A stochastic agent-based model of pathogen propagation in dynamic multi-relational social networks

Bilal Khan, Kirk Dombrowski, and Mohamed Saad,
Journal of Transactions of Society Modeling and Simulation International, SAGE, 2014).

Presented by: Gisela Font Sayeras

Overview

- Network Simulation
- ERGM vs ABM
- Model
- Design of MABUSE, the system
- Validation

Network Simulation

- What is a network?
- What is a dynamic network?
- Networks in the real world
- Networks in simulation

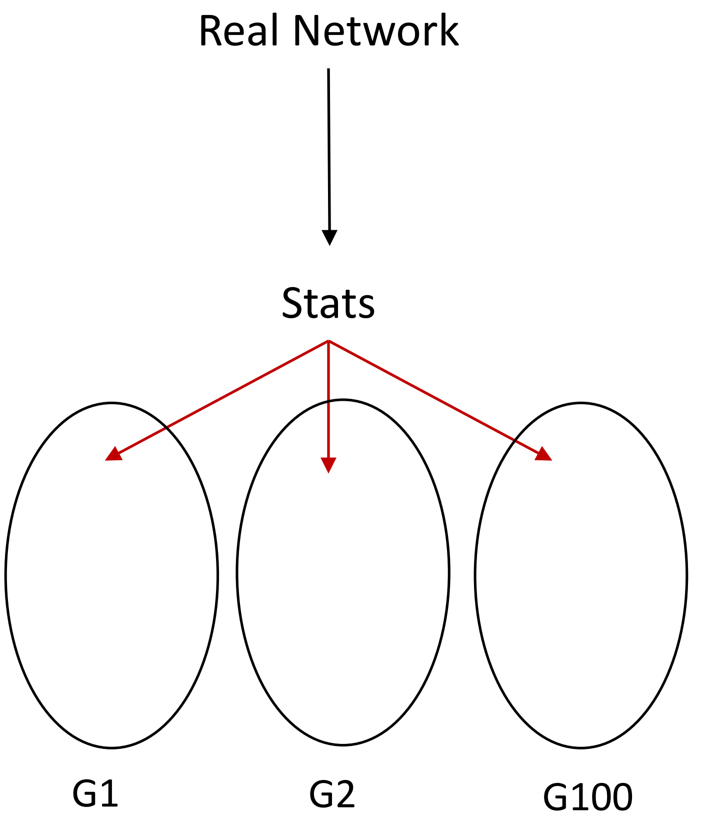
Main network simulation strategies

- ERGM (Exponential Random Graph Modeling)

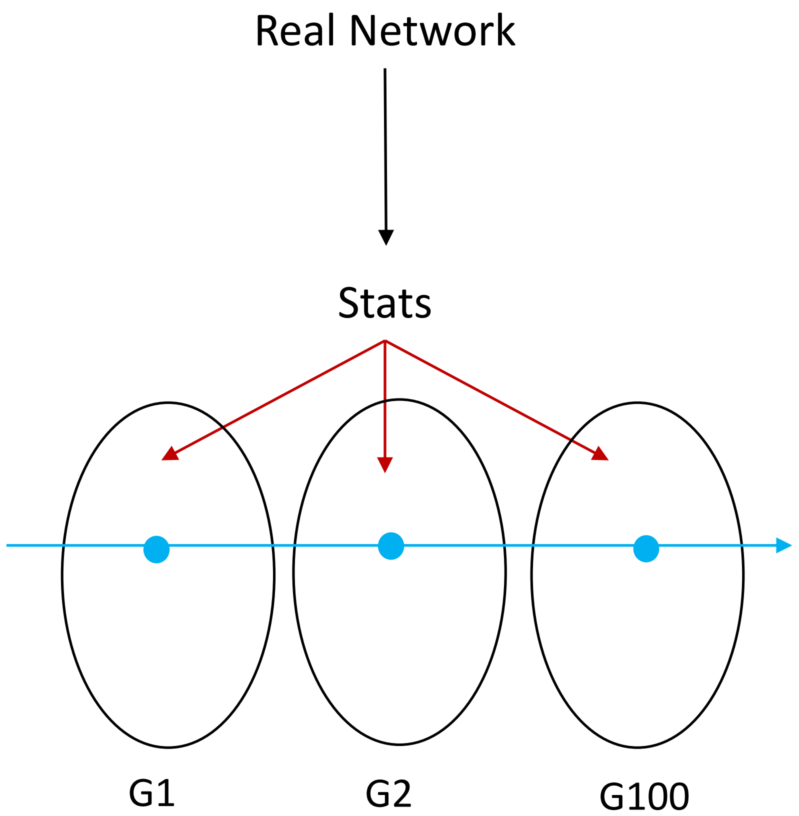
  Hunter DR, Handcock MS, Butts CT, Goodreau SM, Morris M. **ergm: A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks**. *Journal of statistical software*. 2008;24(3):nihpa54860.

- ABM (this class)

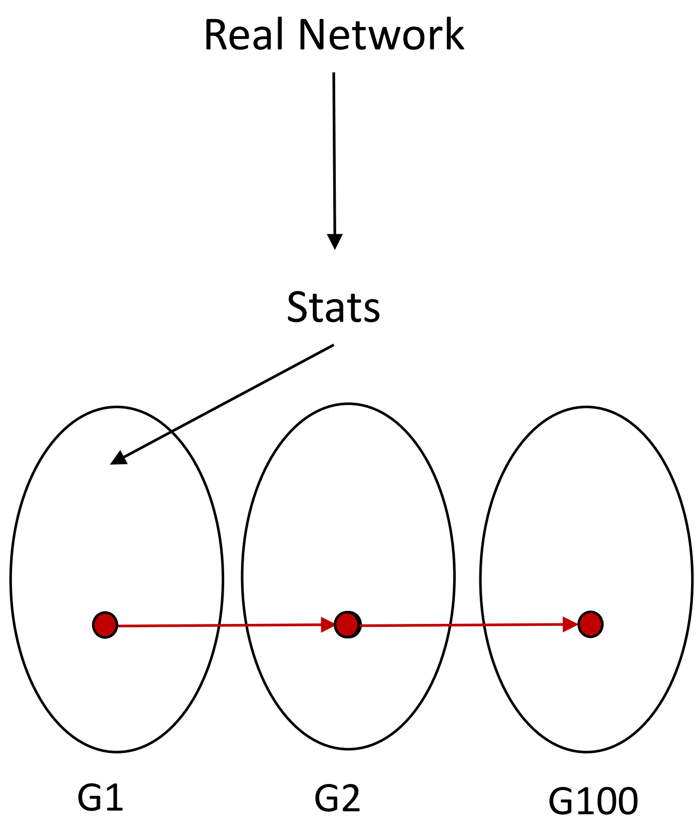
Exponential Random Graph Modeling (ERGM)



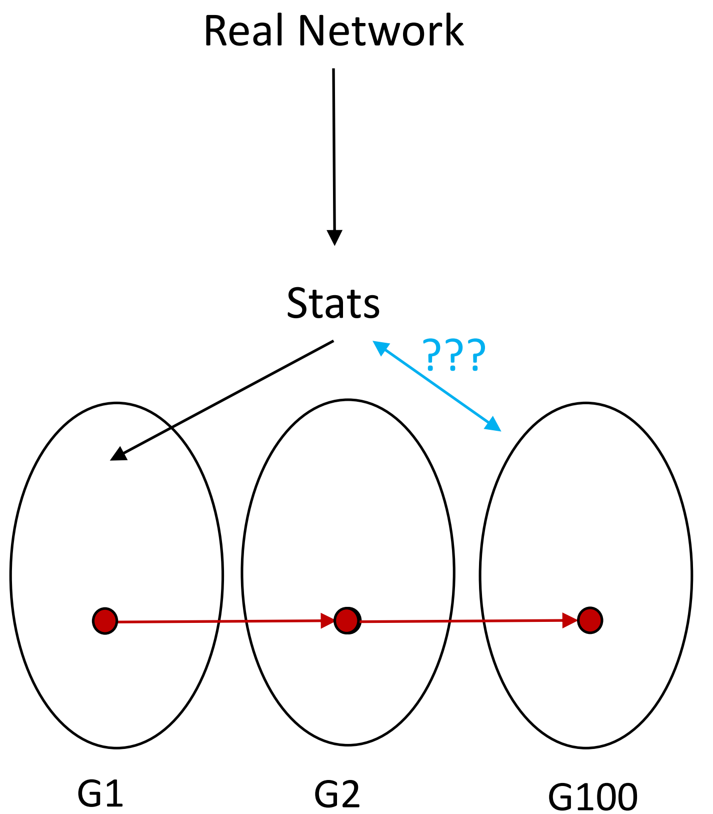
Exponential Random Graph Modeling (ERGM)



Agent Based Model Simulation (ABM)



Agent Based Model Simulation (ABM)



ERGM vs ABM

- ERGM gives you realistic network-level guarantees but lacks realistic individual agency
  - The network gets to change realistically over time
  - What each agent is doing could be unrealistic
- ABM gives you lack realistic individual agency but lacks realistic network-level guarantees
  - Each agent gets to change realistically over time
  - What the network is doing could be unrealistic

MABUSE

- MABUSE gives you
  - The network gets to change realistically over time
  - Each agent gets to change realistically over time

