# Models based on movement networks
Italian experiences

Types of models

- Classical (SIR) -> how large / when
  - Randomness of contacts

- Spatial -> how large / when / where (area)
  - Contacts based on geographical proximity
  - Source of contact incorporated implicitly and statically in the geography

- Network based -> how large / when / where (holding)
  - Contact are represented explicitly on the basis of real data (e.g. livestock movements)
  - Contacts can be considered dynamically in time
  - Do not consider other sources of contacts other than the ones explicitly expressed by the data

- Granularity of the model
  - One population for the area of interest
  - Contacts between holdings
  - Meta-population: contacts between animals at the level of sub-population and between holdings at the level of population
  - Agent based: the single animal

Network definition

- Holdings can be represented as nodes
- Movements of batches of animals can be represented as edges
- Edges have a weight (no. of animals transported) and a date attribute
- Edges represent contacts between holdings along which diseases may spread
- Nodes role in a network can be expressed with several centrality measures

Networks and epidemic models
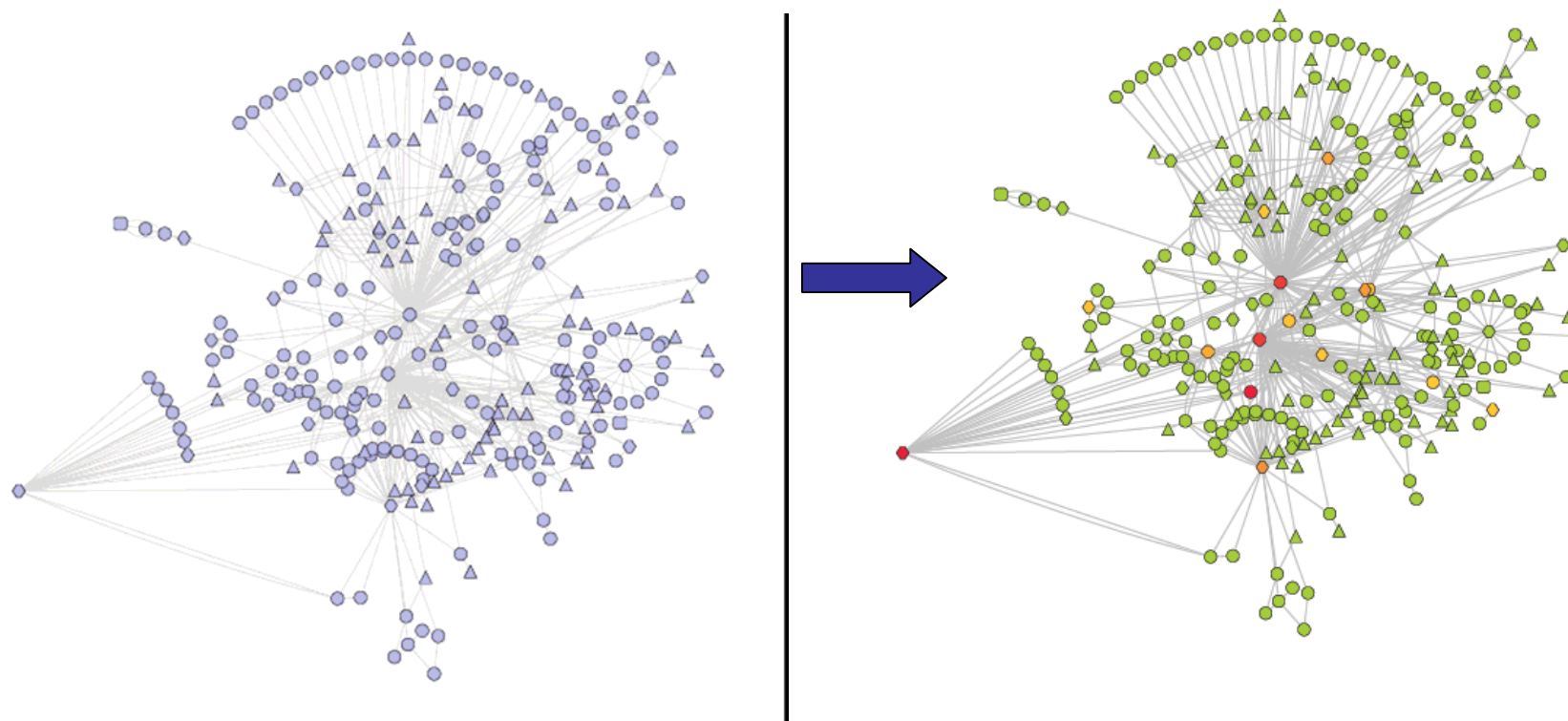
- Risk characterization of holdings based on trade data (preventive)
- Prioritization of field investigation (during the outbreak)

