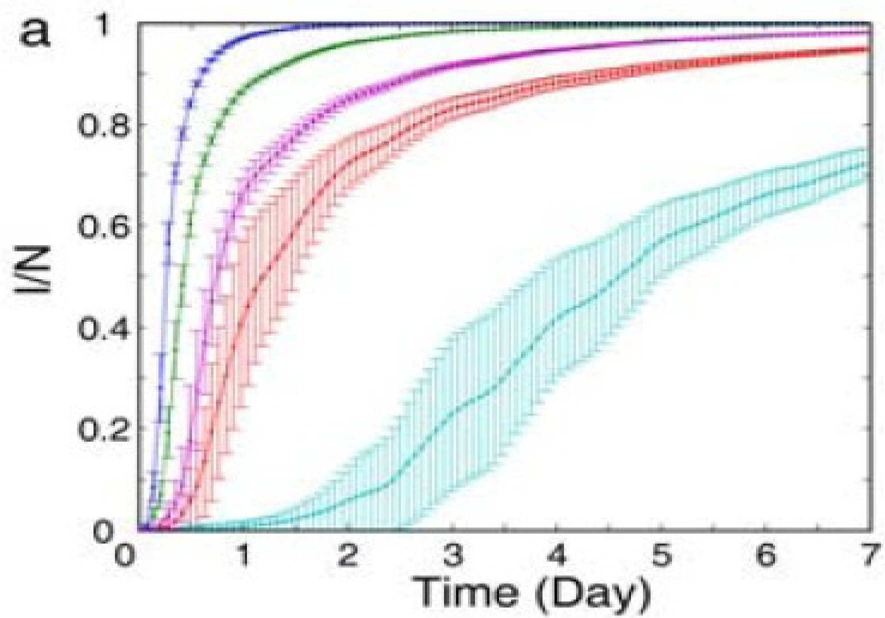
Prediction:

Once the market share of an MMS virus reaches m_c~0.1 (10%), MMS viruses will become a serious concern

