2015 MANEM

Generating Multidimensional Social Network to Simulate the Dissemination of Information

Mathilde Forestier, Jean-Yves Bergier, Youssef Bouanan, Judicael Ribault,

Gregory Zacharewicz, Bruno Vallespir, Colette Faucher

Univ. Bordeaux, IMS, UMR 5218, F-33400 Talence, France.
Univ. Aix-Marseille, LSIS, UMR 7296, F-13000 Marseille, France.

SICOMORES masa ims rapid DGA université de BORDEAUX

The SICOMORES project

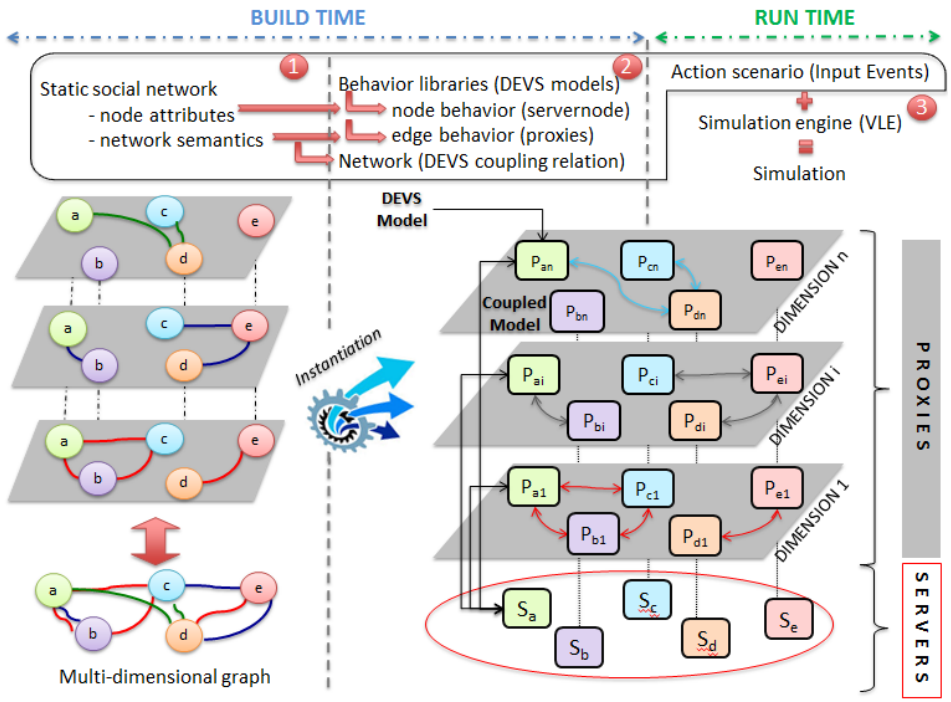
- Project funded by the French DGA (Directorates General of Armaments)
- Aims to
  - Generate a population with
    - cultural features
    - Several relationships between individuals
  - Simulate the diffusion of Psychological Operations (PSYOPS) inside this population

MAIN GOAL :

Train the military in choosing the best actions to obtain a predefined goal

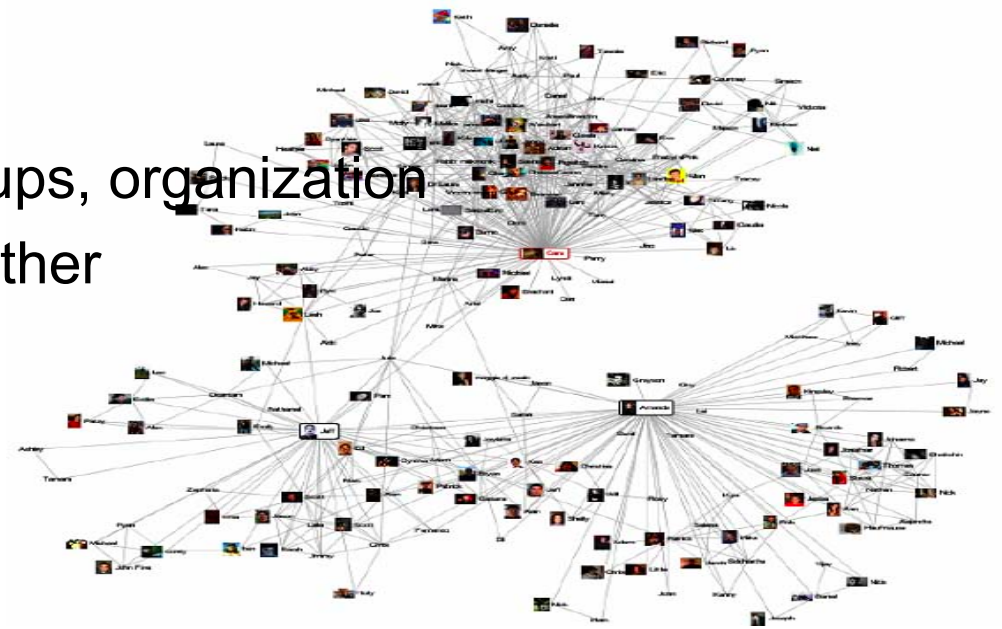
What we want to do ?

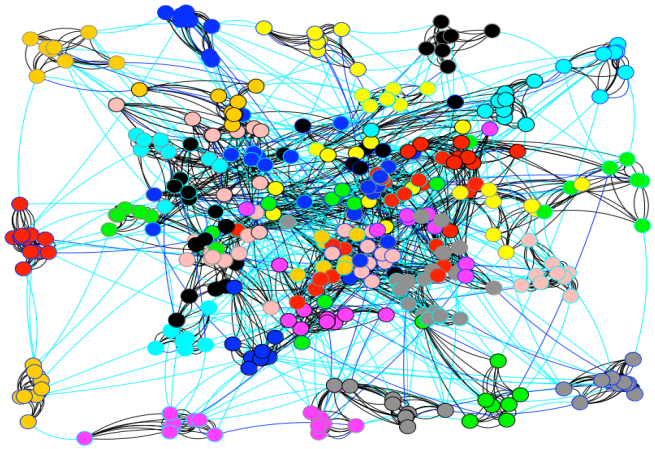
- Generate a population with cultural features using an MSN
- Give an information to several nodes (info-sources)
- Simulate the diffusion of information inside this population starting with the info-sources.