Outline of MABUSE Model

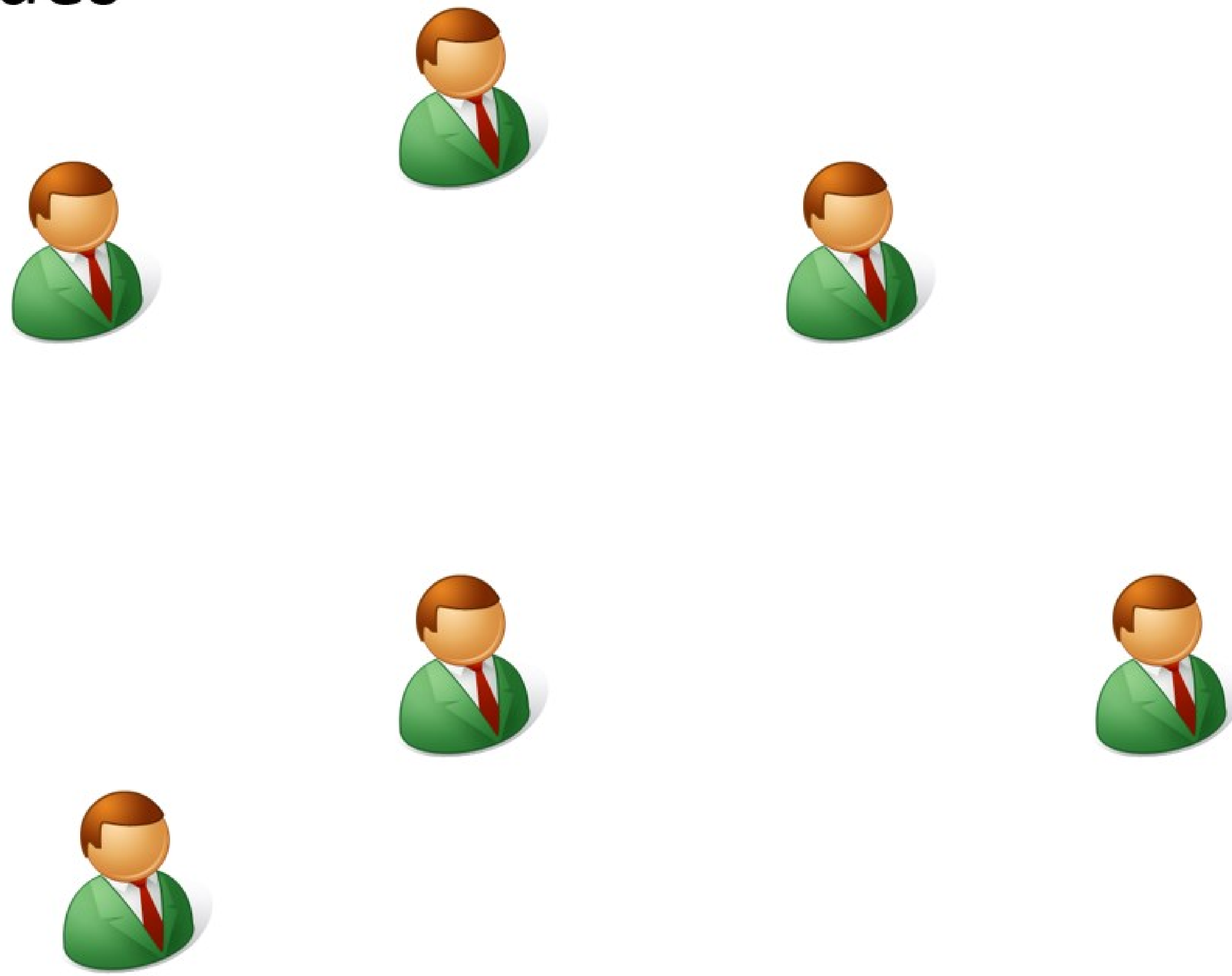
INITIALIZATION

1. Nodes are made (including some infected)
2. Edges are made

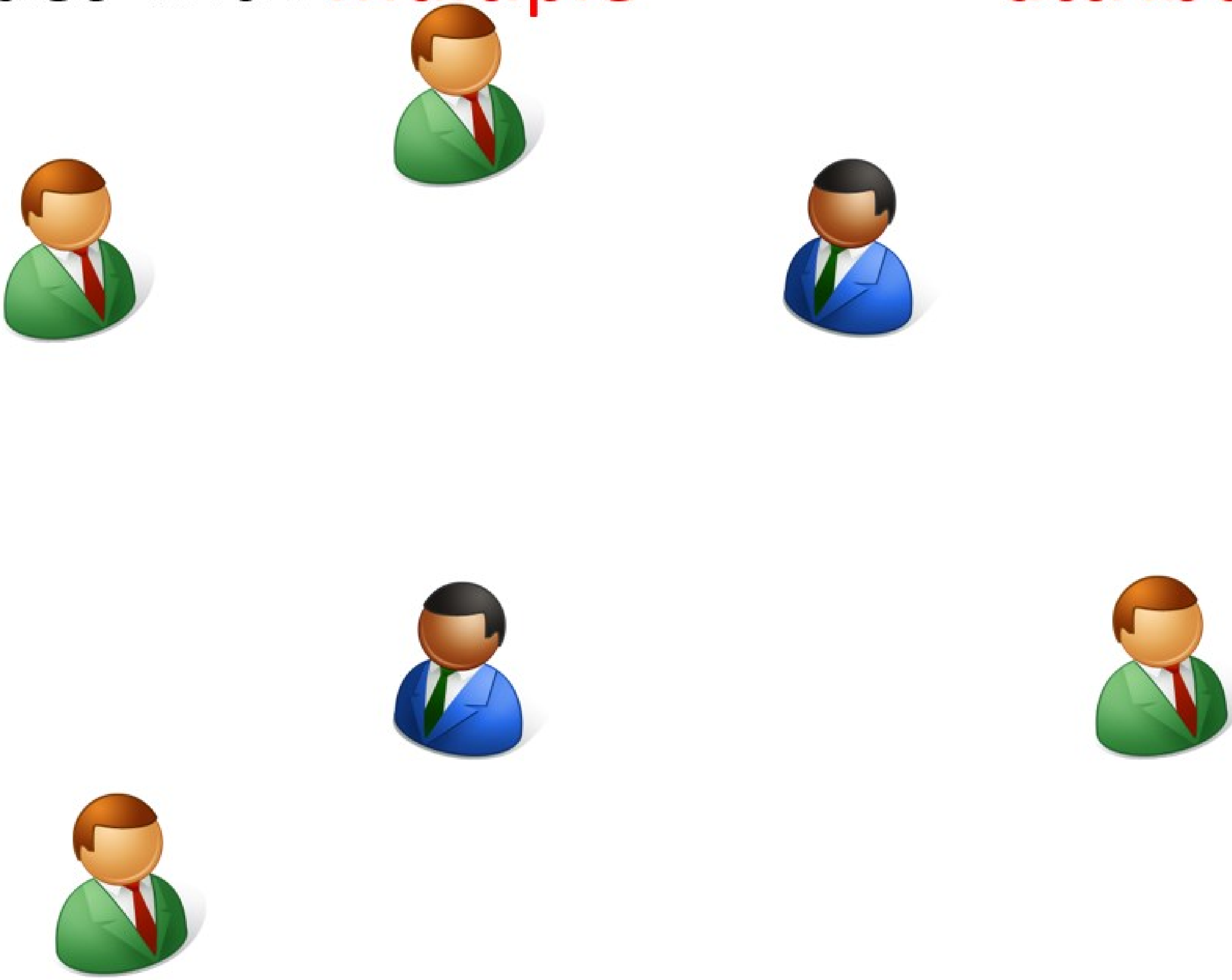
SIMULATION RUN

3. Edges change over time
4. Risk acts happen with neighbors
5. Nodes come and go

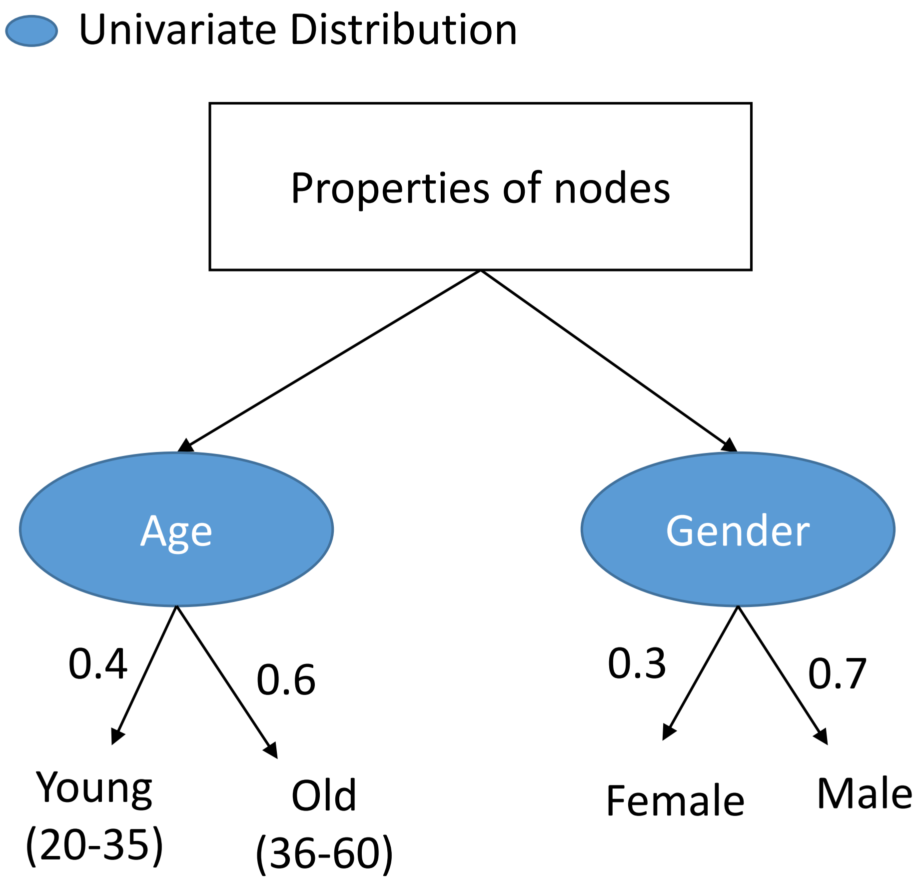
Nodes



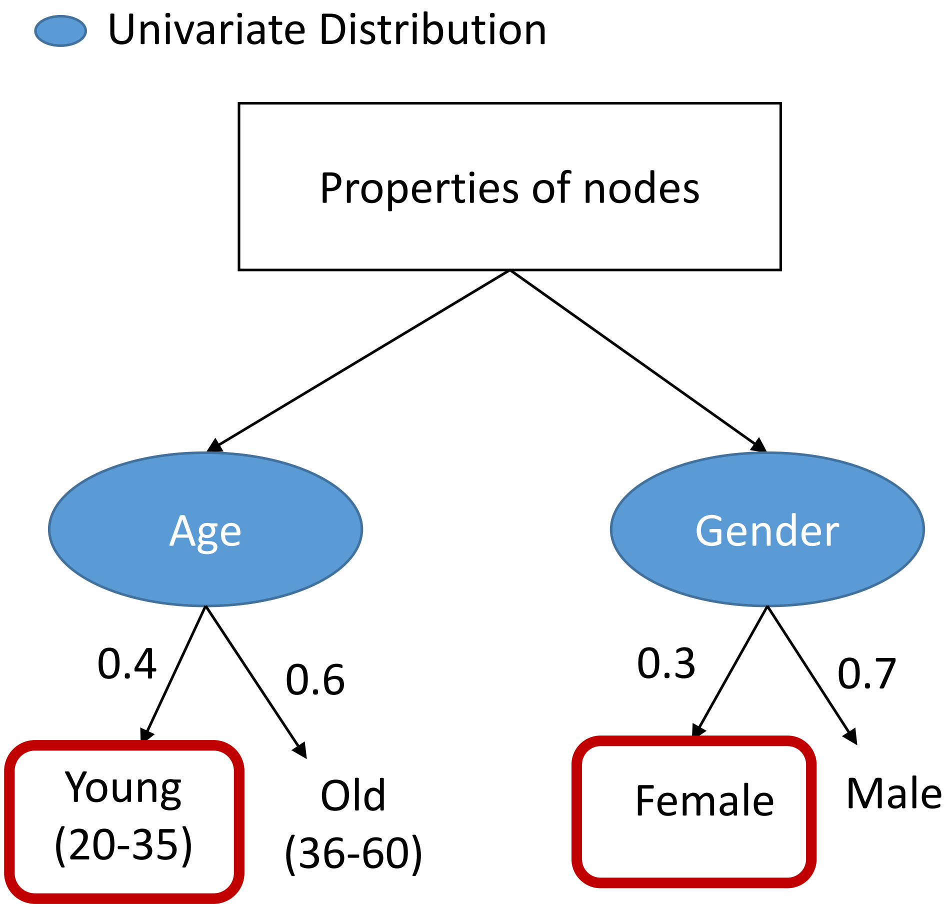
Nodes with multiple attributes



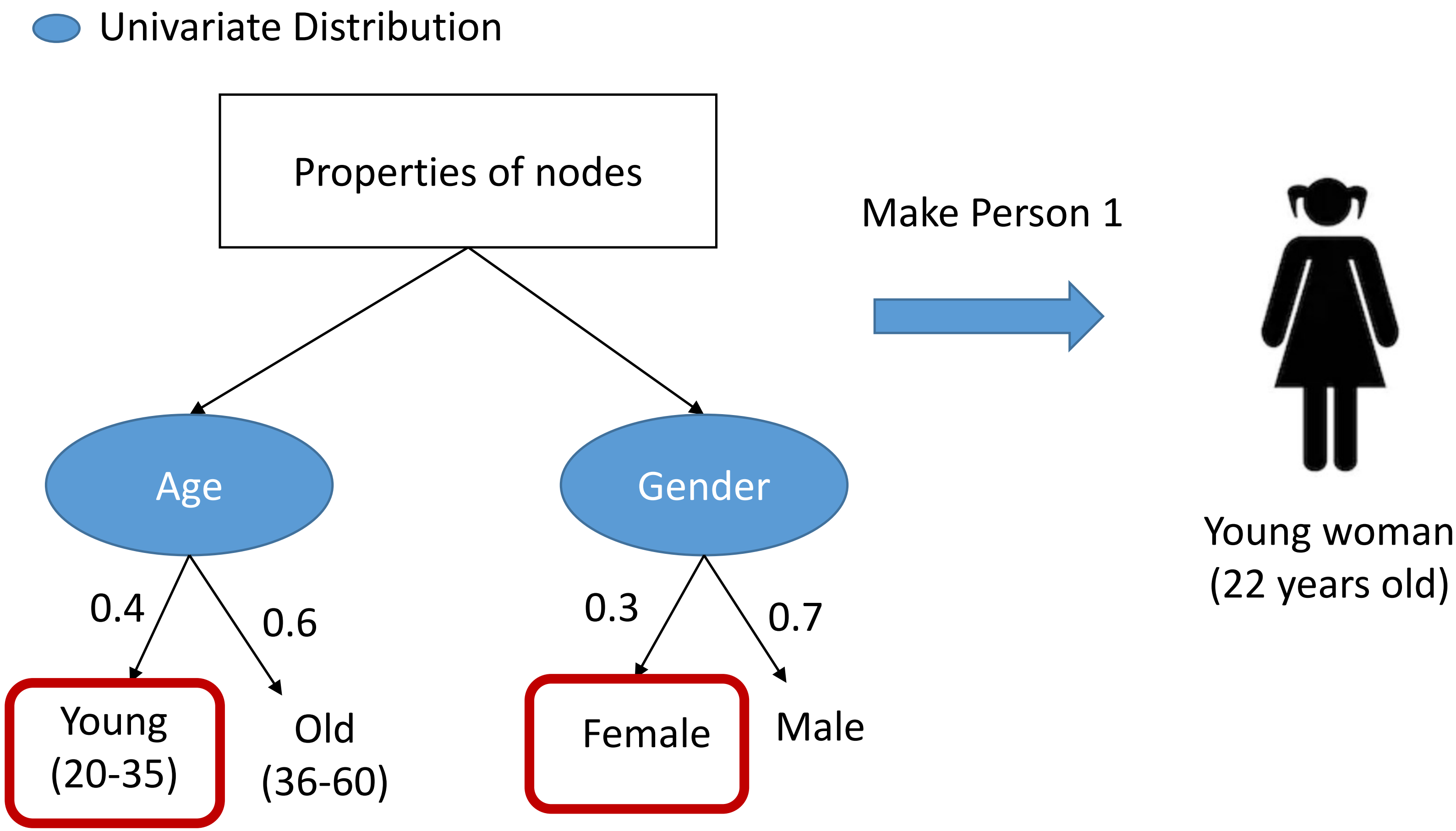
How are the properties of individuals determined?



How are the properties of individuals determined?