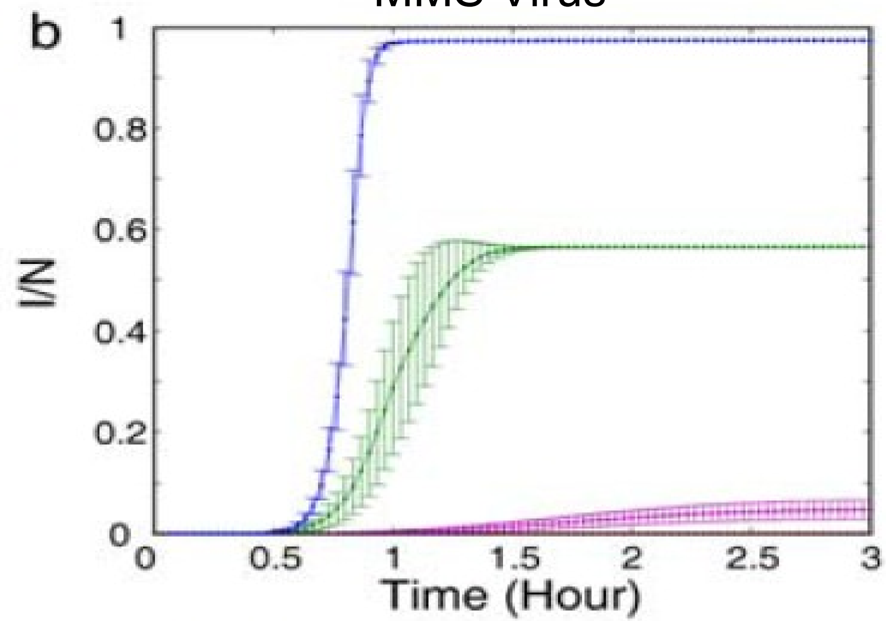
Currently: m_{max}~0.03 <<m_c

Spatial Spreading patterns of Bluetooth and MMS viruses

Bluetooth Virus