Network based SIR

Meta-population SIR

A subpopulation epidemic is started in the source at a given time and with a given level of infection

At the date of a movement, is the level of infection in the source high enough to produce infected individuals in the transport component above a given threshold?
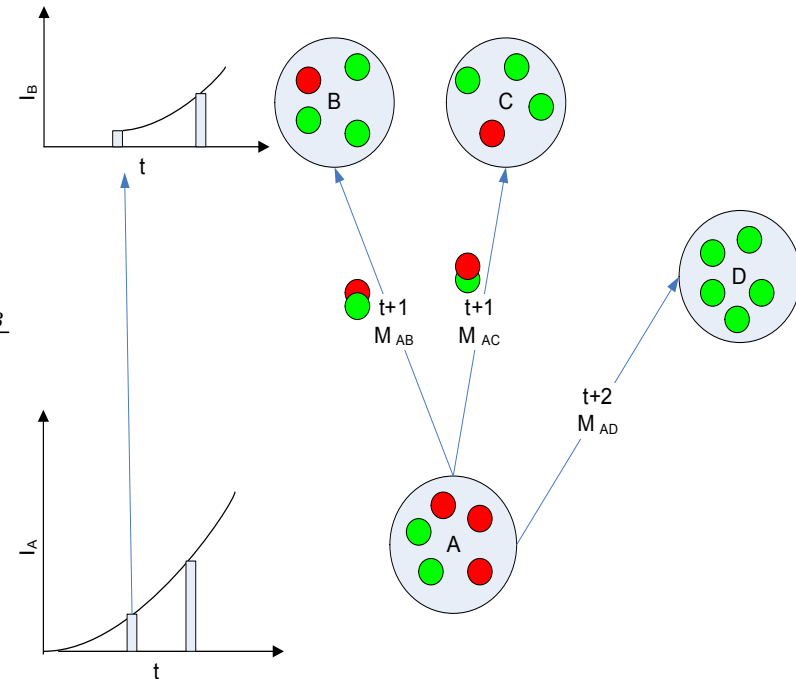
The infection is carried over to the target

β: contact rate
r: recovery rate
S: susceptible animals
I: infected animals
R: recovered animals
P: population
τ: infection threshold
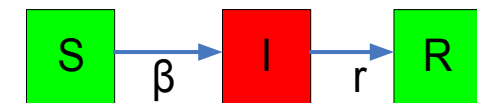M: transported animals

Transport component

$$I_{Bt+1} \geq \tau$$

$$I_{Bt+1} = I_{At} \frac{M_{AB}}{P_A}$$

Intra-population SI/SIR

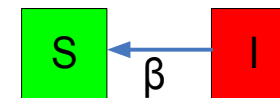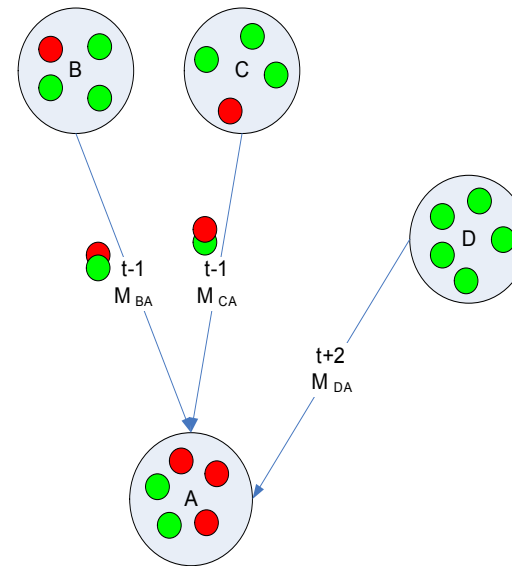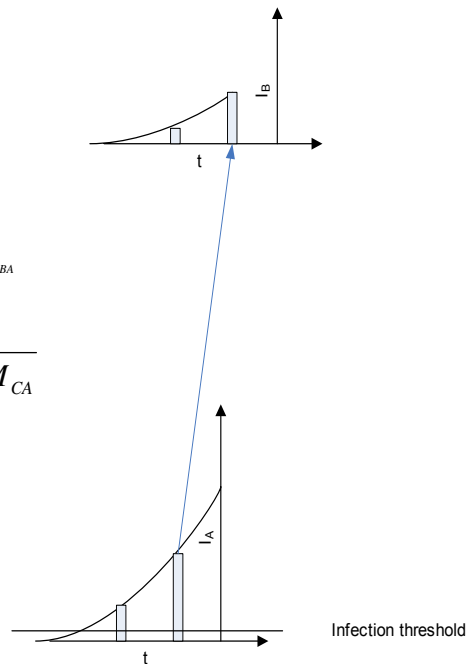$$I_{At+1} = I_{At}(1-r) + \beta S_{At} \frac{I_{At}}{P_A}$$