What is a Social Network ?

- A graph representation
  - Nodes V: individuals, groups, organization
  - Edges E : link nodes together
    - Friendship
    - Family
    - Co-authorship
    - Coworkers
    - Etc…



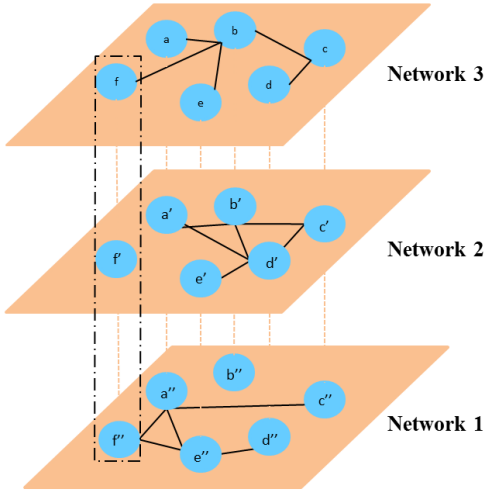


What is a Multidimensional Social Network ?

[Belingerio13]

➔ A Social Network composed by layers/dimensions of social networks : G = (V,E,L)

- V: set of nodes;
- L: set of labels / dimensions
- E: a set of labeled edges
  - Set of triples (u,v,d) where $u,v \in V$ and $d \in L$



Actually, relationships between people are too complex to be modeled by one link, e.g., in real life, people can be friends, kin, neighbors, and so on.

Part 1

Multidimensional social network generation

université de BORDEAUX

Building the multidimensional social network

➔ What is a population ?

- A set of individuals with features
- A set of several relationships
- A set of cultural features: how people link together ?

Composition of the population based on its ethnic groups

➔ Composition of a population in our modelisation

› The ethnic groups are defined by :

- A name
- A religion
- A language
- Its proportion inside the whole population
- the proportion of each social level inside the ethnic group
- Their needs in security, heath care and food
- The opinion about the military

What is an individual ?

A node with attributes :

- Sex
- Age
- Religion
- Ethnicity
- Language
- Social level
- Role inside the family
- Values
- Norms

- Illiteracy
- Reachable by radio
- Reachable by TV
- Food needs
- Security needs
- Healthcare needs
- Opinion

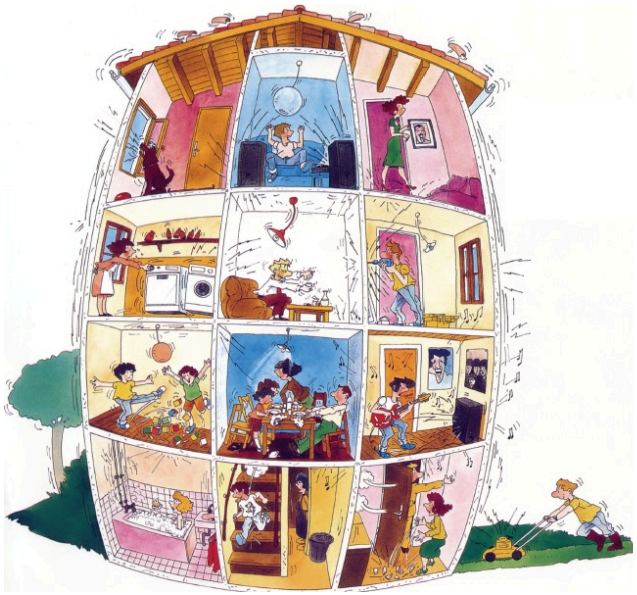
Dynamic variables

The three primary dimensions

[Cooley09]

FAMILY

FRIENDS

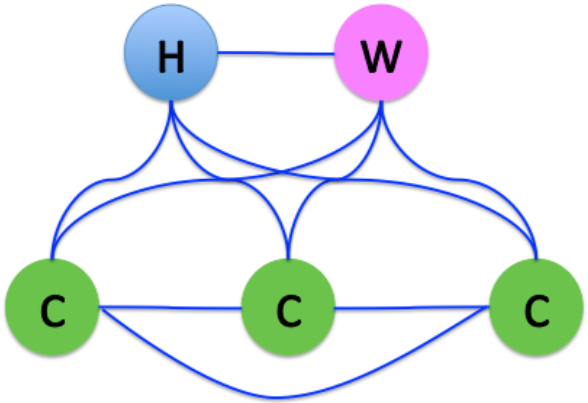


NEIGHBORS

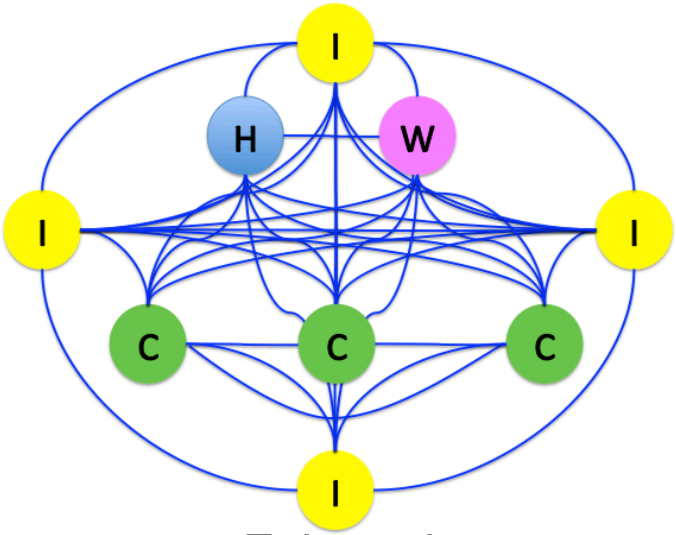
➔ Represent the three first socialization structures of human life

What is a family ?