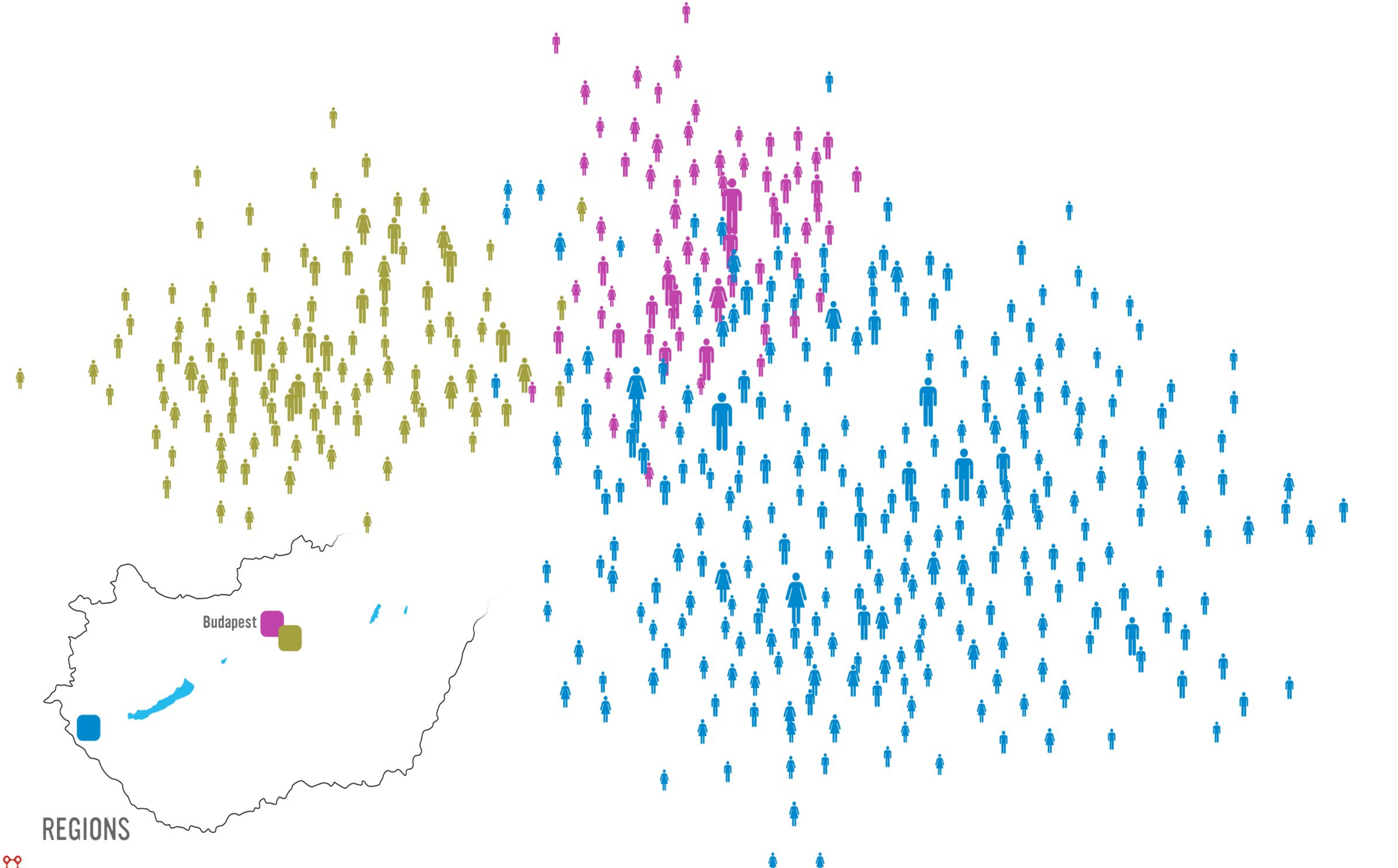
MMS Virus



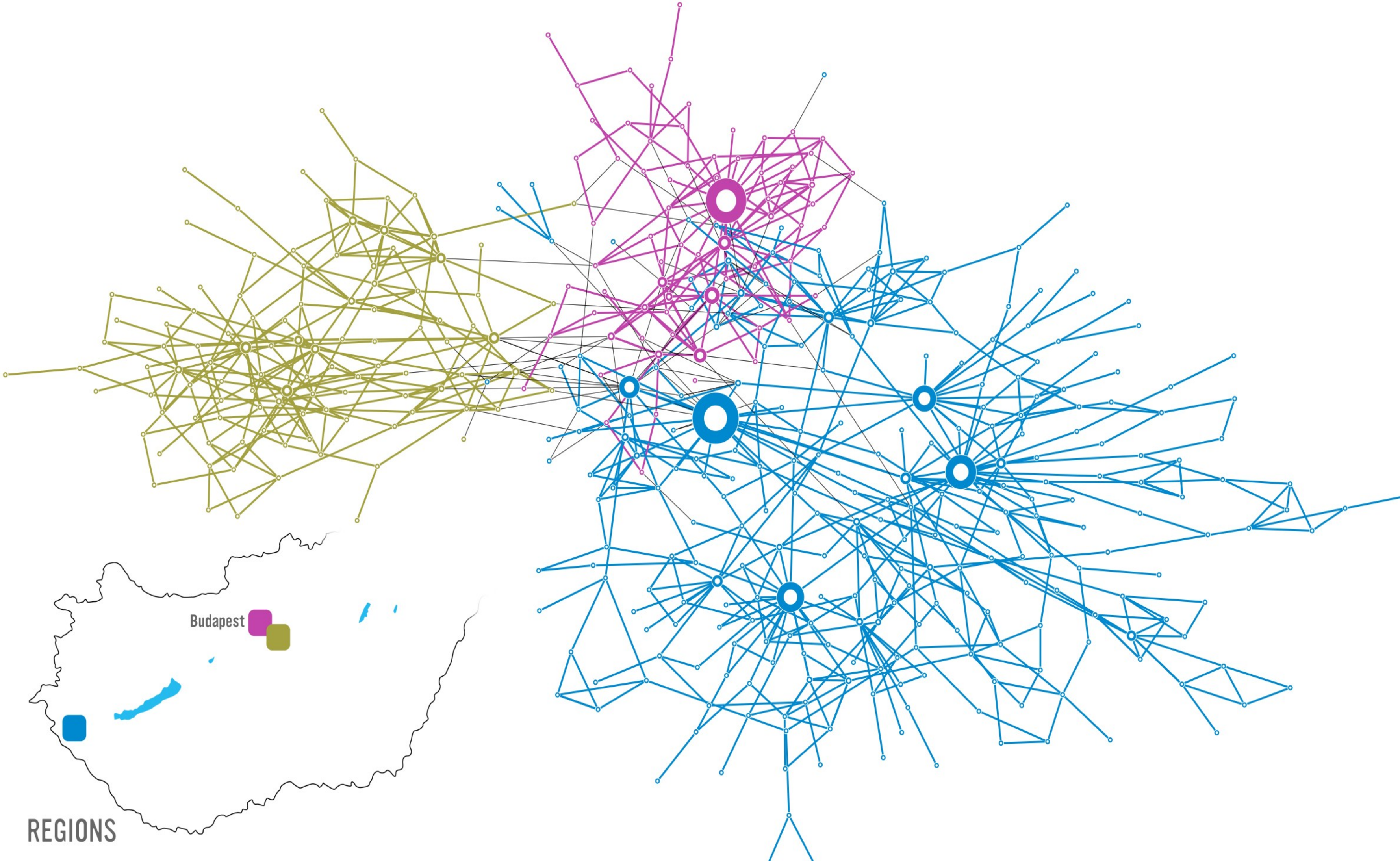
Driven by Human Mobility:
Slow, but can reach all users with time.