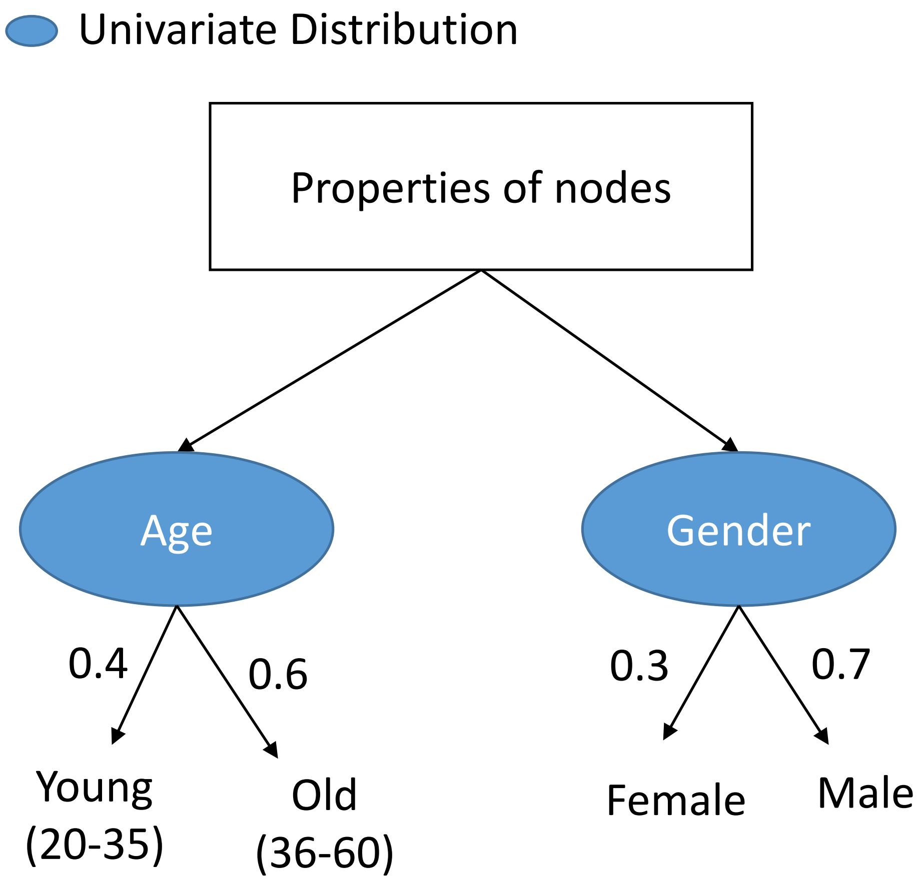


How are the properties of individuals determined?

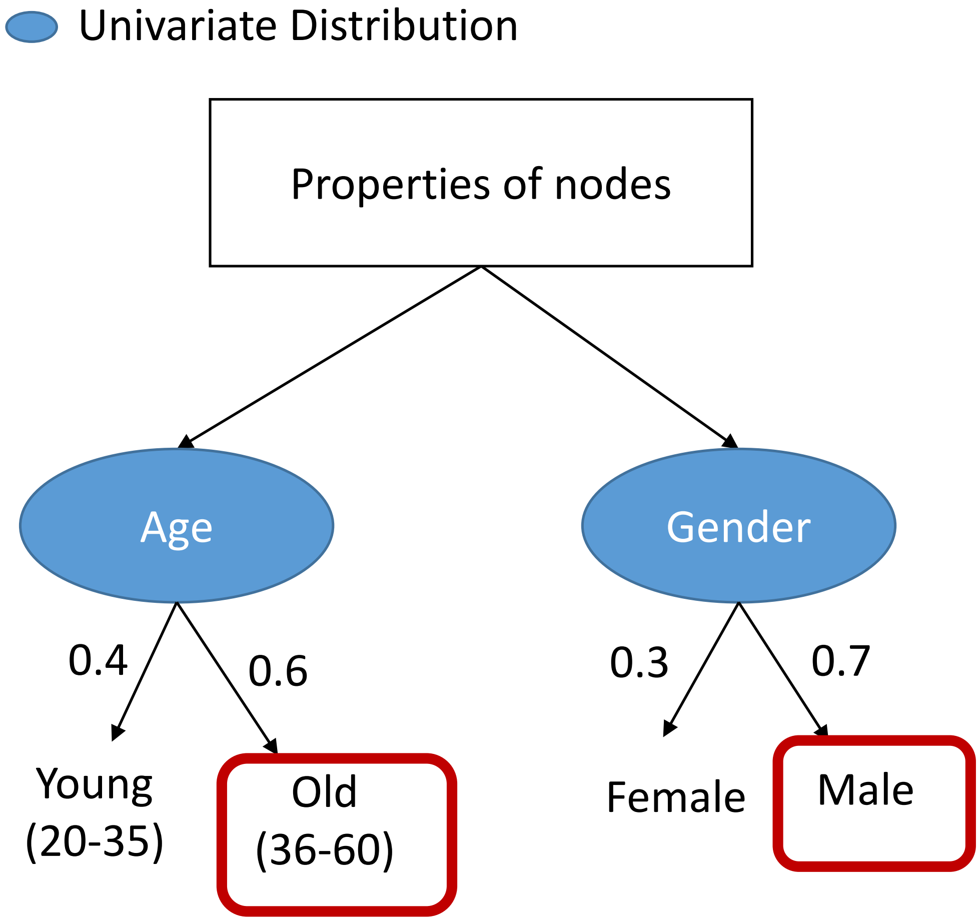


Population

How are the properties of individuals determined?

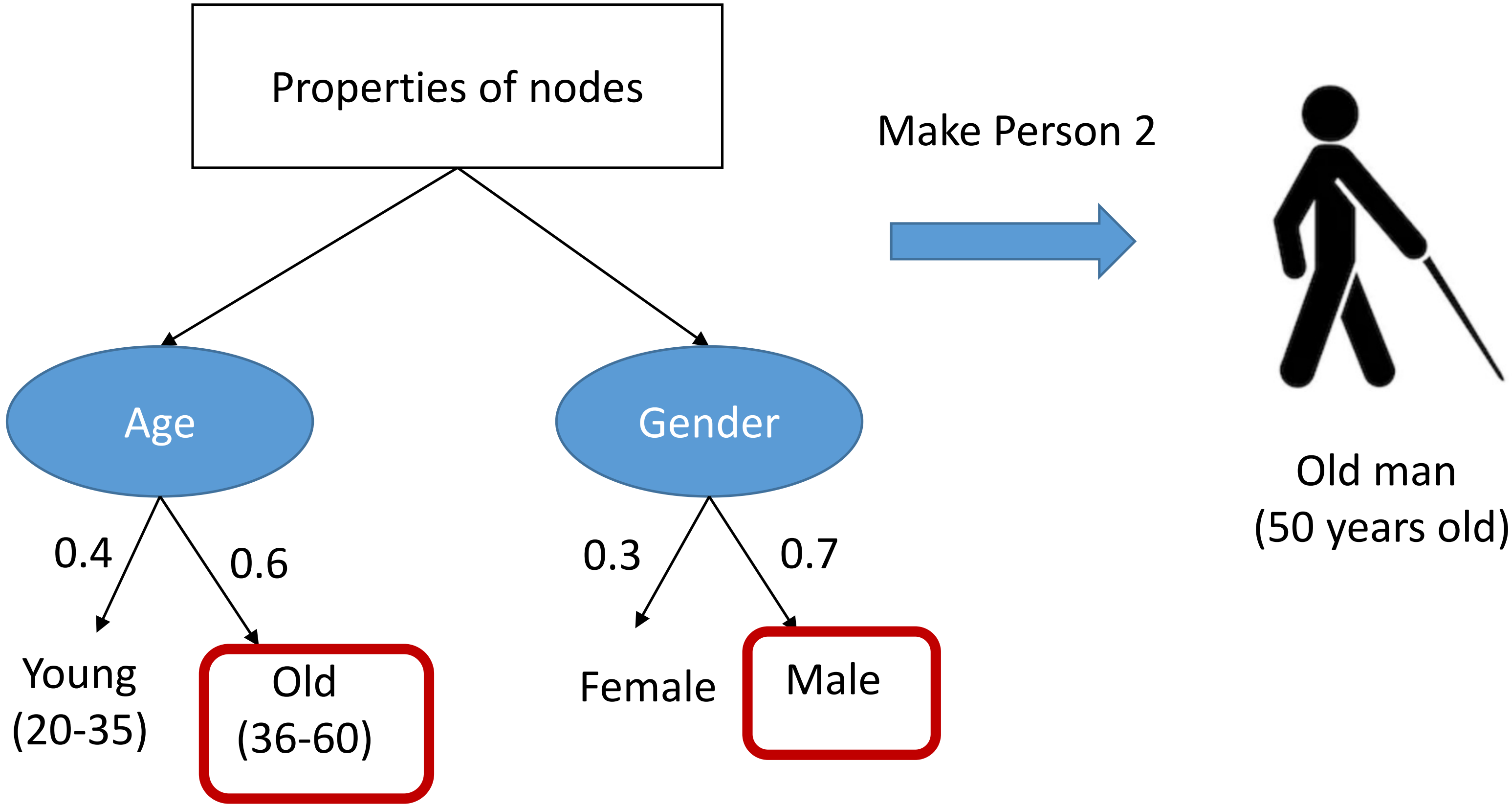


How are the properties of individuals determined?

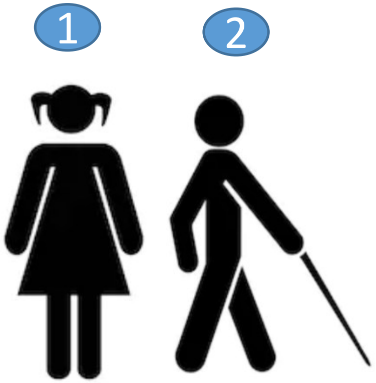


How are the properties of individuals determined?