Backwards SI

β: contact rate
I: infected animals
P: population
τ: infection threshold
M: transported animals

Transport component

$$I_{At-1} \geq \tau \qquad I_{At-1} \leq M_{BA}$$

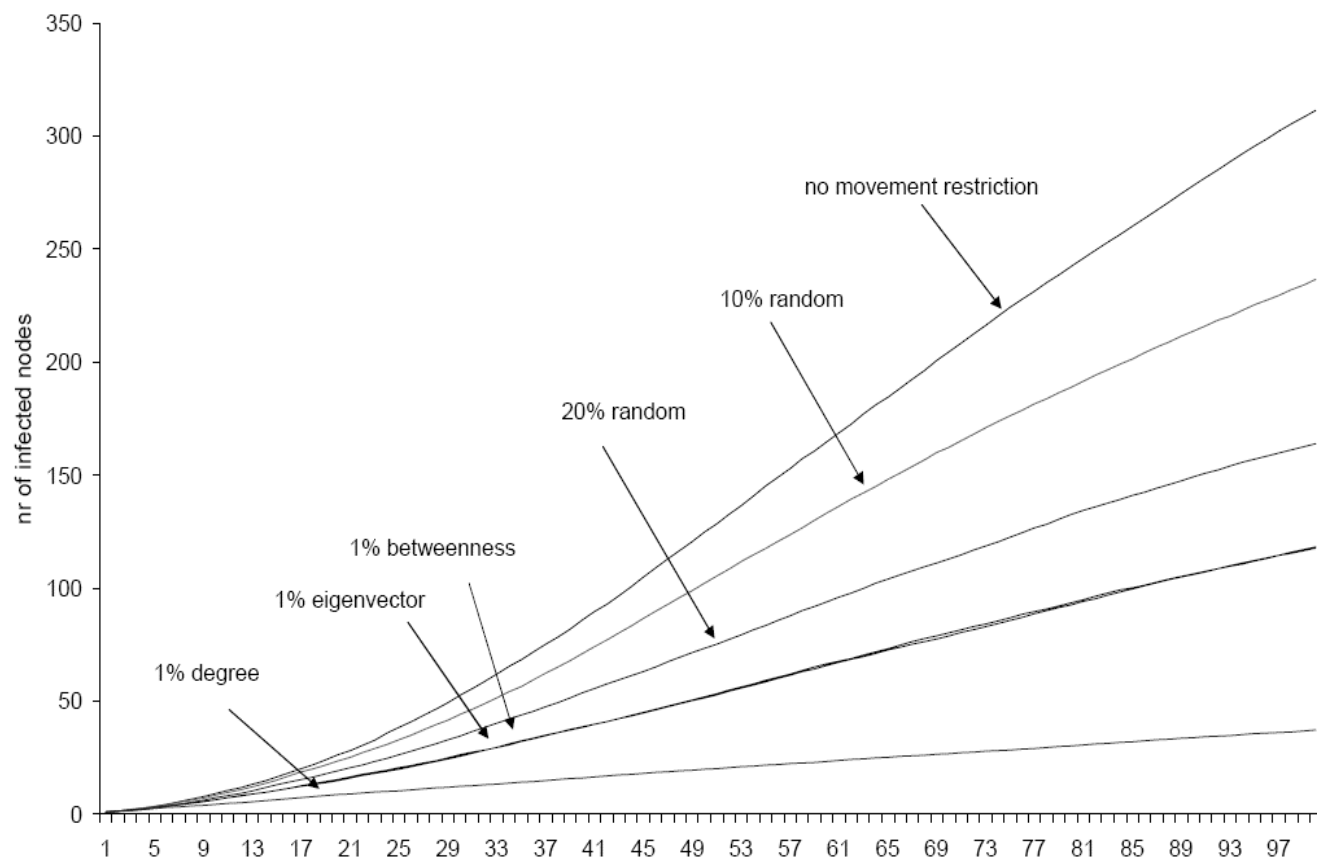$$I_{Bt-1} = I_{At-1} \frac{P_B}{M_{BA} + M_{CA}}$$