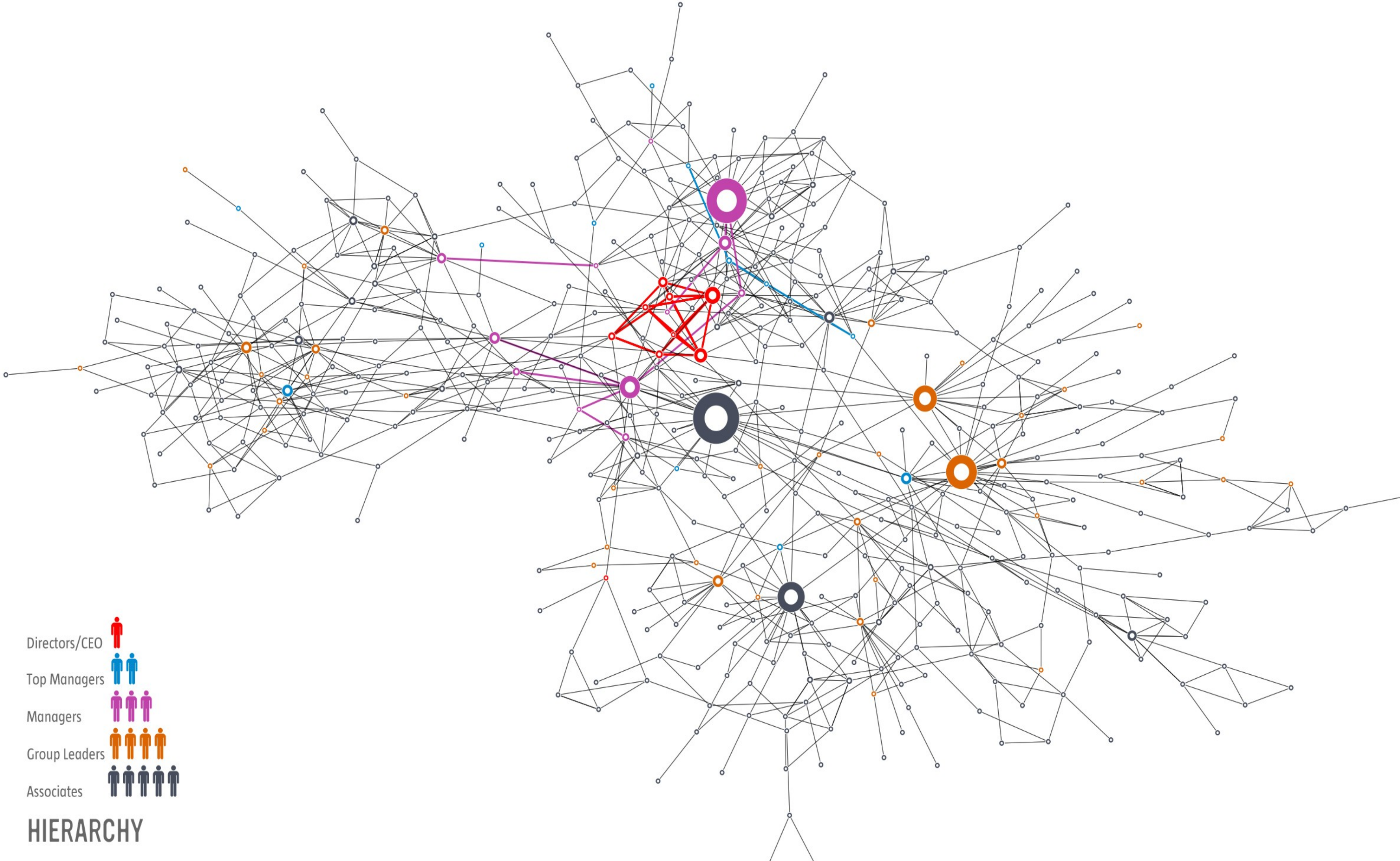
Driven by the Social Network:
Fast, but can reach only a finite fraction of users (the giant component).