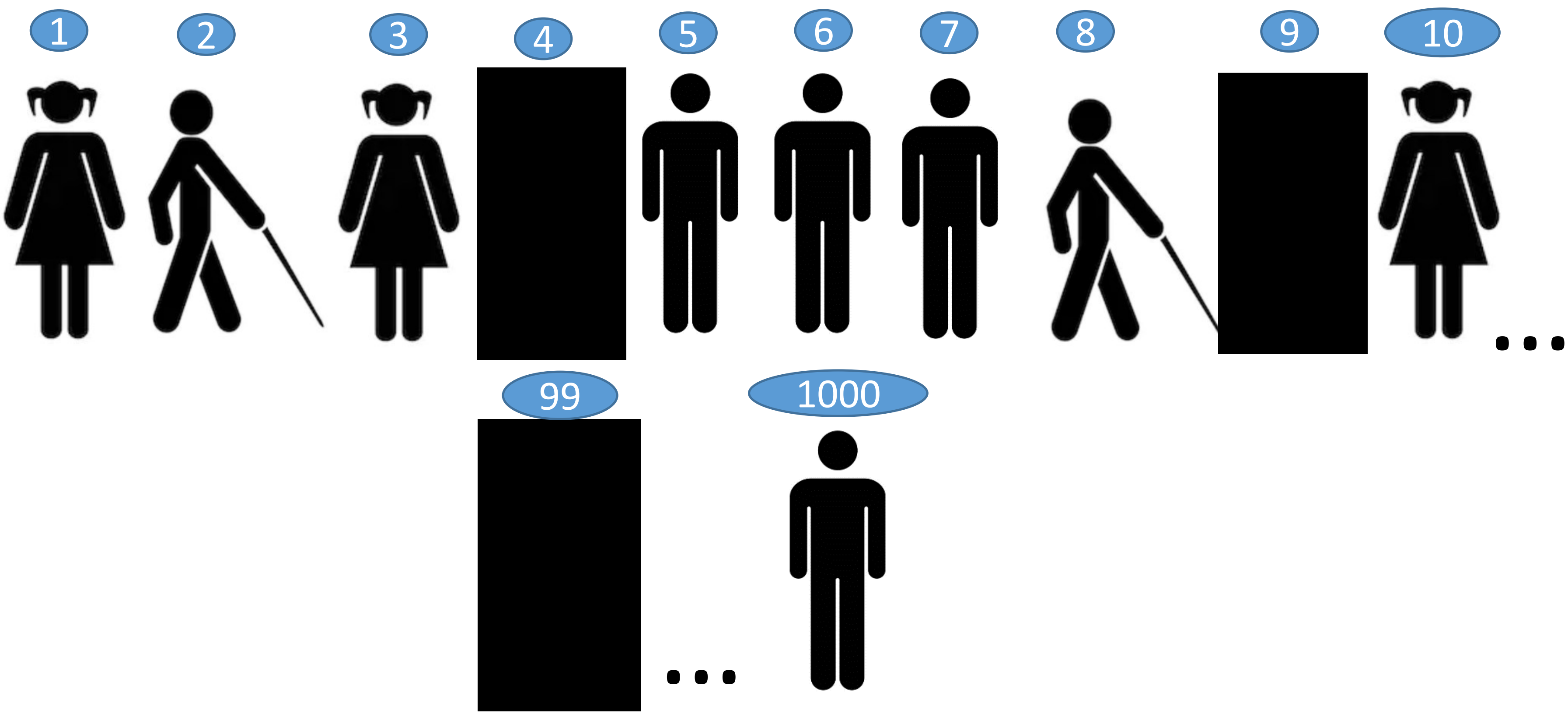
Univariate Distribution



Population



Population



Outline of MABUSE Model

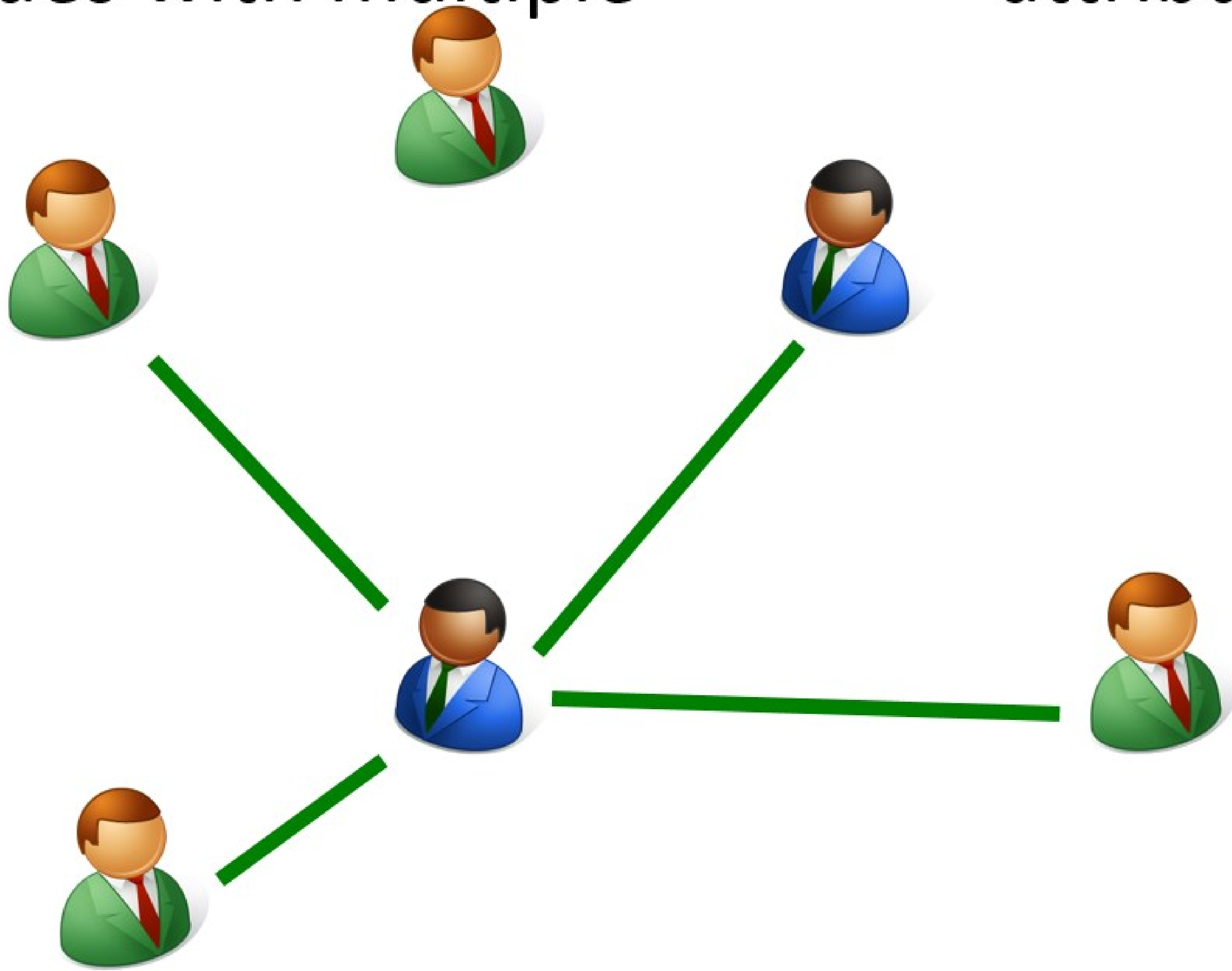
INITIALIZATION

1. Nodes are made (including some infected)
2. Edges are made

SIMULATION RUN

3. Edges change over time
4. Risk acts happen with neighbors
5. Nodes come and go

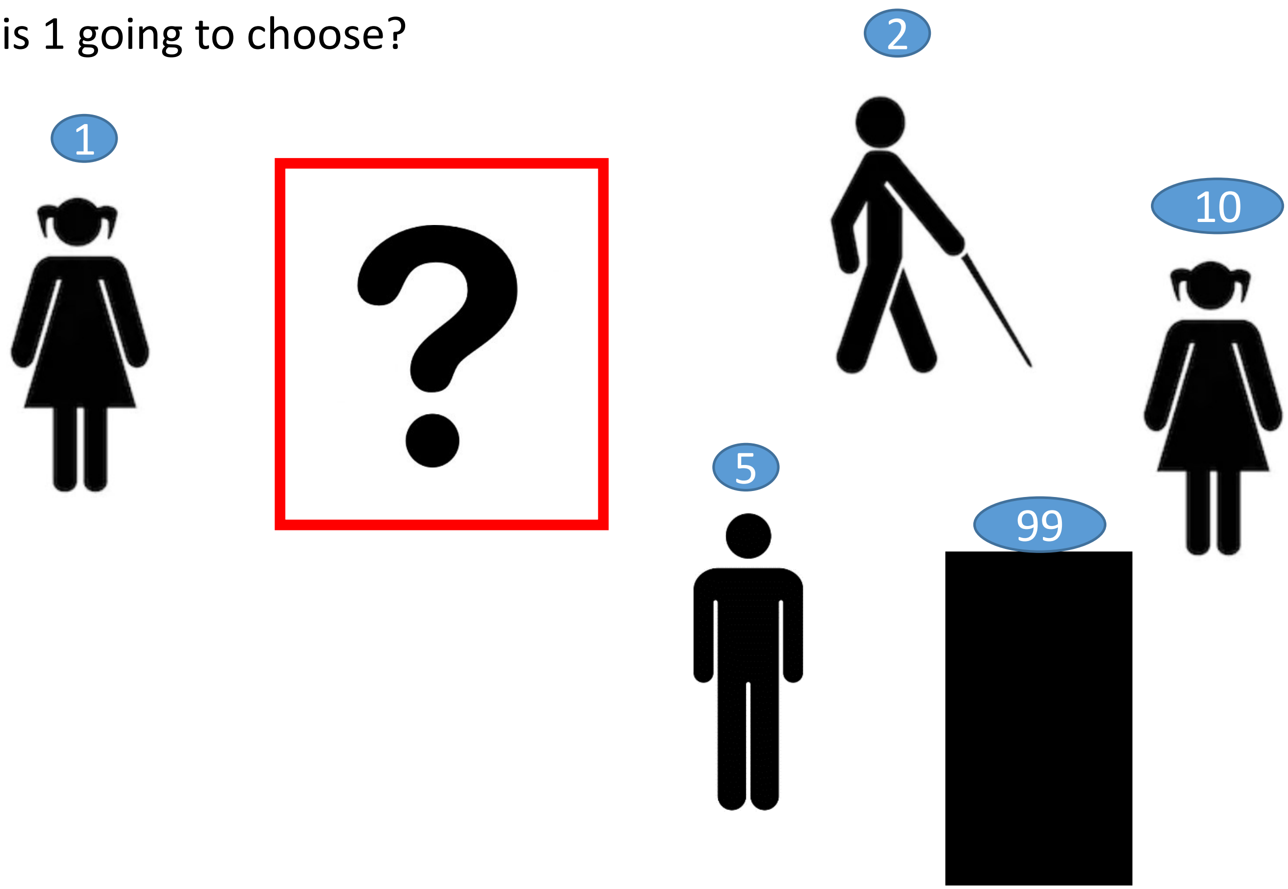
Nodes with multiple attributes



edges

How individuals choose who to form relationships with?

Who is 1 going to choose?



How is the choice going to be made?

Bivariate Distribution

Gender Bivariate

	Male	Female
Male	0.8	0.2
Female	0.25	0.75

Age Bivariate

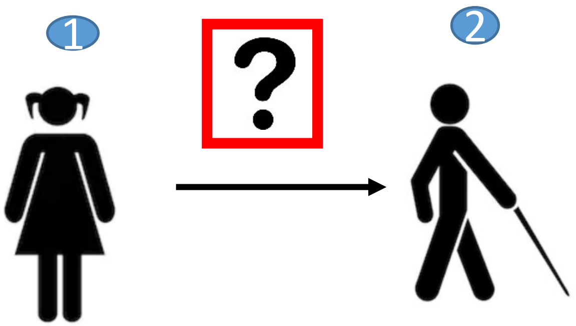
	Young	Old
Young	0.6	0.4
Old	0.3	0.7

Aggregator

- Gender: 0.8
- Age: 0.3

(from ERGM)

How is the choice going to be made?



Gender Bivariate

	Male	Female
Male	0.8	0.2
Female	0.25	0.75

Age Bivariate

	Young	Old
Young	0.6	0.4
Old	0.3	0.7

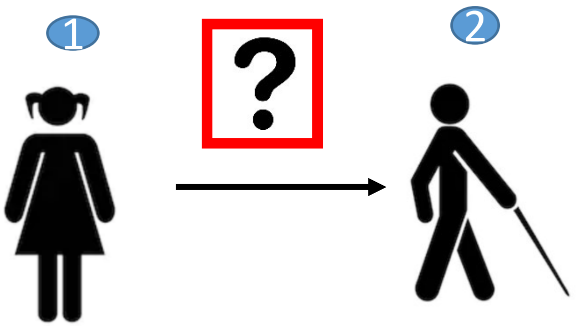
Aggregator

- Gender: 0.8
- Age: 0.3

(from ERGM)

How is the choice going to be made?

Calculate Propensity



Gender Bivariate

	Male	Female
Male	0.8	0.2
Female	0.25	0.75

Age Bivariate

	Young	Old
Young	0.6	0.4
Old	0.3	0.7

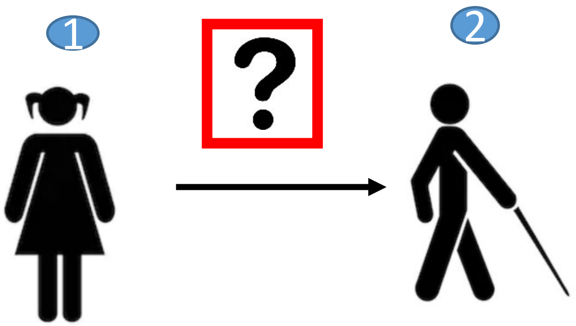
Aggregator

- Gender: 0.8
- Age: 0.3

(from ERGM)

How is the choice going to be made?

Calculate Propensity



Gender Bivariate

	Male	Female
Male	0.8	0.2
Female	0.25	0.75

Age Bivariate

	Young	Old
Young	0.6	0.4
Old	0.3	0.7

Aggregator

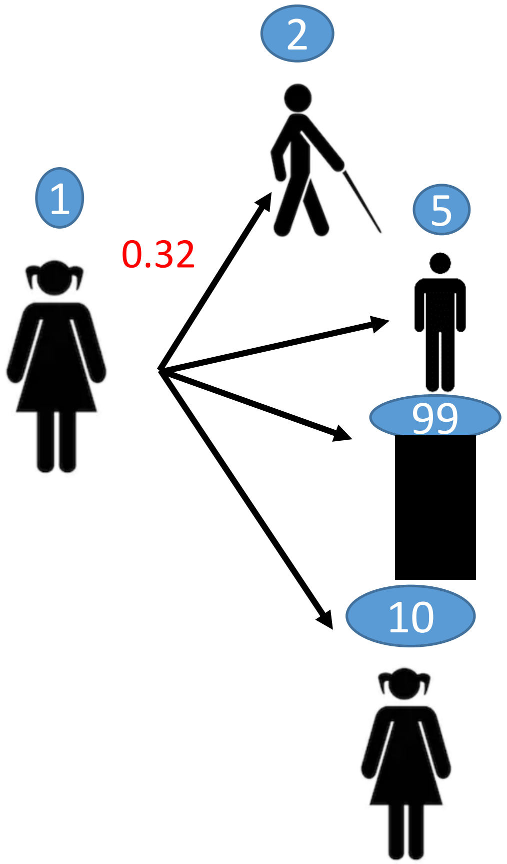
- Gender: 0.8
- Age: 0.3

(from ERGM)

Propensity 1-2 = 0.25*0.8 + 0.4*0.3 = 0.32

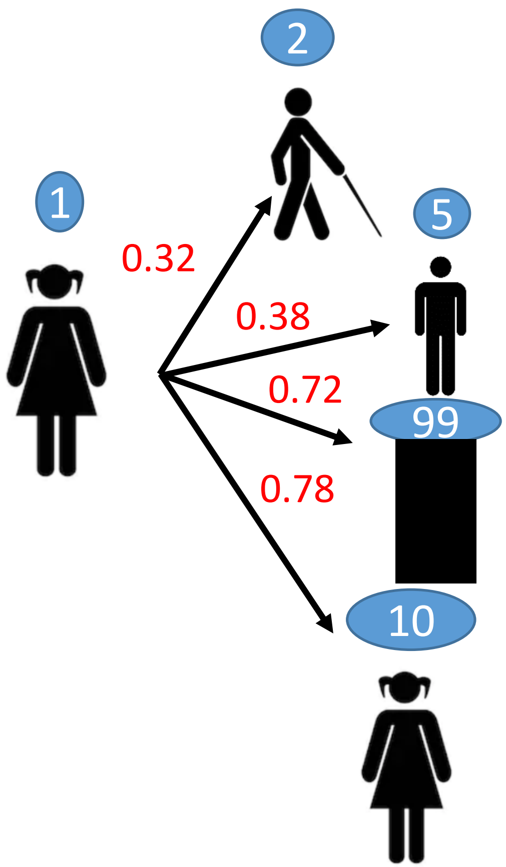
How is the choice going to be made?