Infection threshold

Intra-population epidemic (backwards SI)

$$I_{At-1} = \frac{P_A + \beta P_A - \sqrt{(P_A + \beta P_A)^2 - 4\beta P_A I_{At}}}{2\beta}$$

Centrality and epidemics

Targeted control measures

Number of infected nodes by time for different scenarios of movement restriction measures (suspension of outgoing movements from cattle holdings selected on the basis of their values of centrality)

Spatial distribution

HERD

GENETIC CENTER

OTHER

MARKETS

ASSEMBLY CENTER
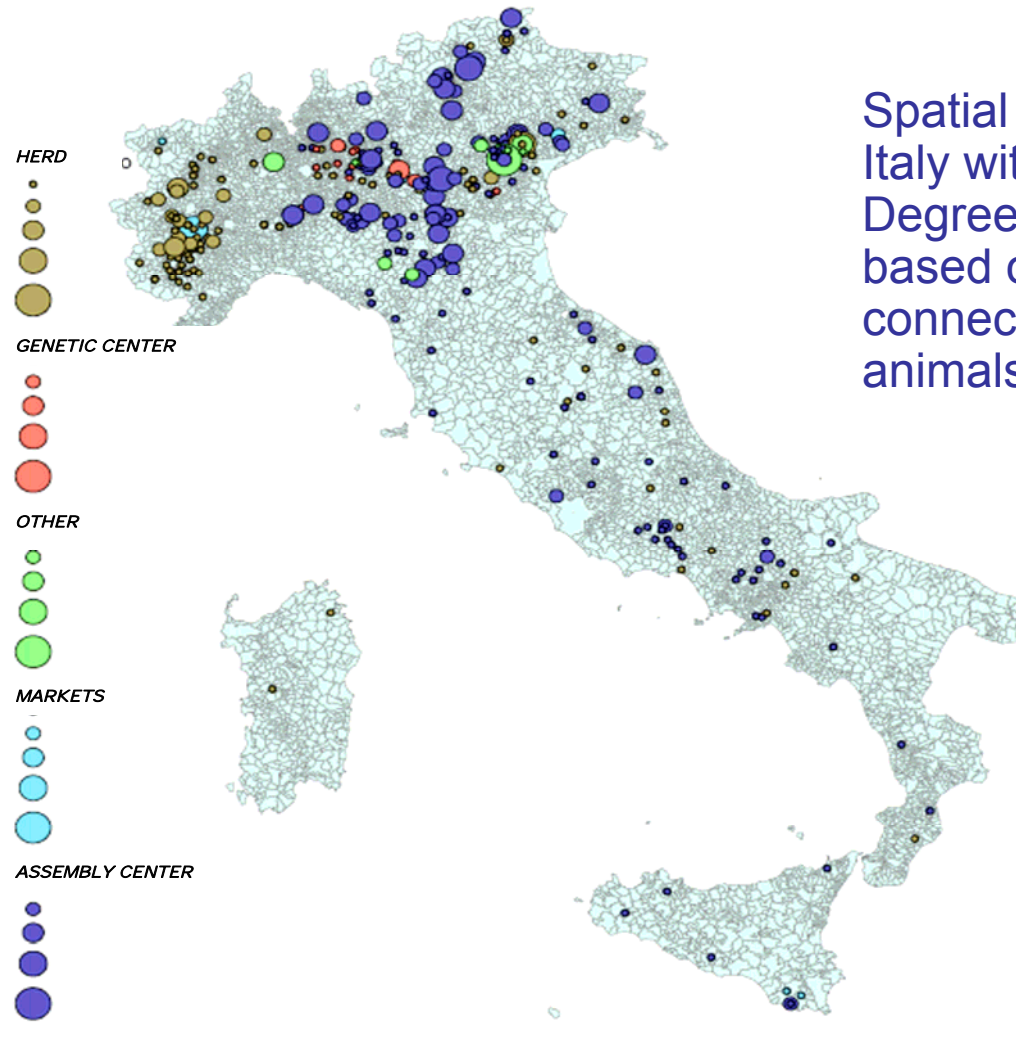
Spatial distribution of cattle premises in Italy with the highest CDC (Complex Degree Centrality) values in 2007. CDC is based on geometric mean of the number of connections (degree) and the number of animals moved

Dynamic vs static network

- livestock networks exhibit temporal instability of edges, nodes centrality, paths (P. Bajardi et al., to be published).