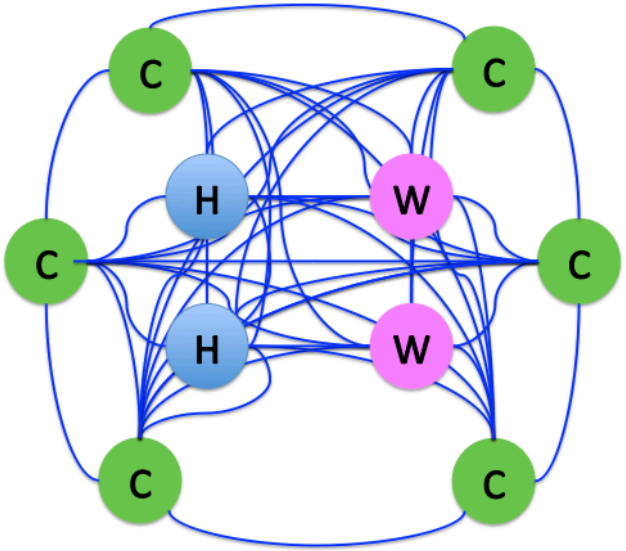
➔ We set three family structures based on social science theories



Nuclear



Enlarged



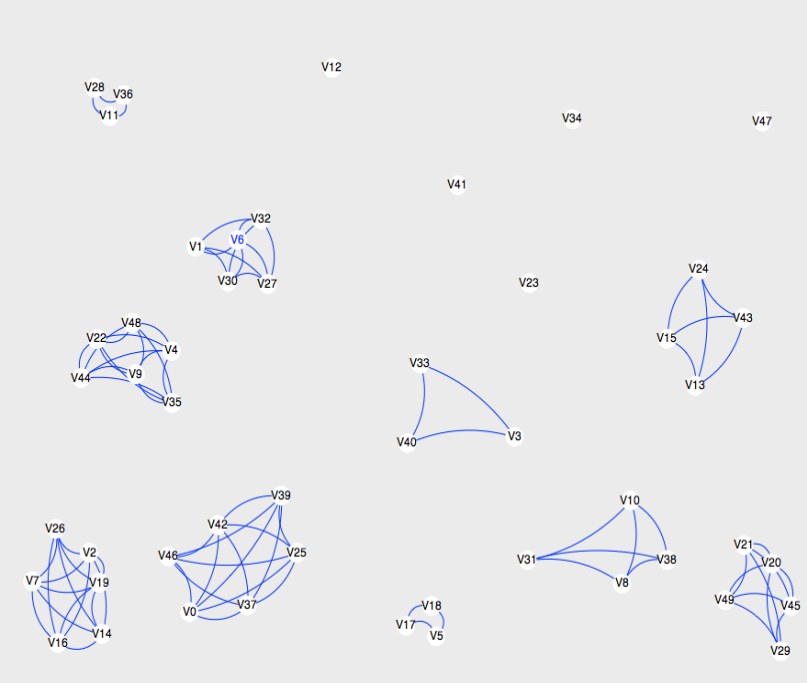
Extended

Family links generation

Inputs
- % of each kind of family in the city
- % of lonely people
- % of polygamist families
- % of matriarchal families

Define for the family
- an ethnicity
- a religion,
- a language,
- how the family is reachable (TV, radio, etc.),
- the needs,
- Its type (nuclear, extended, or enlarged)

Build a family:
- ➢ Assign to nodes the features defining for the family,
- ➢ Assign to nodes specific features such as a sex, an age, the role inside the family
- ➢ Generate a clique between all nodes inside the family



Friendship link generation

- Birds of a feather flock together
  - friendship is based on the concept of homophily

$$\text{friendshipHomophily} = w_s * \text{sexe} + w_a * \text{age} + w_{sc} * \text{socialLevel} + w_e * \text{ethnicity} + w_l * \text{language} + w_r * \text{religion}$$

Friendship Link generation (1/2)

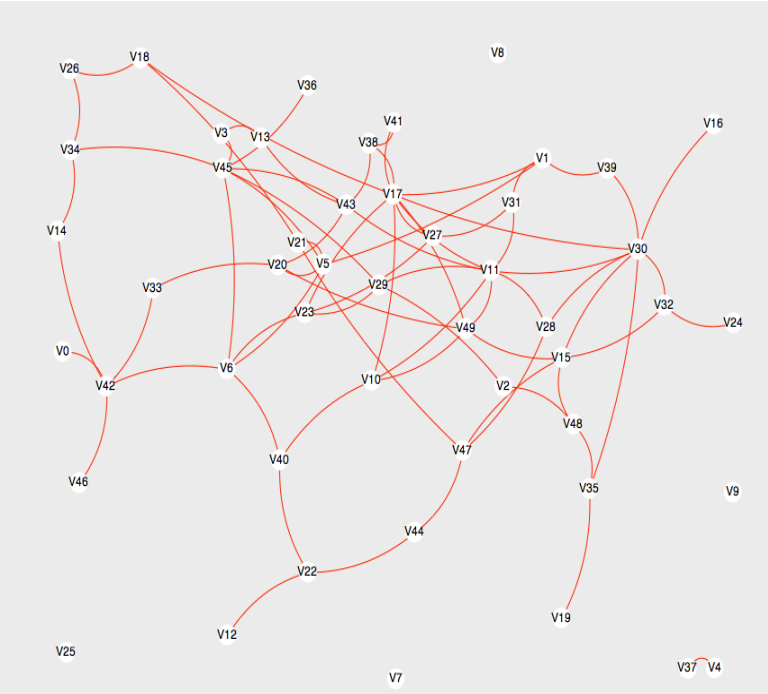
Inputs
- Average # of friends
- Threshold of homophily
- Weights for each features