Calculate Propensity

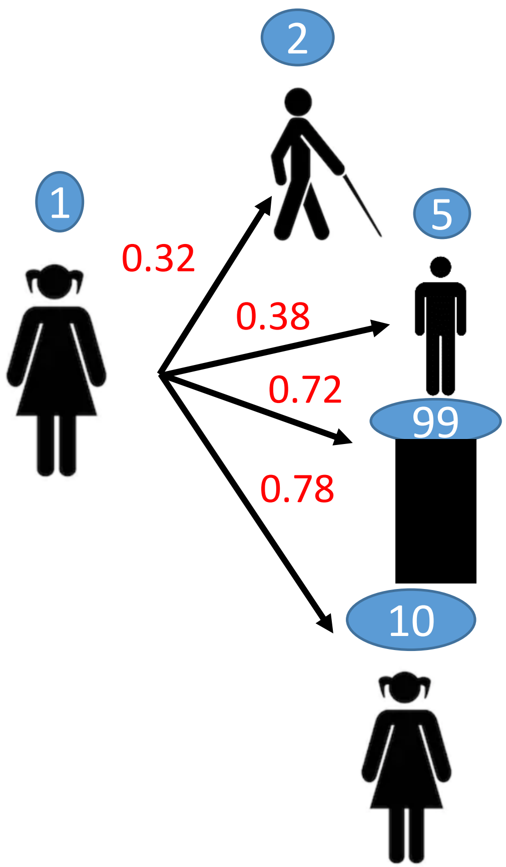


How is the choice going to be made?

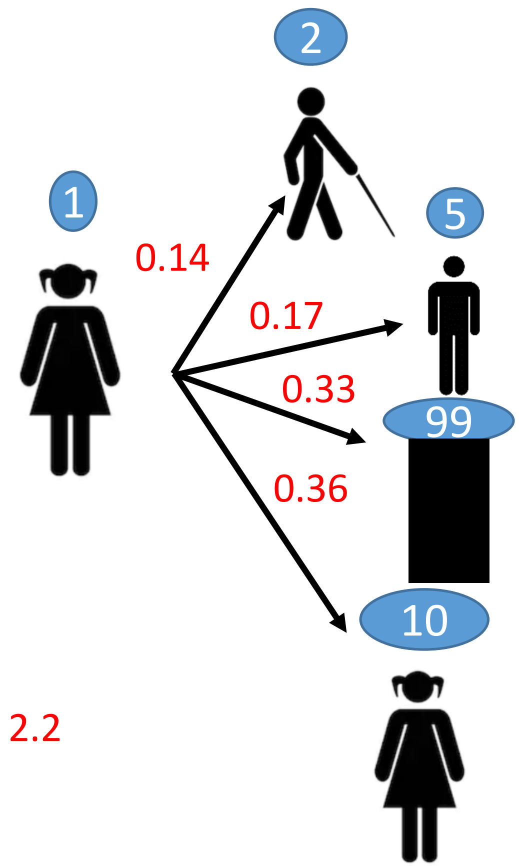
Calculate Propensity



How is the choice going to be made?



Normalize

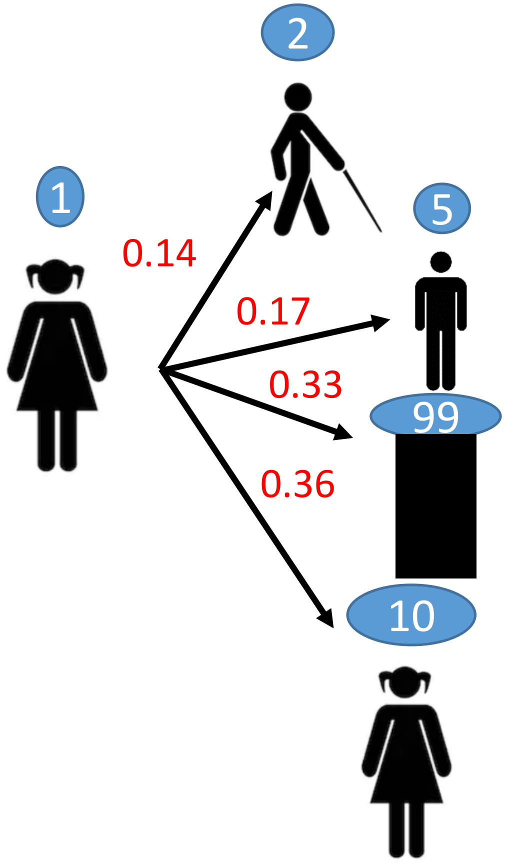


0.32+0.38+0.72+0.78 = 2.2

0.32 / 2.2 = 0.14

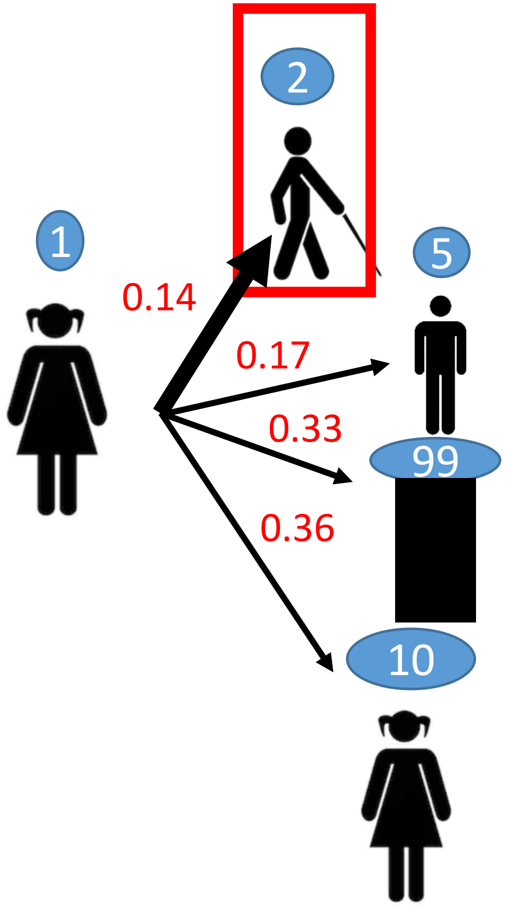
How is the choice going to be made?

Random Selection

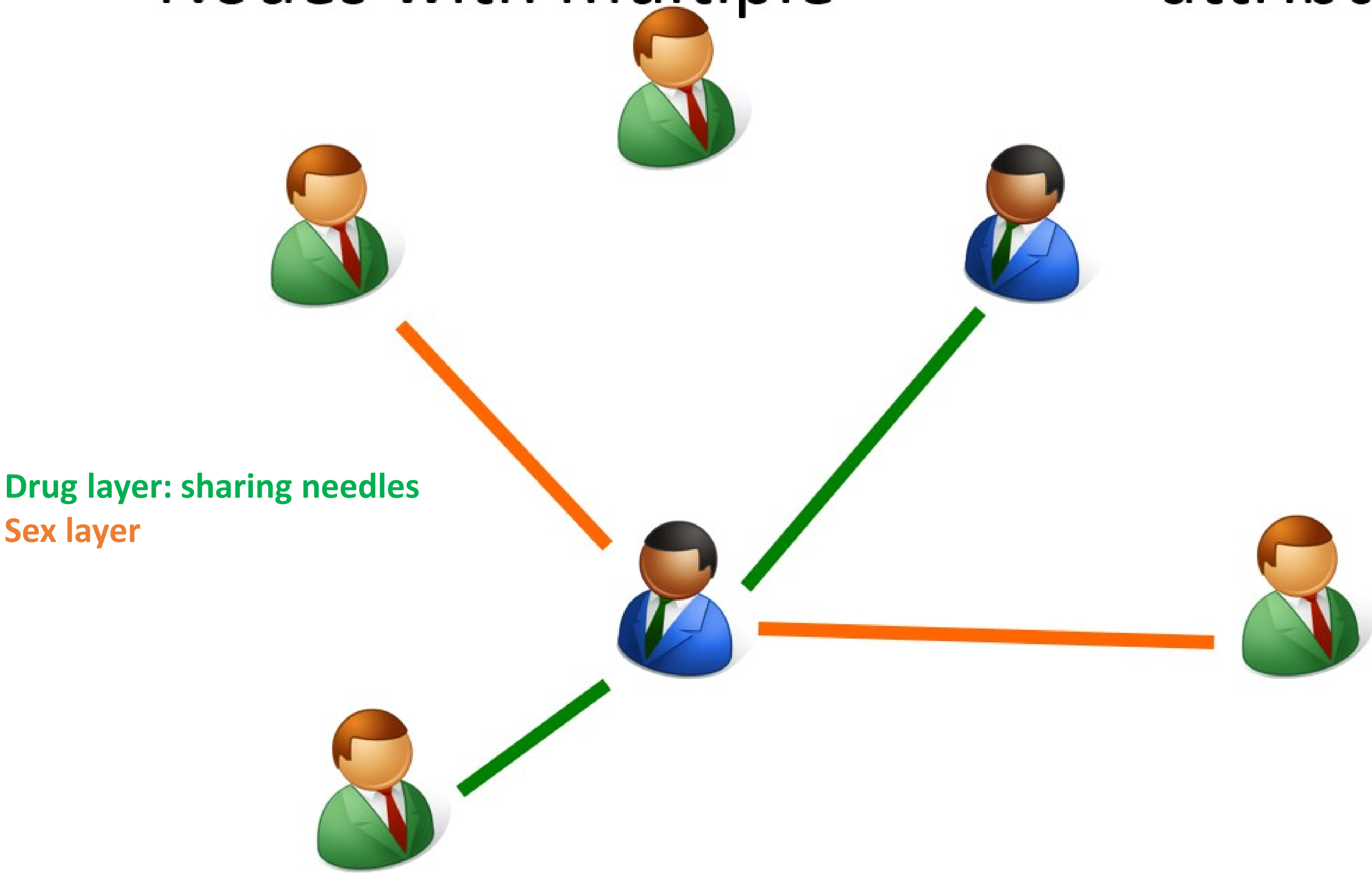


How is the choice going to be made?