- temporal dynamics should also be considered when using traversal based centrality measures (betweenness, closeness) and other traversal based properties (average path length, components, clustering):

Issues to solve

- Model correctly with the traversal the process of interest (epidemic)

- Respect the temporality in the path sequence

- Include all possible paths and not only geodesics

- Reflect the temporality of some nodes (e.g. animal markets)

- Incorporate weights of contacts in the path

- Consider dilution of the process along the path

Disease Flow Centrality

- Traversing a network built from historical trade contacts, respecting specific disease flow rules
  - to simulate a simplified epidemic process

- Count the number of times nodes are contributing to the path
  - to characterize disease spread potential

- Count the number of times nodes are reached
  - to characterize vulnerability
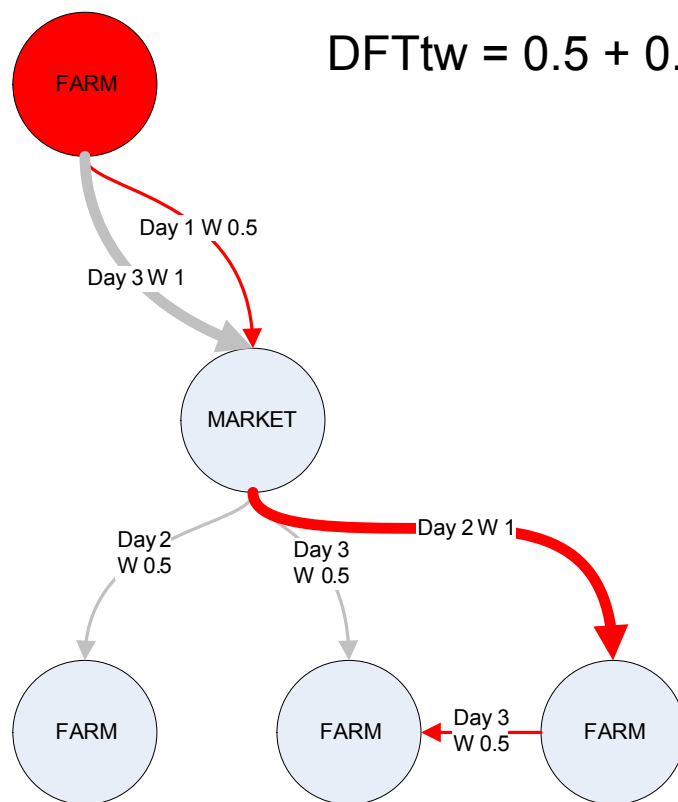
Traversal rules

Betweenness

DFC
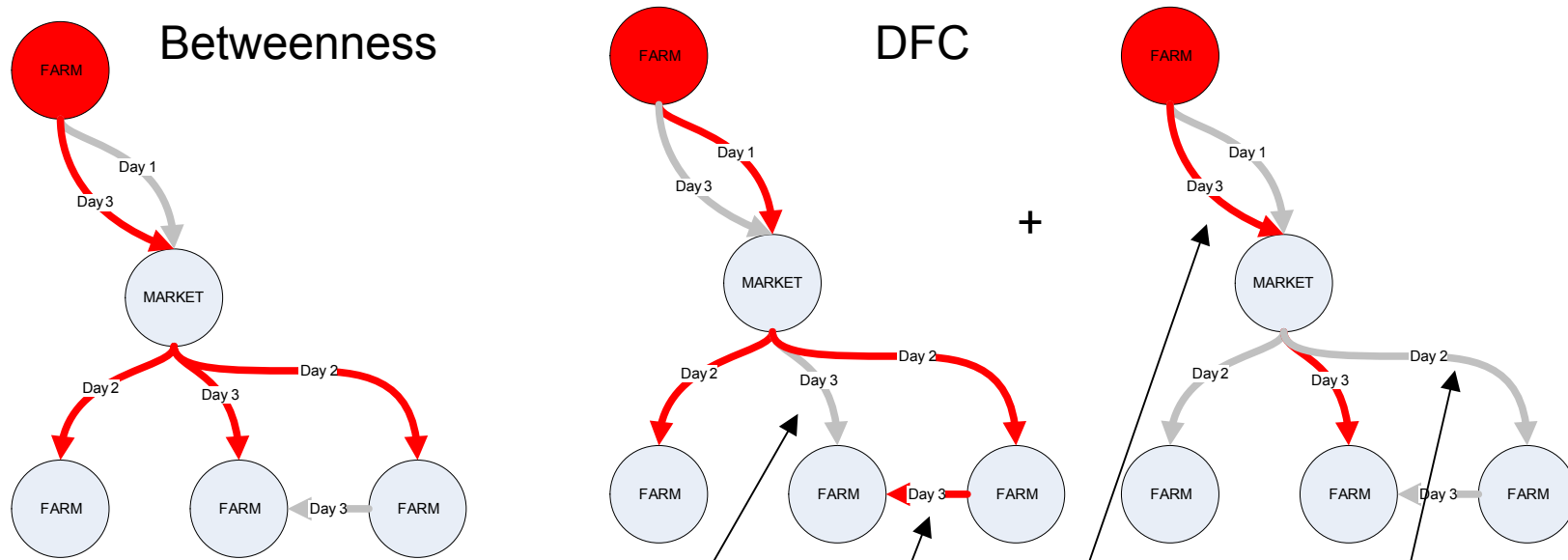
The market cannot preserve the infection lacking stable population