Pick two nodes randomly

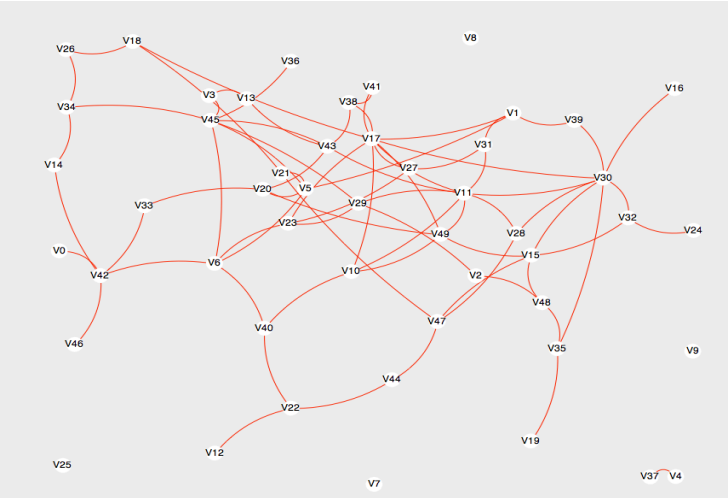
Calculate the homophily between the two picked nodes

If (homophily > threshold)
Then
    Generate a link between the two nodes
End if

Repeat until obtaining the average/2 number of friends



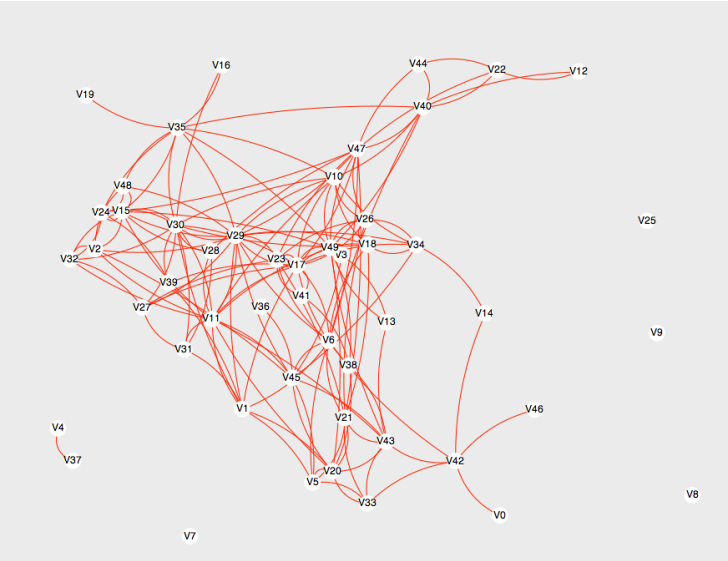
Friendship Link generation (2/2)



Pick one node randomly

Retrieve its ego-network in the friendship graph dimension

Generate links between nodes inside the ego-network



Repeat until obtaining the average number of friends

Neighborhood link generation

➔ The neighborhood dimension is also generated on the concept of homophily

$$\text{neighborHomophily} = w_{sc}*\text{socialLevel} + w_e*\text{ethnicity} + w_l*\text{language} + w_r*\text{religion}$$

Neighborhood link generation

Inputs
- Average # of neighbors
- Threshold of homophily
- Weights for each features

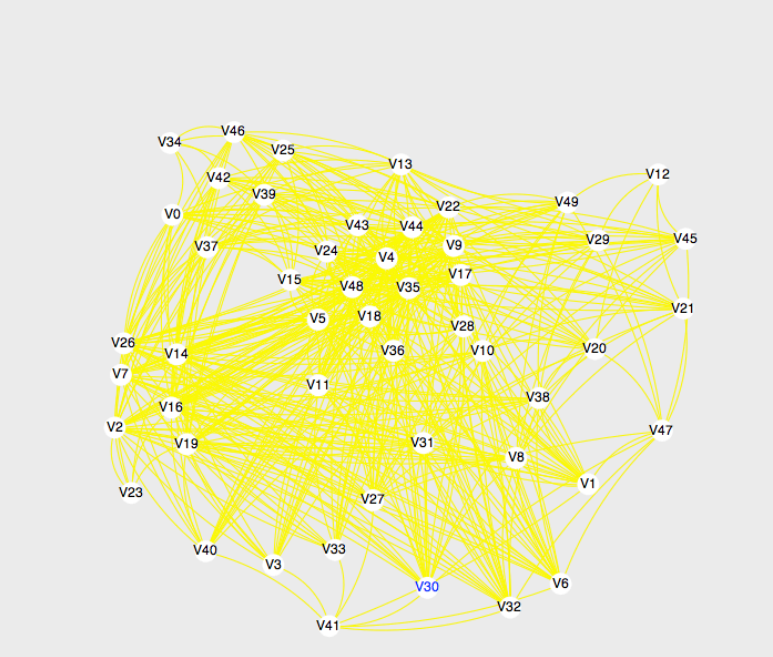
Pick two nodes randomly

Calculate the homophily between the two picked nodes

If (homophily > threshold) Then
- Generate a link between the two nodes
- Generate links between all members of each family

End if

Repeat until obtaining the average number of neighbors



The war time dimension

- During a war, people can group differently than as in peace time
- We define two ways to group :
  - According to the ethnicity (e.g., the Rwanda genocide)
  - According to the religion (e.g., the Central African Republic civil war)
- This dimension should be activated during a situation of chaos defined during the simulation by a cohesion social threshold