Case Study 3: The spread of information at workplace



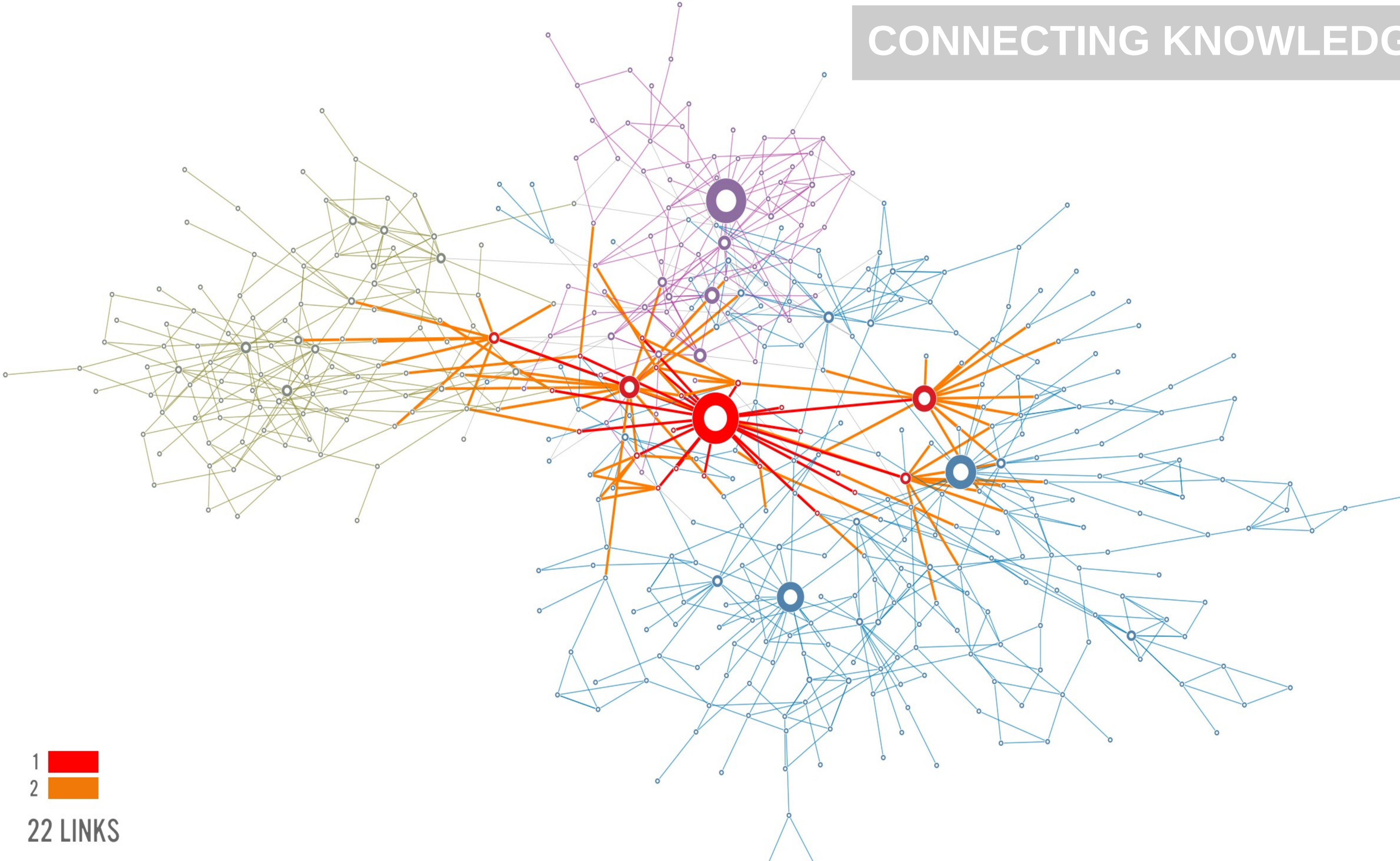
maven7
connecting knowledge


Budapest
REGIONS



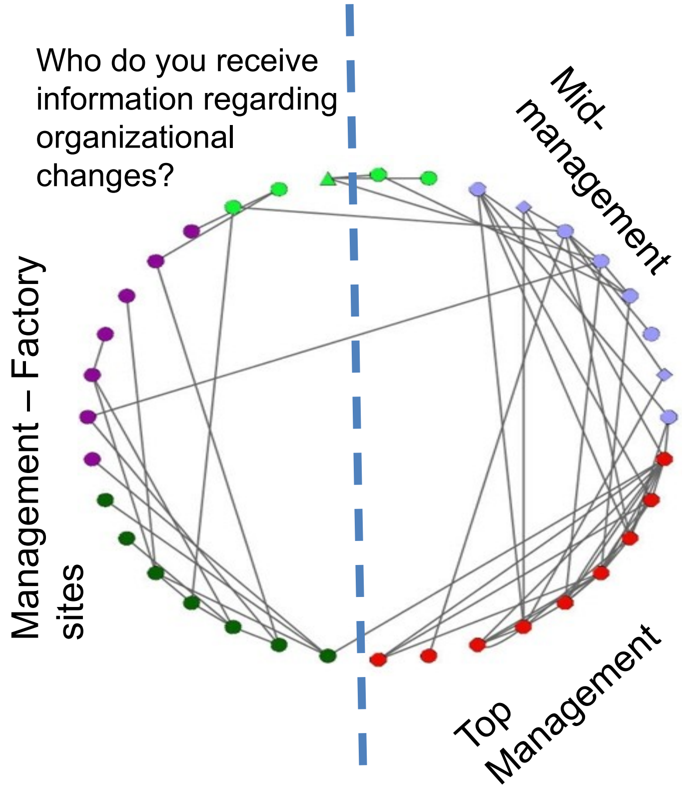
Directors/CEO
Top Managers
Managers
Group Leaders
Associates
HIERARCHY

CONNECTING KNOWLEDGE



Improving Information Flow

Who do you receive information regarding organizational changes?



Links are indicating information flow between individuals about organizational changes.

Easy-to-recognize gap between management levels

Manufacturing company with about 800 employees

<u>Issues:</u>

(1)Information gaps and gossip about organizational changes;

(2) Strategic decisions miss-understood;