Also not shortest paths are considered

Illogic time sequences are not considered

Multiple paths are considered independently

DFC measures

- On top of the basic traversal algorithm several measures can be computed
  - DFC: accumulation number of contacts
    - as source -> responsibility of infection
    - as target -> vulnerability
  - DFCw:
    - conditional probability based on edges weights e.g. transported-population ratio
  - Selection of seeding sites
    - from all nodes in the network as seeding sites
    - from a selection of seeding sites in a geographical area

Weighted version

$$DFTtw = 0.5 + 0.5*1 + 0.5*1*0.5$$

FARM

Day 1 W 0.5

Day 3 W 1

MARKET

Day 2 W 0.5

Day 3 W 0.5

Day 2 W 1

FARM

FARM

FARM

Day 3 W 0.5

1. Specific risk for each contact
2. Dilution of probability along the path

- Local (in-degree, out-degree, strength)
  - ☺ simple to calculate
  - ☹ limited to the neighborhood

- Based on network traversal (betweenness, closeness)
  - ☺ reflect entire network structure
  - ☺ mimic a diffusion process
  - ☹ do not consider temporality of relations

- Disease Flow Centrality
  - ☺ reflect entire network structure
  - ☺ mimic a diffusion process for an epidemic considering temporality of relations
  - ☹ heavy to compute

| ID	Type of holding	DFC	Betweenness	Out degree	In degree
59655	Assembly centre	101476146	7592971	215	2500
73929	Other farm	60989409	400573	75	79
177	Market	48179111	8410178	133	564
30567	Dairy farm	46853207	11138817	64	264
45327	Dairy farm	44799584	1234901	112	147
59631	Fattening farm	44174380	1622632	9	10
173691	Assembly centre	31360217	17939747	152	1769
189	Market	30717751	12619719	24	1102
12790	Dairy farm	26028603	111739	10	1
89083	Other farm	20354523	575616	37	220
211688	Other farm	14343992	2801	1	2
57894	Dairy farm	14056882	618908	102	131
13324	Dairy farm	13155248	476991	10	6
21045	Assembly centre	13092463	4729504	41	172
166168	Fattening farm	13079296	1346552	66	89