The war time dimension generation

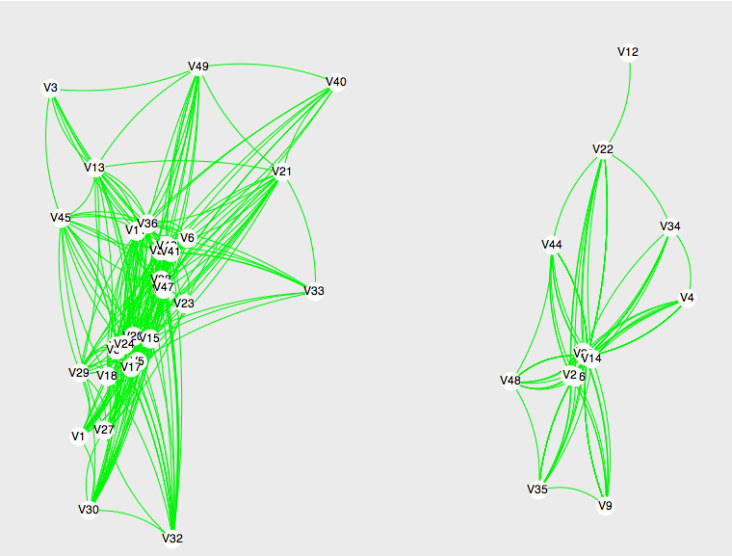
Inputs
- Average # of relations
- Segregation choice

Pick one node randomly

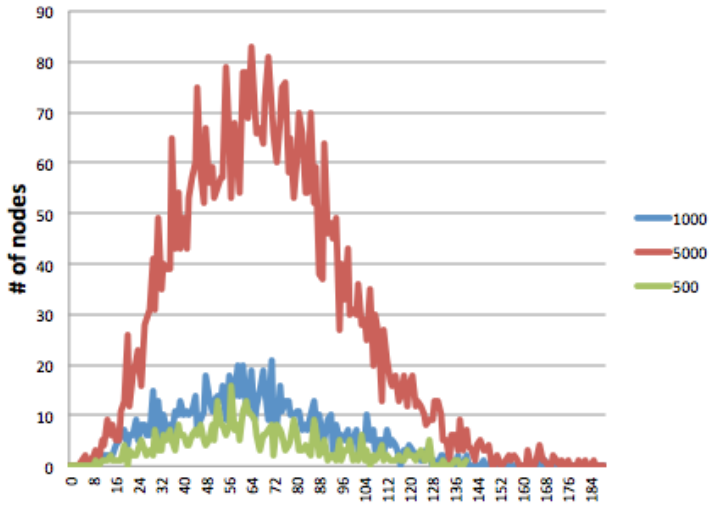
Retrieve its ego-network from friend and neighborhood dimensions

Generate the links between the picked node and its friends and neighbors sharing the same religion or ethnicity

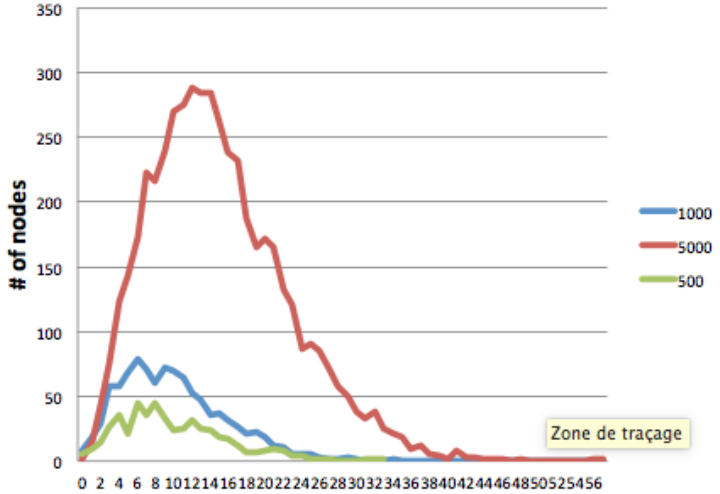
Improve the coefficient clustering in adding links between the previously created graph



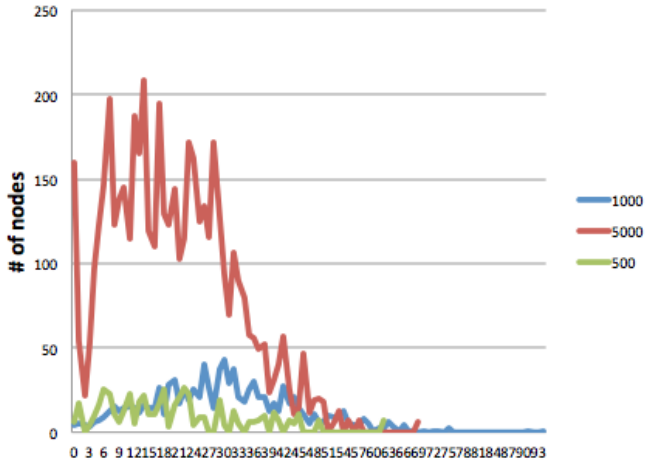
Results (1/2)