Random Selection



Nodes with multiple attributes



edges of different types

Outline of MABUSE Model

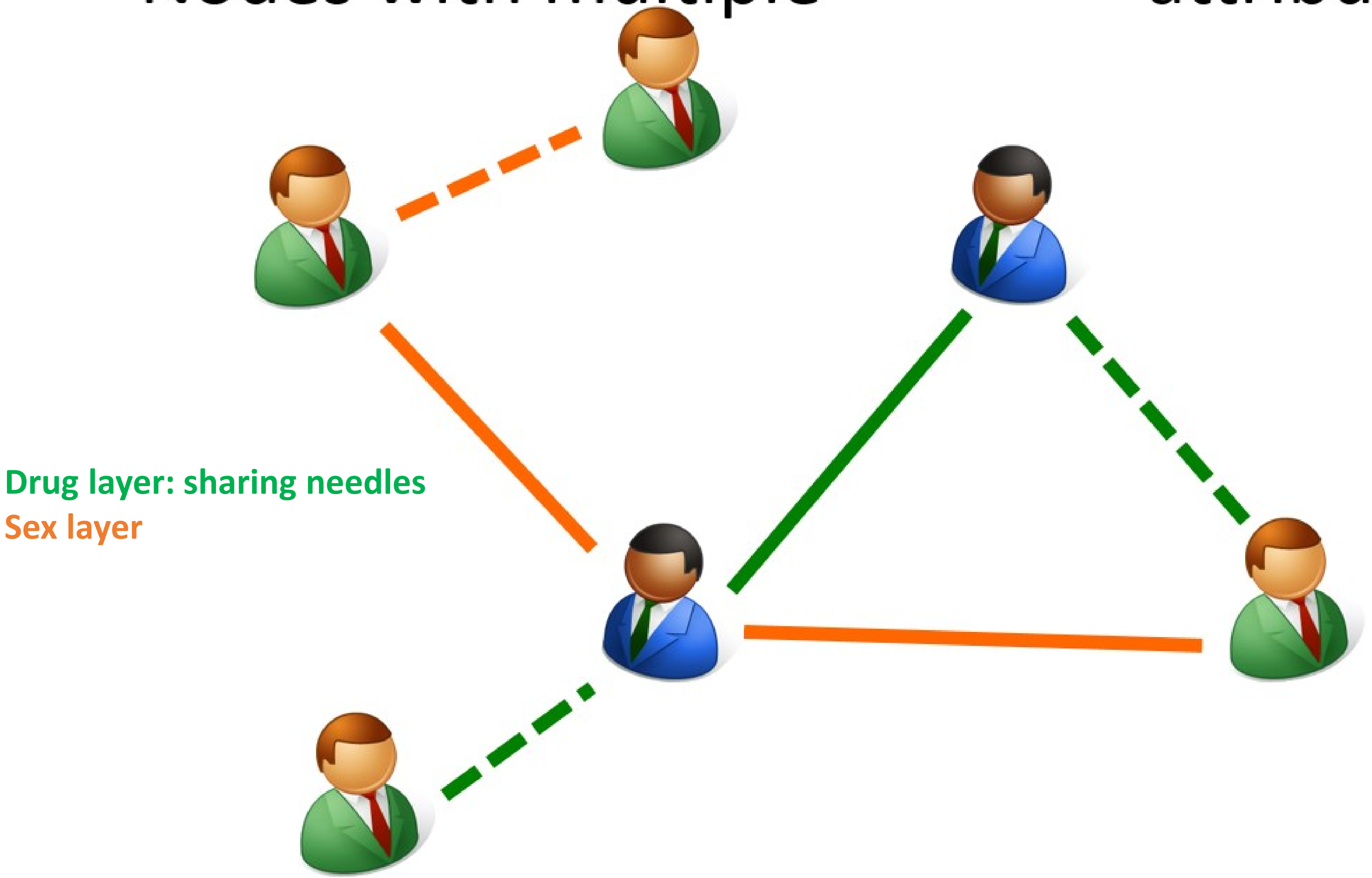
INITIALIZATION

1. Nodes are made (including some infected)
2. Edges are made

SIMULATION RUN

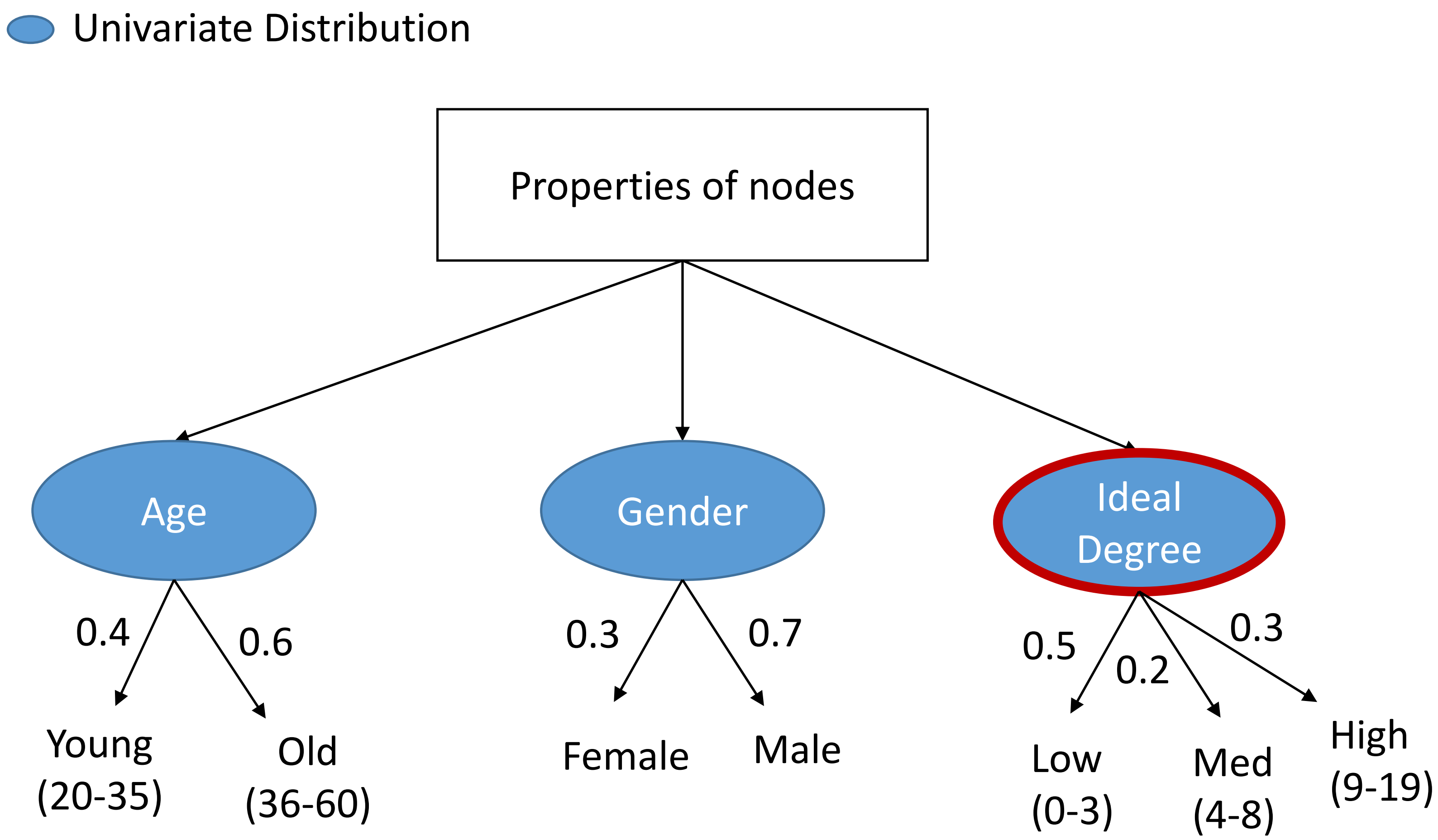
3. Edges change over time
4. Risk acts happen with neighbors
5. Nodes come and go

Nodes with multiple attributes

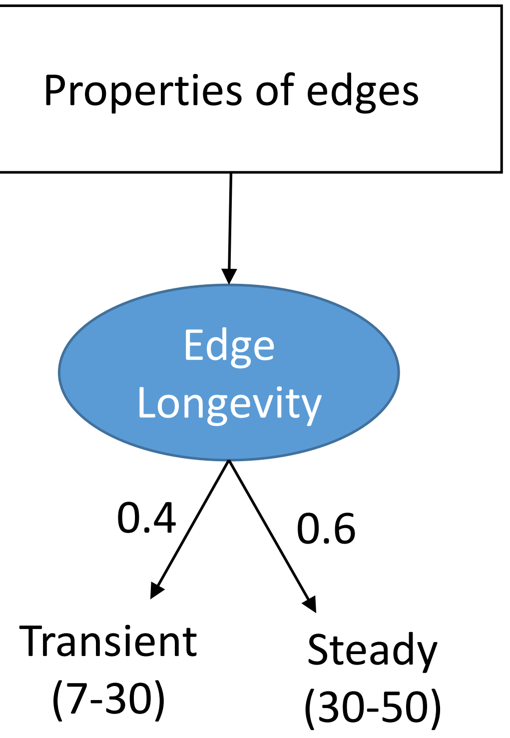


Dynamic edges of different types

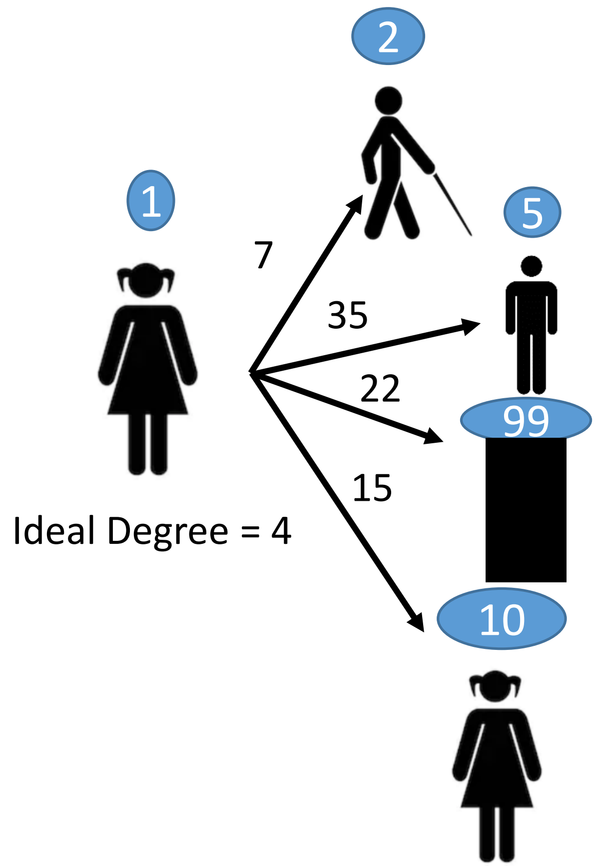
Every node has an "ideal degree"



Every Relationship has a "longevity"

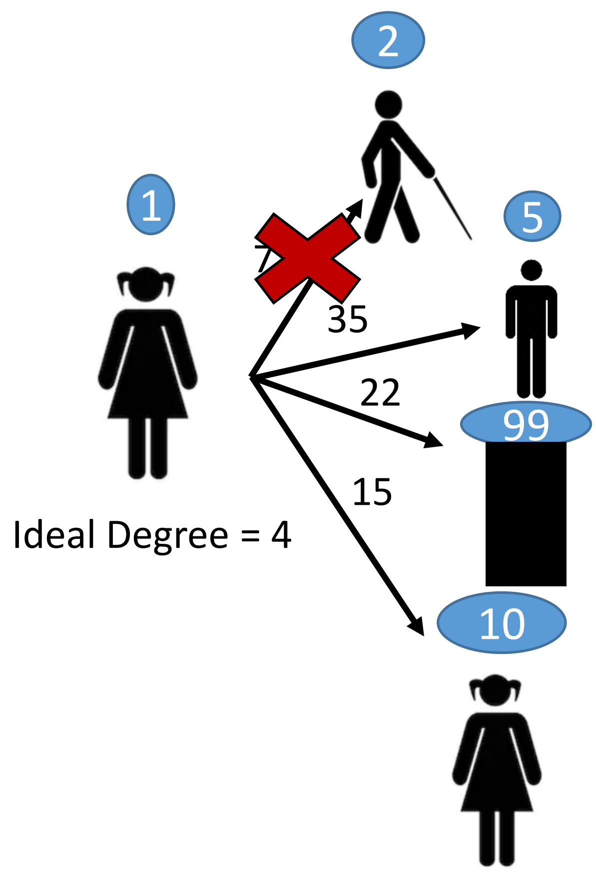


How do network edges change?



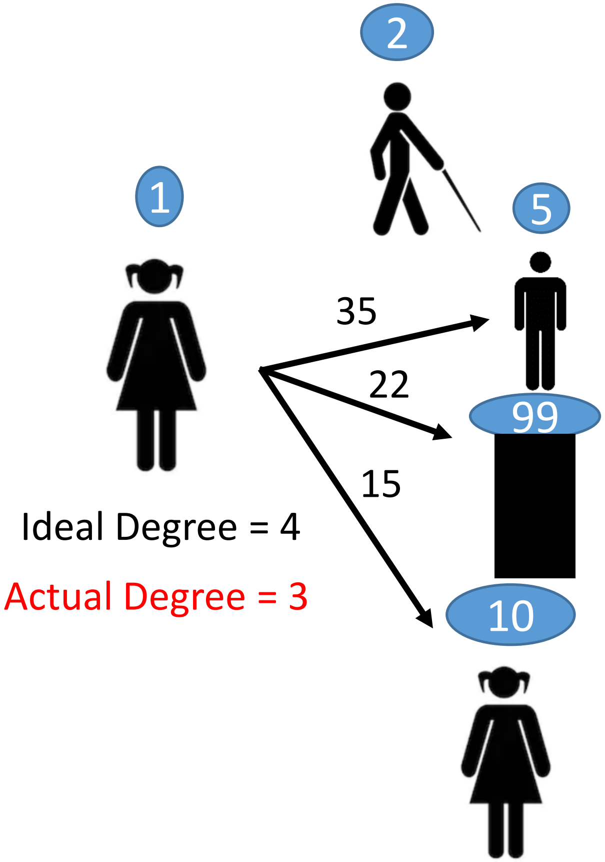
How do network edges change?

Day 7: Relationship 1-2 ends

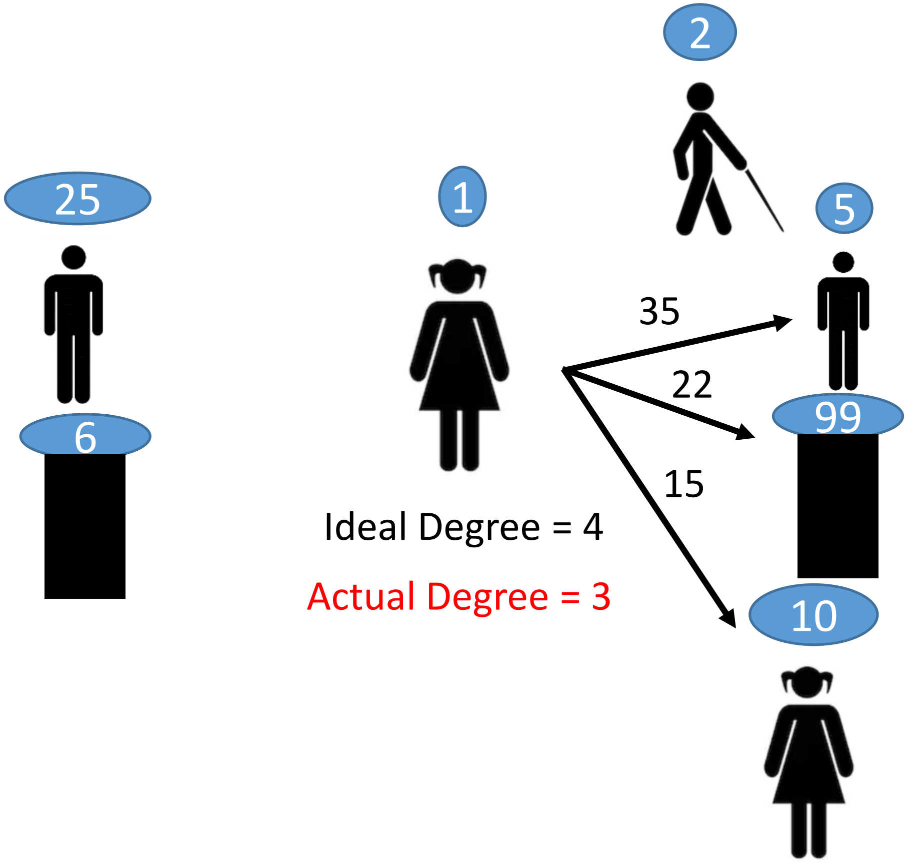


How do network edges change?