Centrality of the top 15 cattle holdings classified on DFC
(transports of Jan 2009)
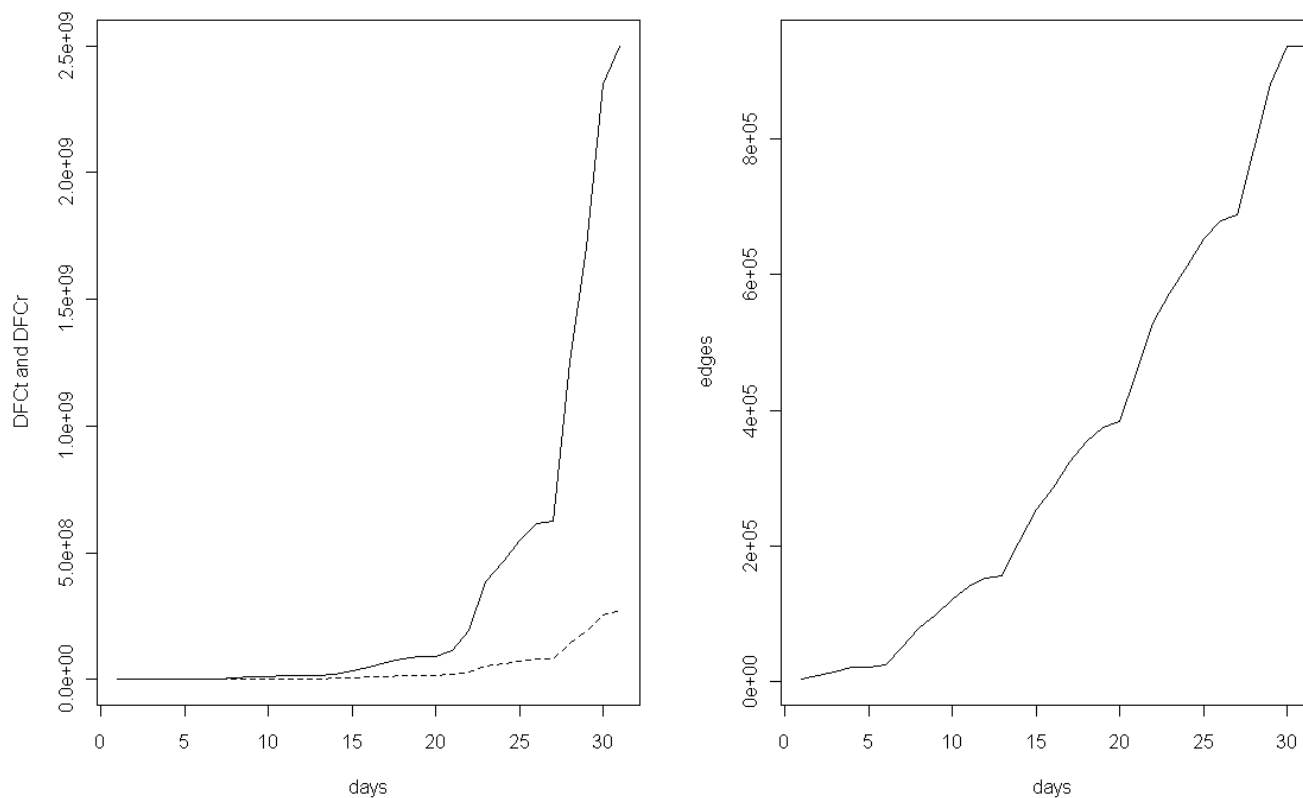
Total DFCr (solid line), DFCr (dashed line) and number edges (right graph) for networks with expanding time window