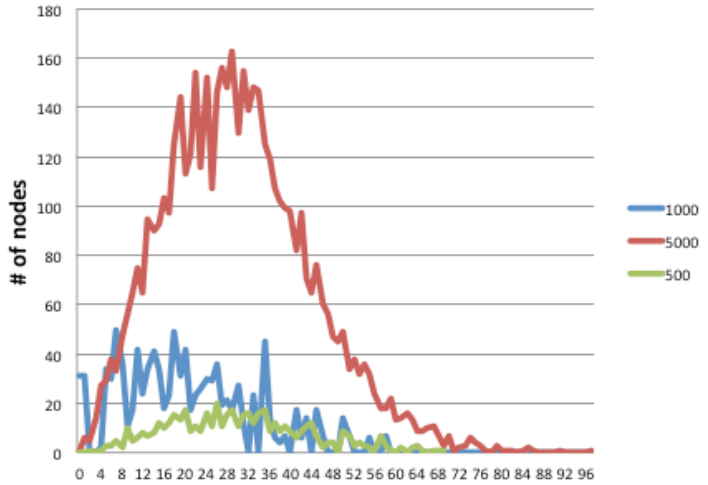
MSN



Friendship dimension



Neighborhood dimension



Religious dimension

Results (2/2)

# of people	Population #1 500	Population #2 1000	Population #3 5000
average # of degree	64	64	69
min	9	6	4
max	140	168	190
	average of shortest past between two nodes		
MSN	2.09	2.33	2.76
family dimension	1	1	1
friendship dimension (1)	4	4.47	4.59
friendship dimension (2)	3.08	3.83	3.54
neighbor dimension	3.4	4.6	5.17
religious dimension (1)	2.3	2.6	2.95
religious dimension (2)	2	2.27	2.78

(1) Before improving the clustering coefficient
(2) After improving the clustering coefficient

Conclusion about the MSN generation

➔ Modify the algorithms to obtain a power law distribution degree

› Use less randomly chosen nodes to build the dimensions

➔ Add other dimensions

› e.g., the religious dimension

› Temporary dimensions such as a market time ➔ new links for a faster information propagation

Part 2

SIMULATION

université de BORDEAUX

Choices to the simulation architecture

➔ Shared node across networks

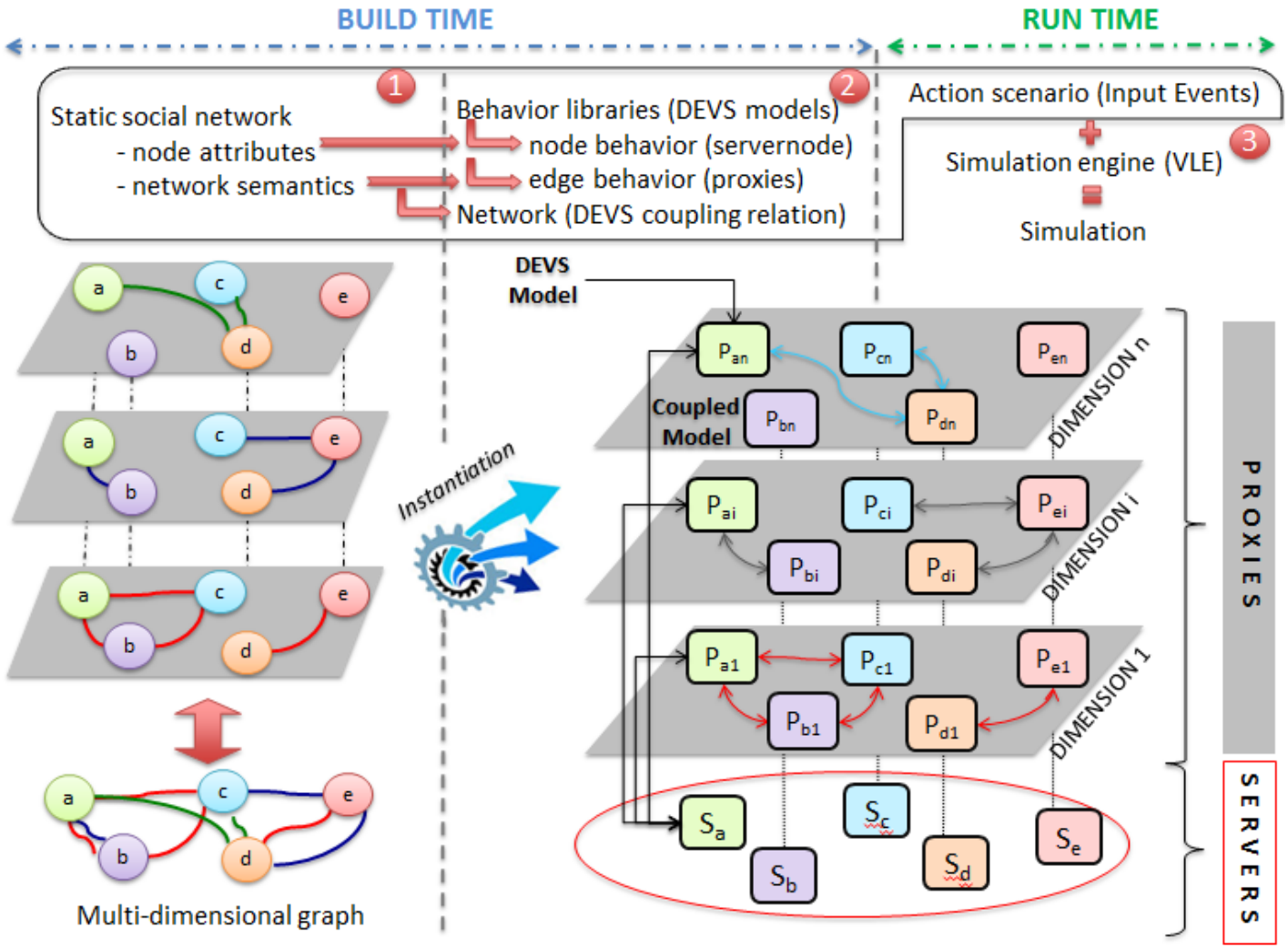
› Solution 1: flatten dimensions into one social network

- One agent with constraints.
  - Hard to implement and to reuse: all the code are in one place (the individual).