Day 7: Relationship 1-2 ends

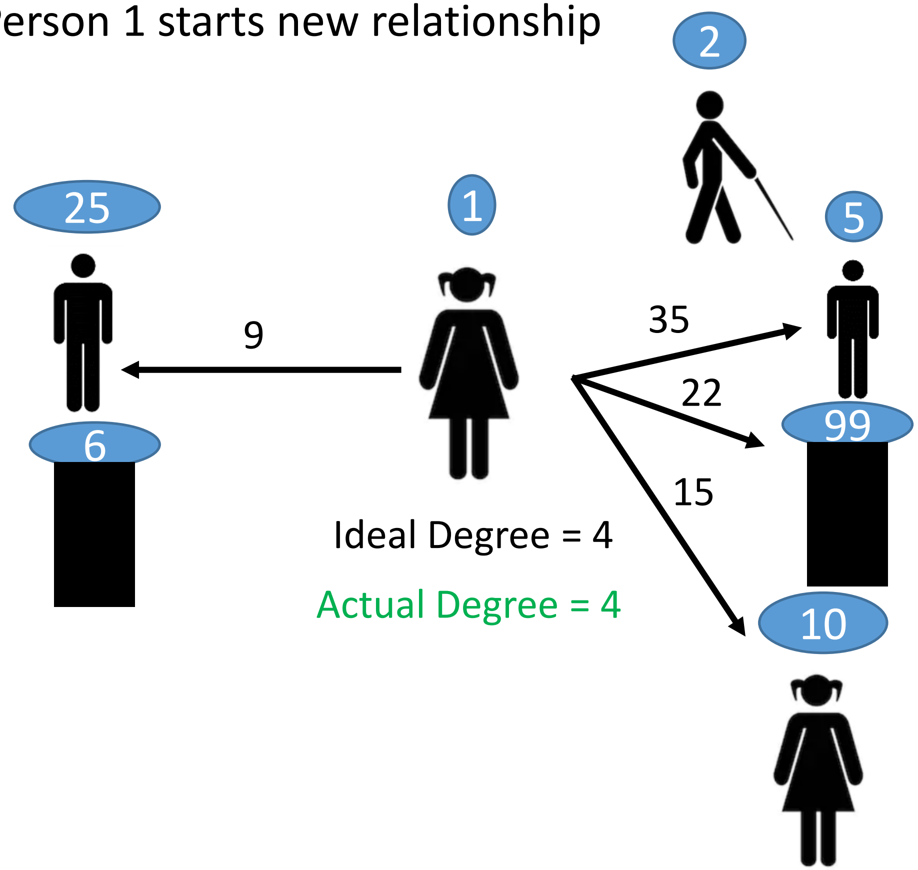


How do network edges change?



How do network edges change?

Person 1 starts new relationship



Outline of MABUSE Model

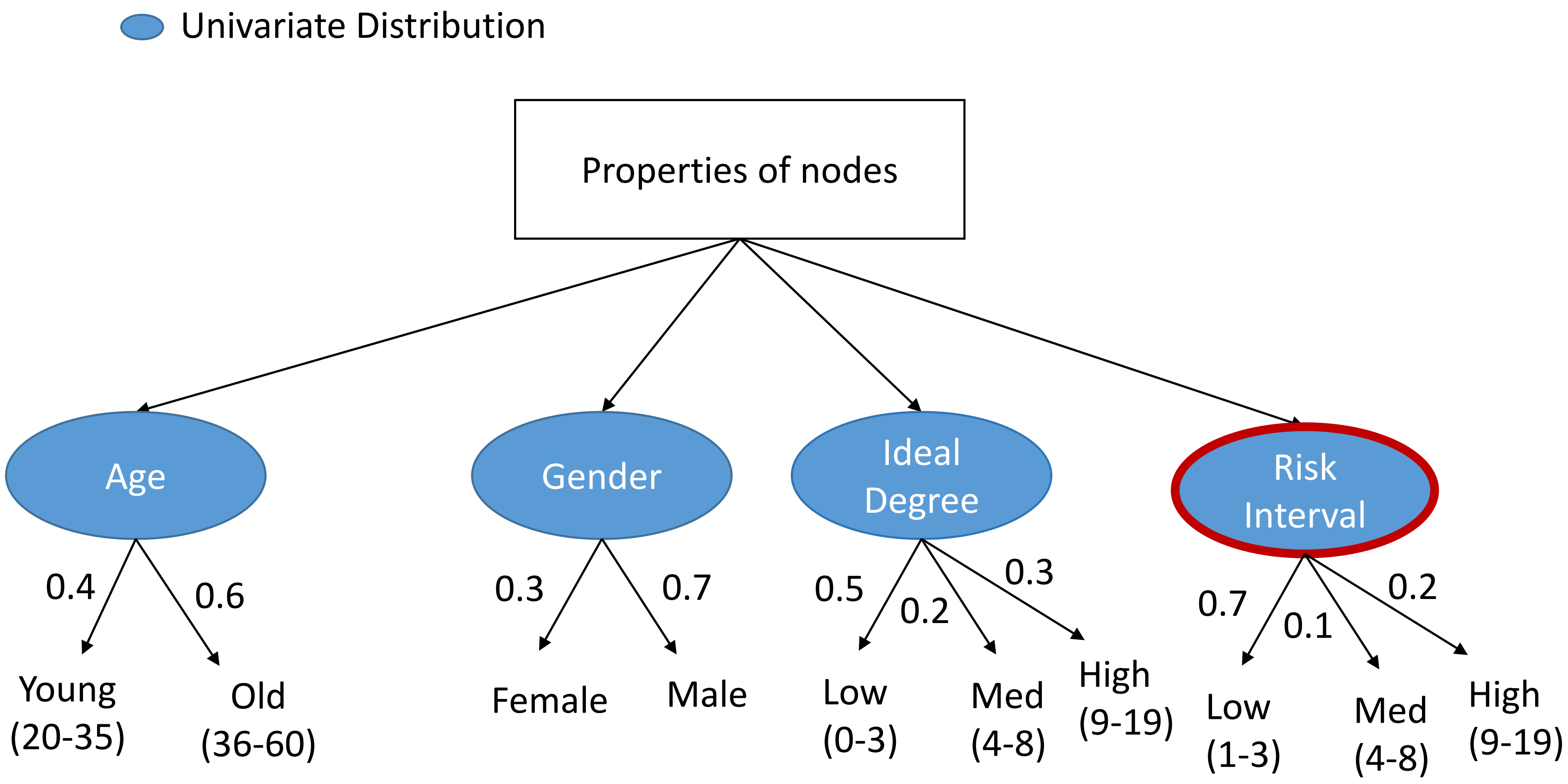
INITIALIZATION

1. Nodes are made (including some infected)
2. Edges are made

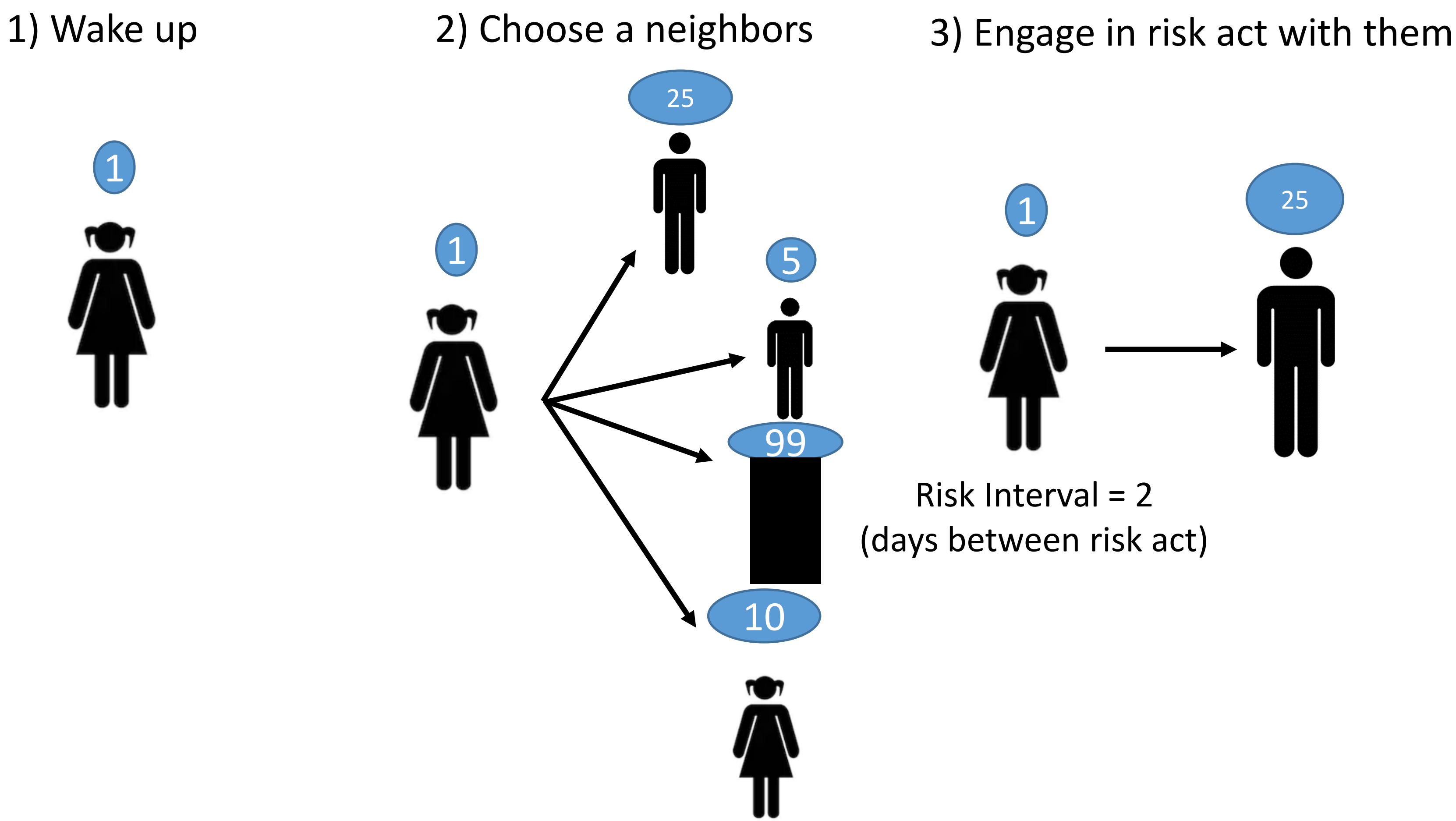
SIMULATION RUN

3. Edges change over time
4. Risk acts happen with neighbors
5. Nodes come and go

Every node has a “risk interval"