Overall network risk over time
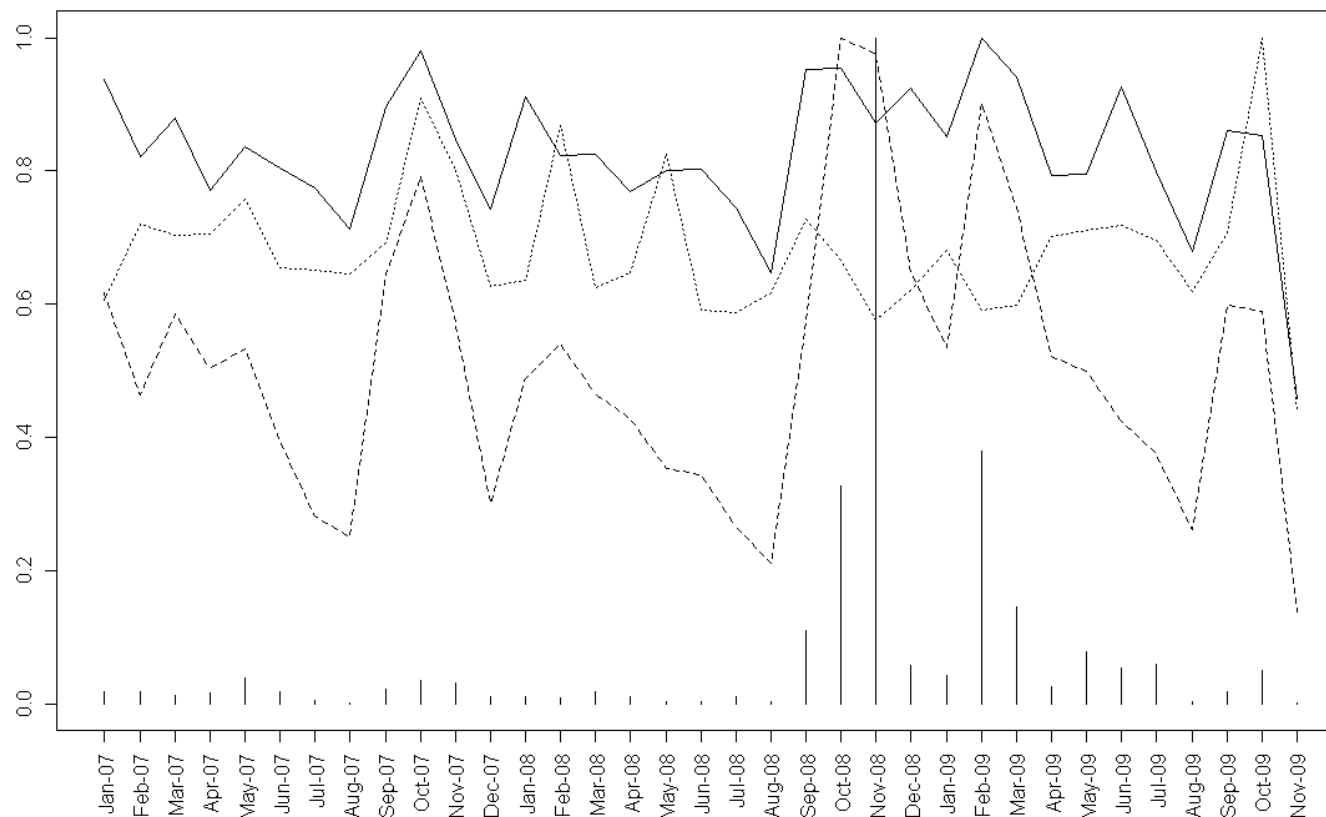
Total DFCt (bars), number of edges (solid line), average path length (dotted line), number of geodesics weighted on the inverse of their path length (dashed line)
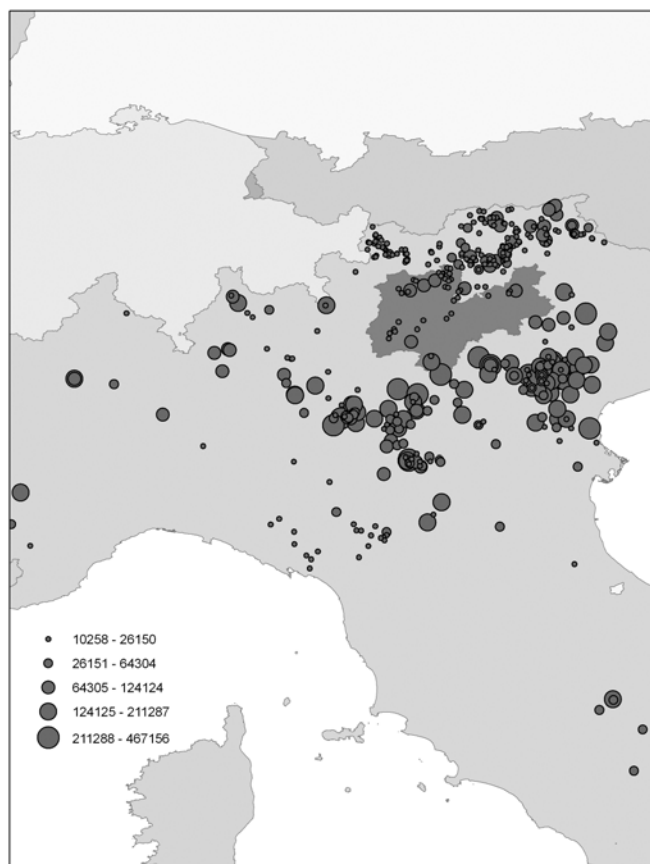
Holdings reached by epidemic paths from the Province of Trento, on the basis of cattle movements in January 2009

Future work

- More efficient DFC algorithm

- More complex path rules on-top of the basic DFC algorithm

- Network based scan statistics (can we scan for anomalies in groups of connected holdings using network adjacency instead of geographical adjacency?)

- To integrate stochastic components in SIR models based on real cattle networks