› Solution 2: use a Server/Proxy architecture

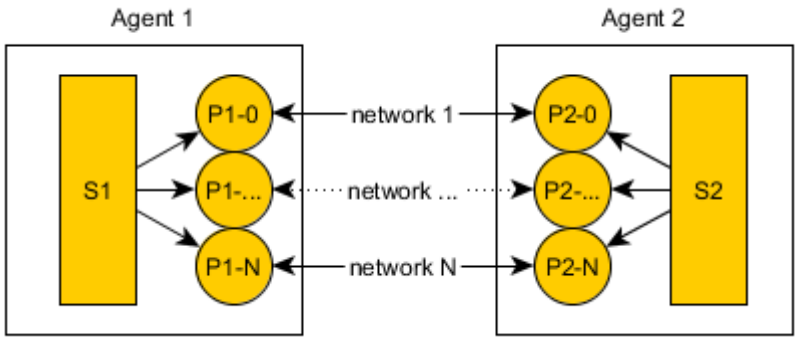
- One Server per node.
- One Proxy per node dimension.
  - N Server (N=size of the population under study)
  - M Proxy (M=N*P, P=number of dimensions)

The general architecture



Chosen architecture

- Shared node across networks
  - The code is split between the Server and the Proxy
    - Better separation of concerns
    - Each network can have their own acceptance and propagation rules
    - The Server just maintain the state of the individual



The Discrete Event System Specification

[Zeigler76]

➔ We use the DEVS formalism

› formalism for modeling and analysis of discrete event systems.

› Low level formalism

- Allow to define all constraints we need
- Allow to have the full powers on what is implemented
- Allow to define libraries for reusability of the code

Diffusion Models

- First mathematical models
  - [Schelling70/78, Granovetter78]
- Large body of subsequent work:
  - [Rogers95, Valente95, Wasserman/Faust94]
- Two basic classes of diffusion models:
  - The linear threshold model
  - The independent cascade model
- General operational view:
  - A social network is represented as a directed graph
  - Nodes start either active or inactive
  - An active node may trigger activation of neighboring nodes
  - Monotonicity assumption: active nodes never deactivate

Linear Threshold Model

[Granovetter78]

- A node v has random threshold $\theta_v \sim U[0,1]$
- A node v is influenced by each neighbor w according to a *weight* $b_{v,w}$ such that

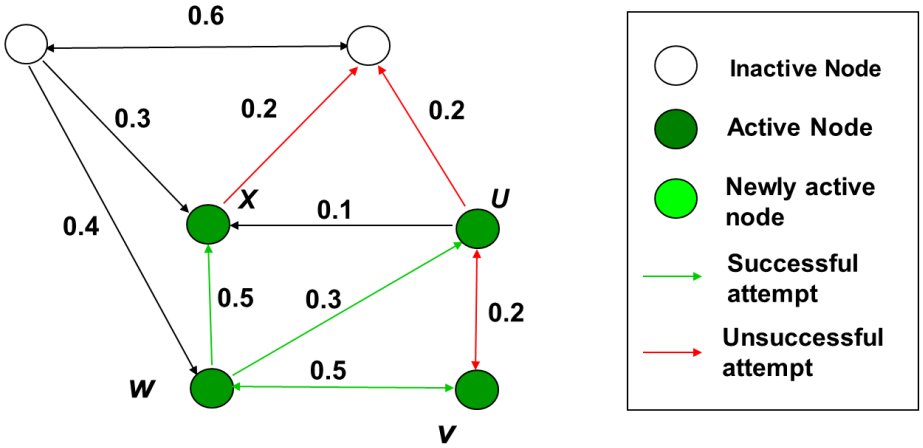
$$\sum_{w \text{ neighbor of } v} b_{v,w} \leq 1$$

- A node v becomes active when at least (weighted) θ_v fraction of its neighbors are active