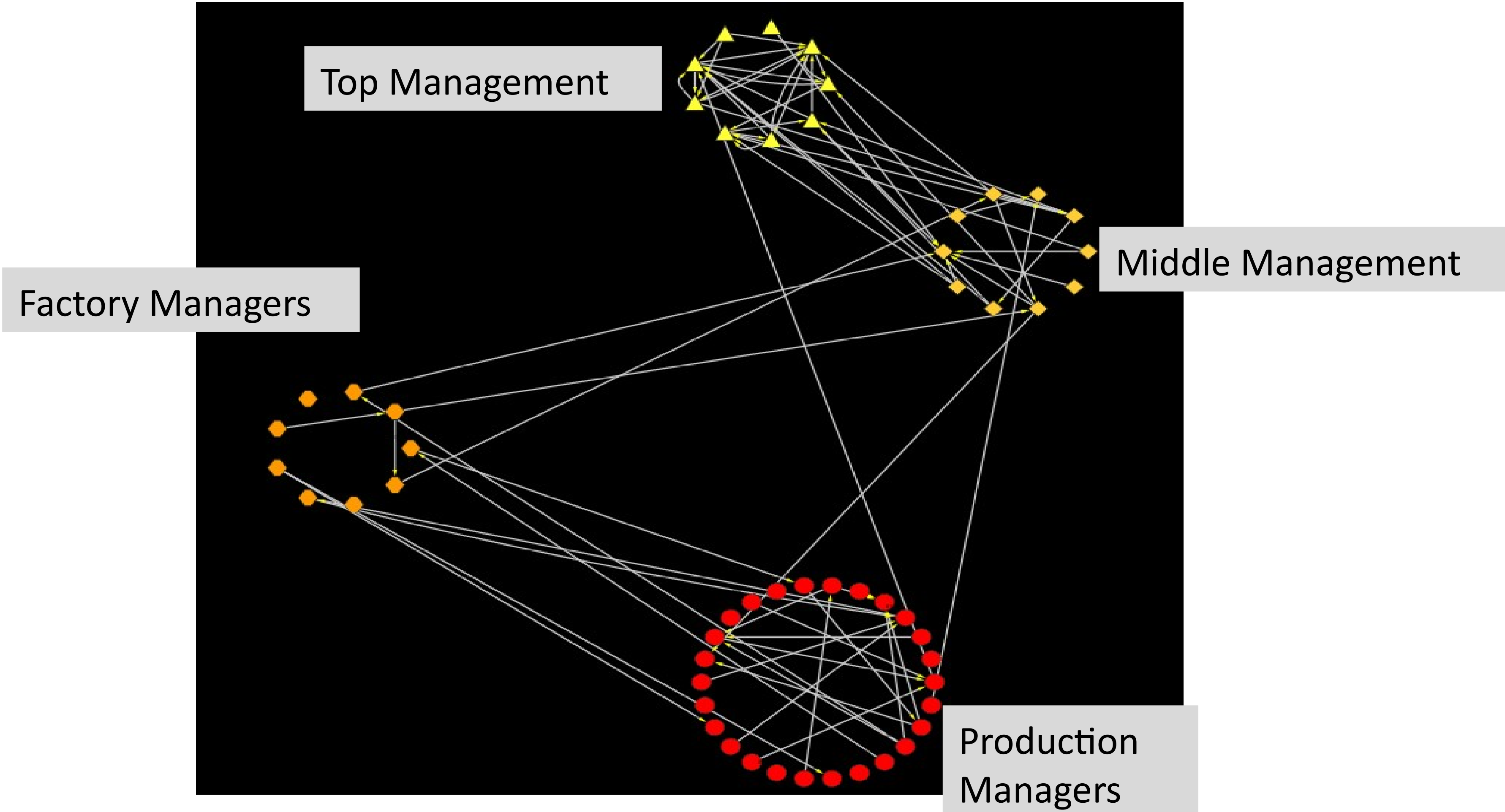
(3) Lack of trust in management

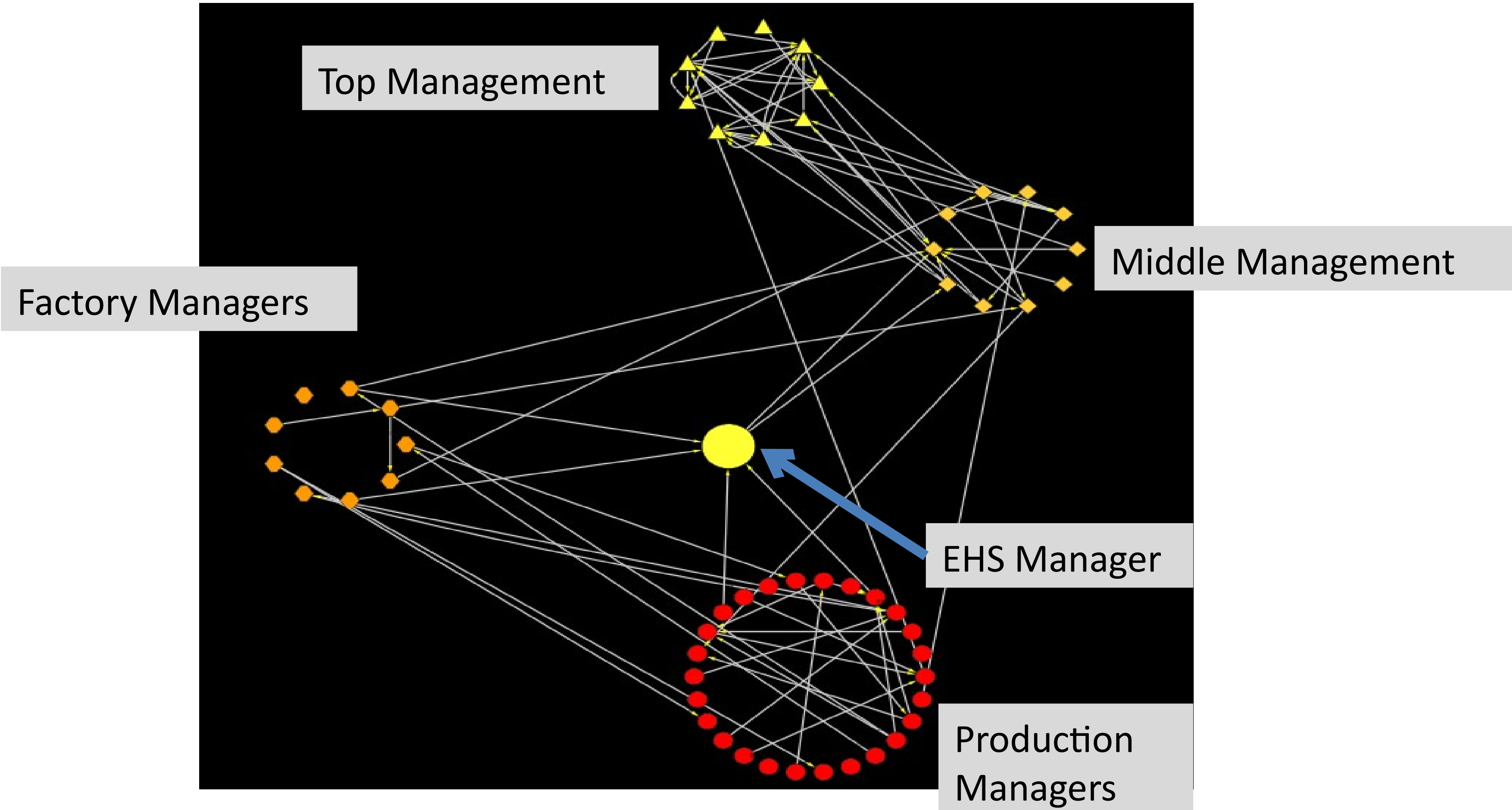
<u>Aim:</u>

Reduce time for accepting changes;

Gossip management;

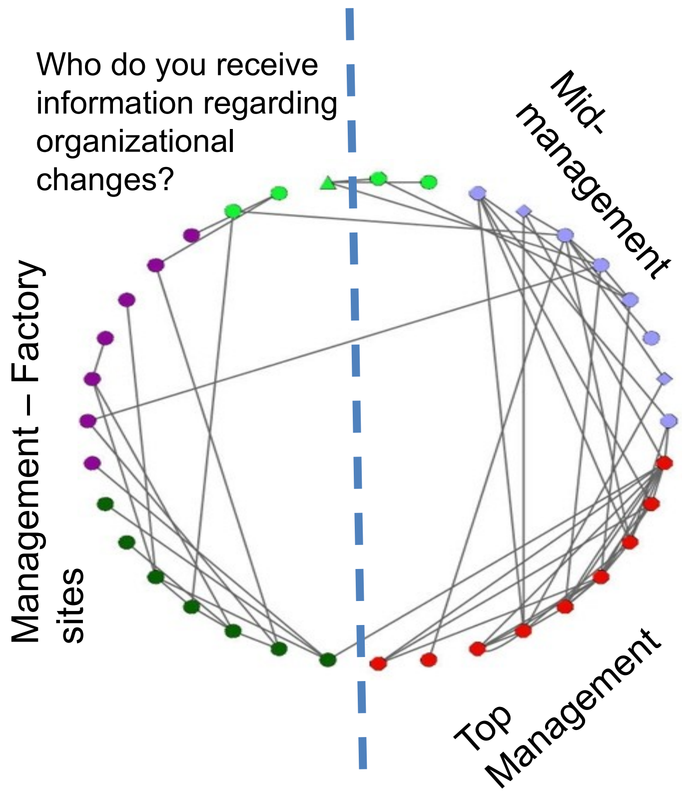
Build trust


Top Management
Middle Management
Factory Managers
Production Managers


Top Management
Middle Management
Factory Managers
EHS Manager
Production
Managers

Improving Information Flow

Who do you receive information regarding organizational changes?



Links are indicating information flow between individuals about organizational changes.

Easy-to-recognize gap between management levels

Manufacturing company with about 800 employees

Issues: (1) Information gaps and gossip about organizational changes; (2) Strategic decisions miss-understood; (3) Lack of trust in management

Aim: Reduce time for accepting changes; Gossip management; Build trust

Findings:

Robust communication between mid and senior management *BUT* Lack of information flow between mid-management and management of manufacturing sites.

Main source of information for Factory Management: EHS Manager – no connection to management, no career plan and frustrated about own possibilities

The end