$$\sum_{w \text{ active neighbor of } v} b_{v,w} \geq \theta_v$$

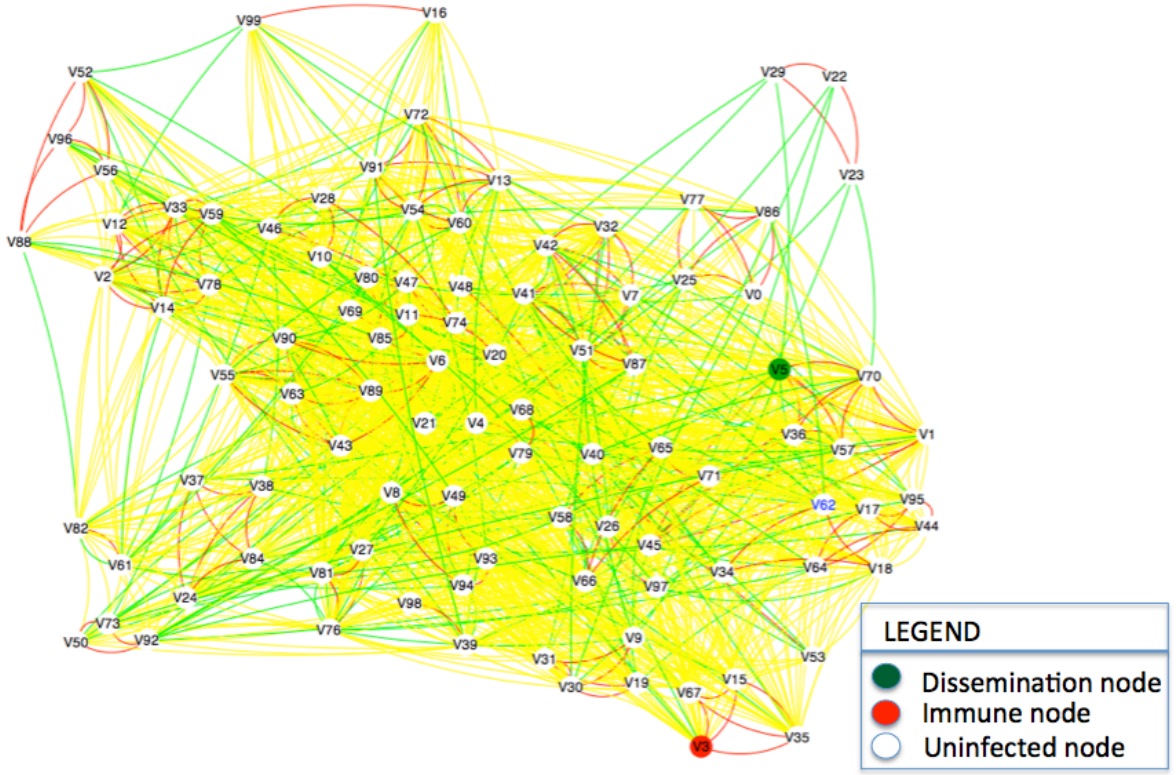
Independent Cascade Model

- When node *v* becomes active, it has a single chance of activating each currently inactive neighbor *w.*
- The activation attempt succeeds with probability p_{vw} .

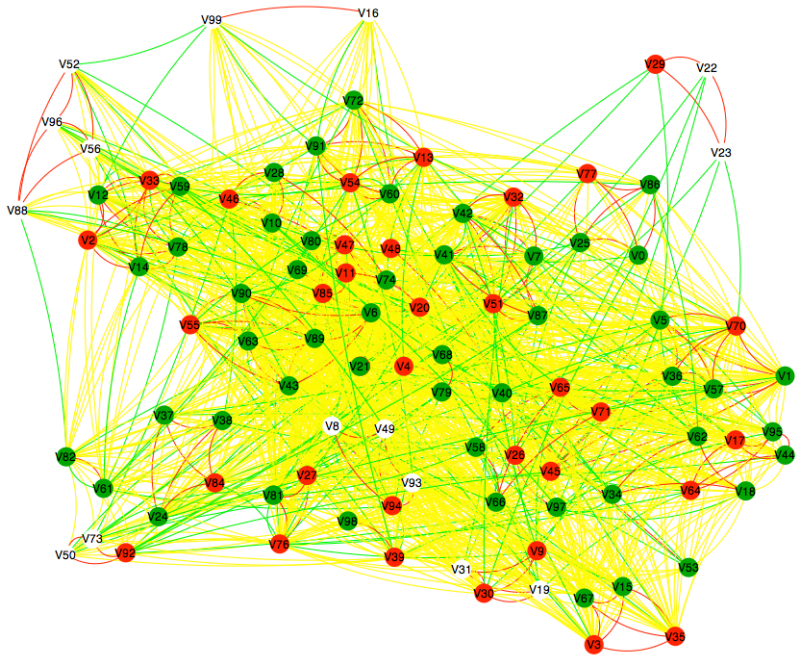


Experiments

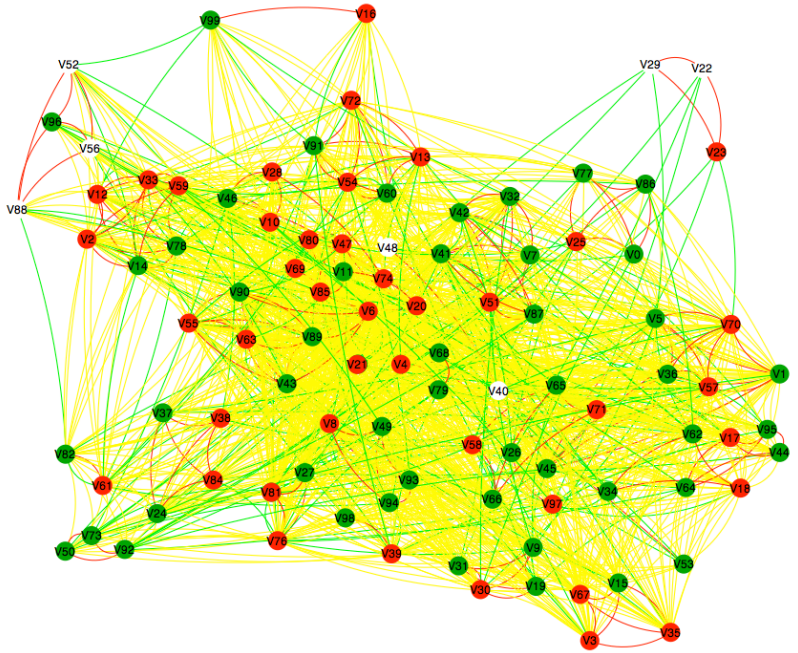
DIMENSION	PROBABILITIES	MESSAGE STRENGTH
Family	0.9 security O.9 health care	message strength -1
Friendship	O.4 security 0.6 health care	message strength -2
Neighbor	O.1 security 0.4 health care	message strength -3



Results of the simulations



Result of the simulation with a message about security



Result of the simulation with a message about health care

Conclusion of simulation

➔ Our architecture allows to:

- › Reuse parts of the model for other studies
- › Add new concern easily
- › Keep code simple
- › Improve the VV&A process
  - Once a Server or a Proxy has been validated, we don't need to modify it to add a new network (with its own specific rules).
    - – We just have to create a new proxy

➔ The information diffusion will be improved with social science studies to better model the human behavior

General conclusion and perspective

➔ Human behavior is complex to simulate

- Using an MSN allows to separate the diffusion ways
- Using a proxy/server architecture allows to model an MSN with several rules for each node and for each relation.

➔ This work can be used in plenty other fields such as

- In marketing to simulate the adoption of a new product
- In politics to simulate the diffusion of an idea or the way that a politician's reputation change

THANK YOU FOR YOUR ATTENTION

Mathilde Forestier
mathilde.forestier@univ-bordeaux.fr