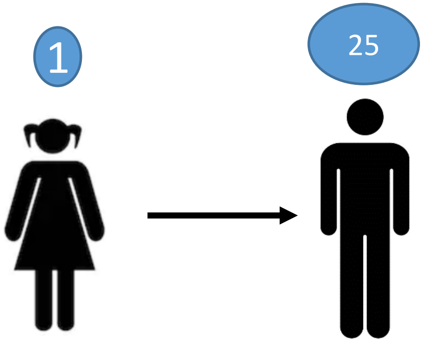


What are the people doing? Risky things

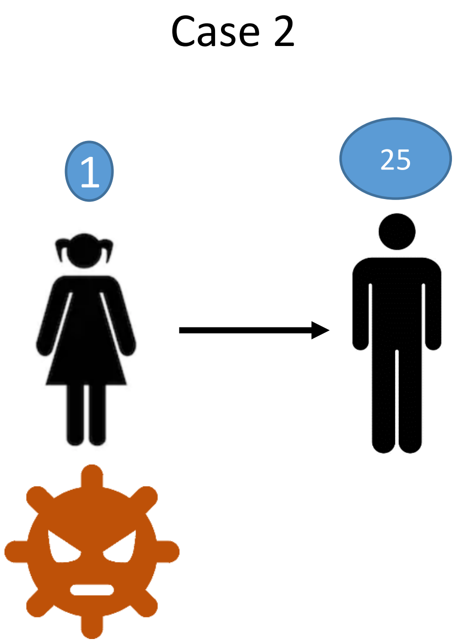


Disease transmission during a risk act

Case 1

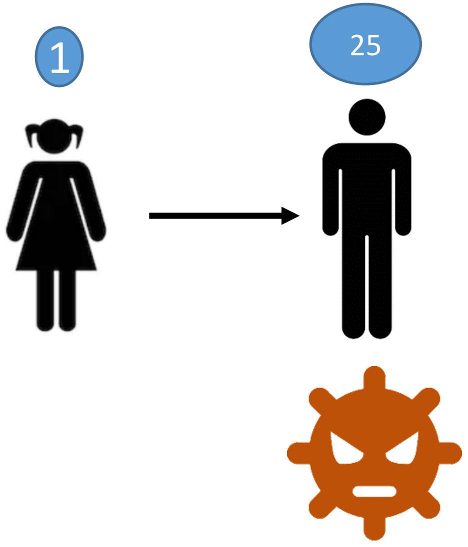


Disease transmission during a risk act



Disease transmission during a risk act

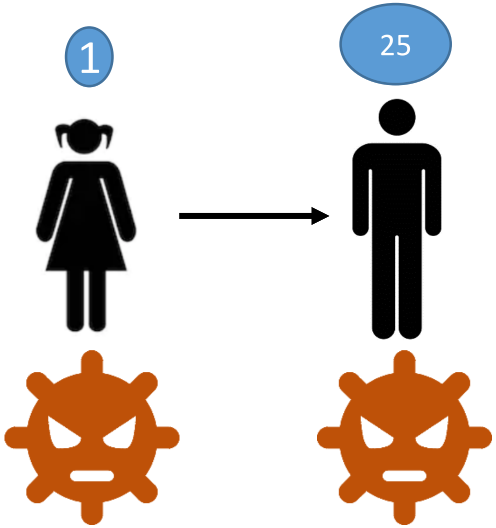
Case 3



Transmission probability

Disease transmission during a risk act

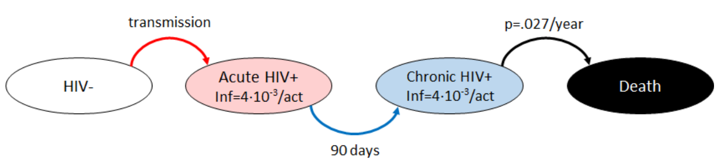
Case 4

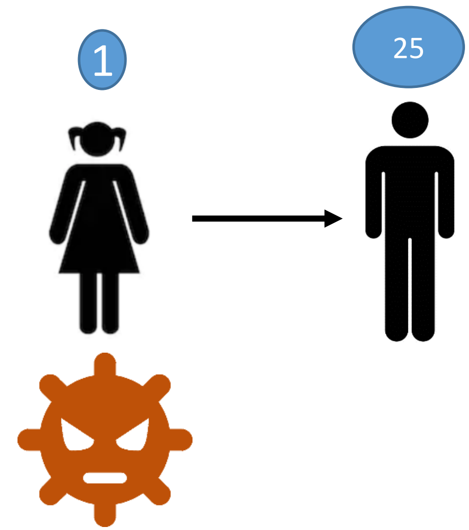


Transmission probability

The "natural history" of disease

HIV



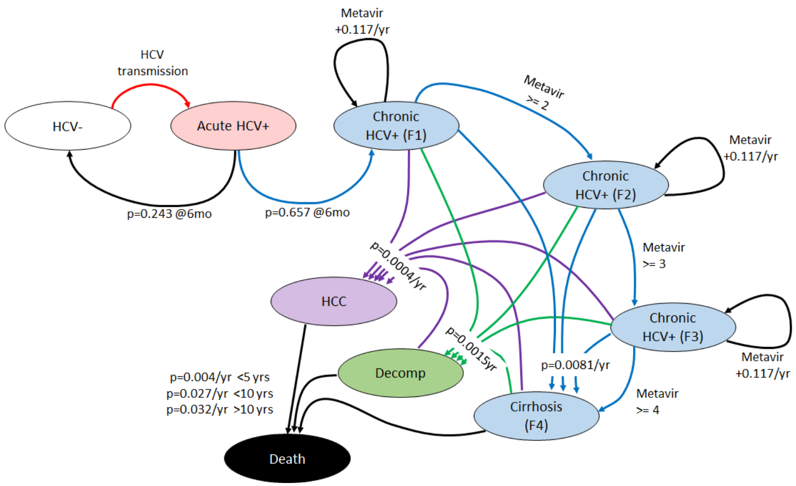


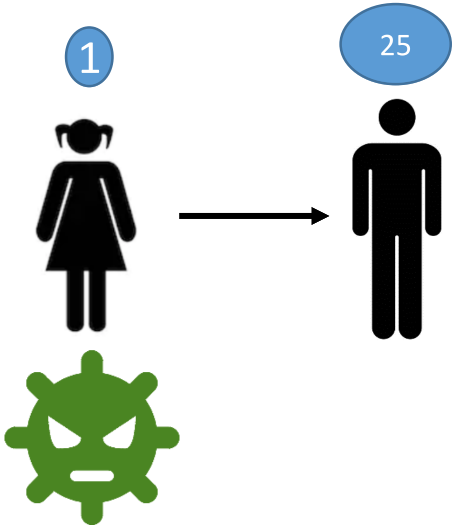
=0.004 in acute
=0.00004 in chronic

Transmission probability

The "natural history" of disease

HCV





Transmission probability

Outline of MABUSE Model

INITIALIZATION

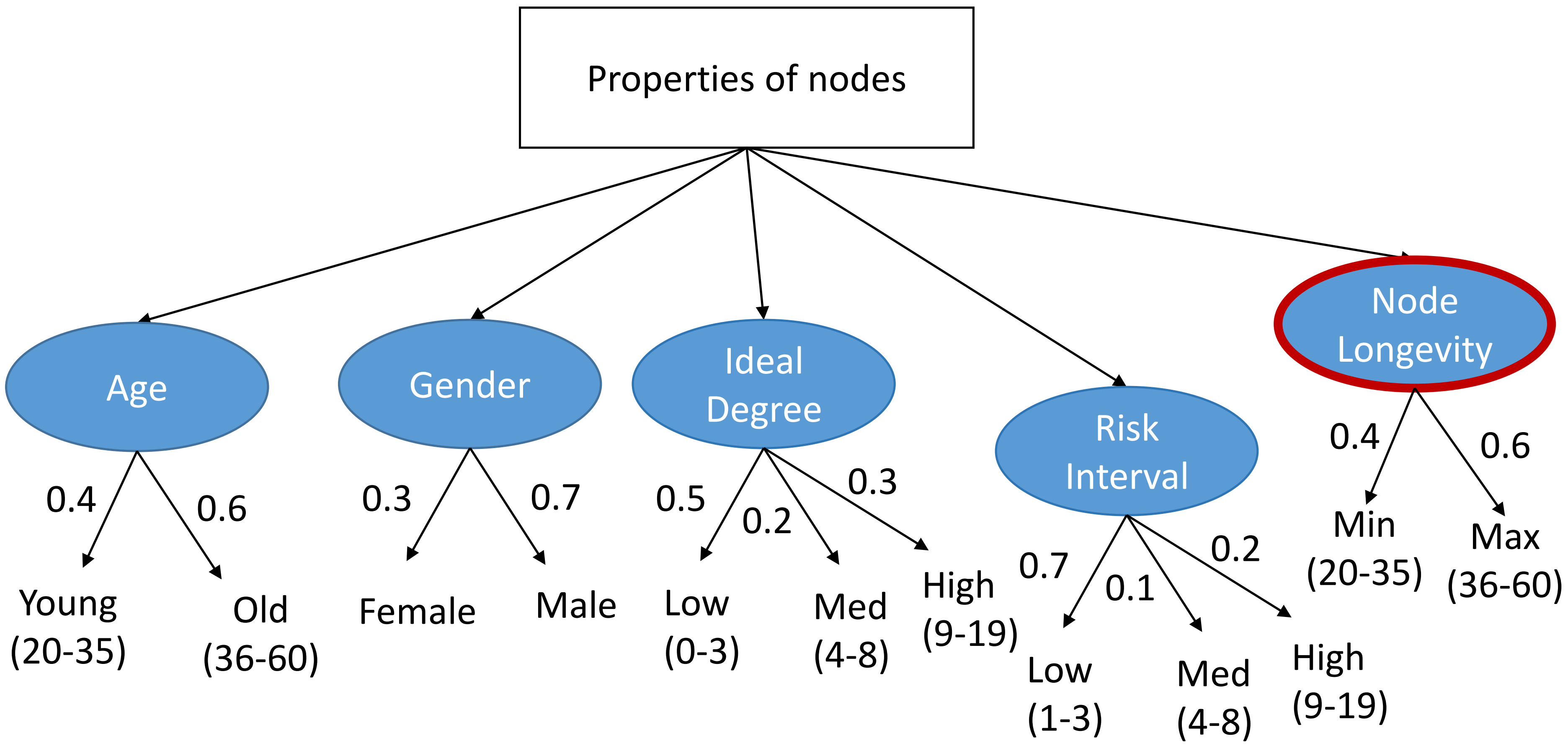
1. Nodes are made (including some infected)
2. Edges are made

SIMULATION RUN

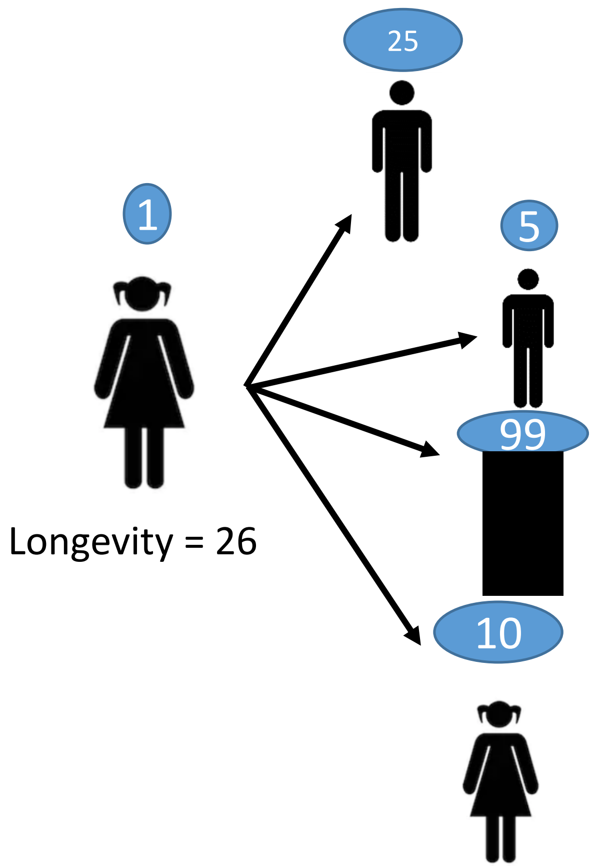
3. Edges change over time
4. Risk acts happen with neighbors
5. Nodes come and go

How are people coming and going?

Univariate Distribution

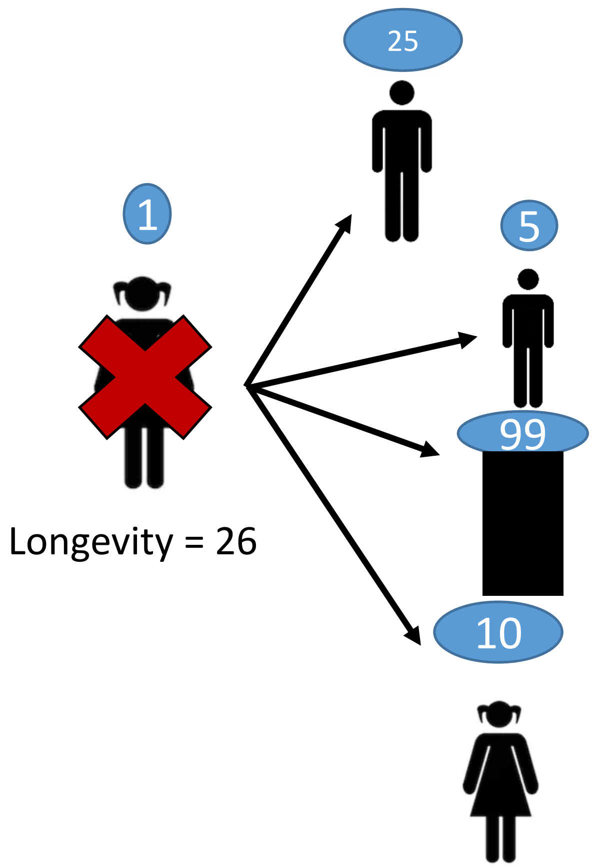


How are people coming and going?



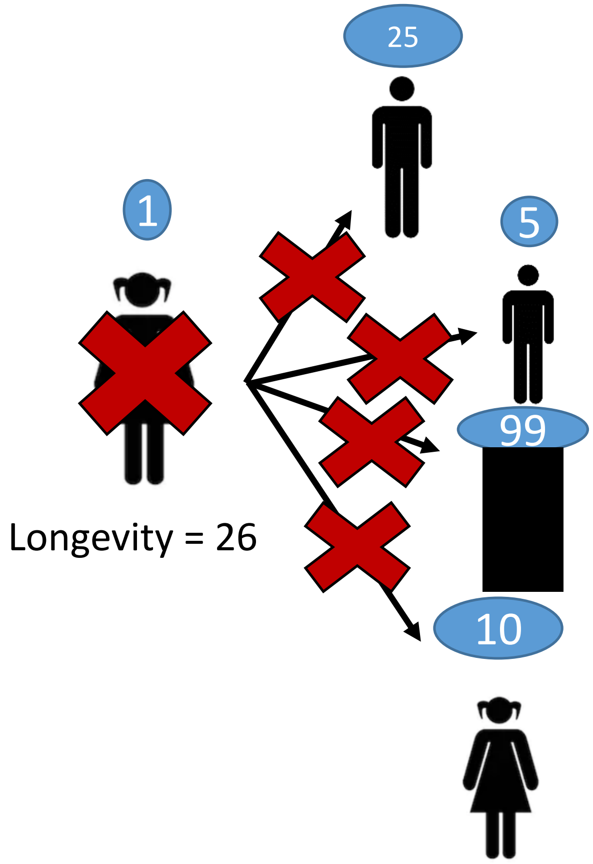
How are people coming and going?

Day 26: Person 1 dies



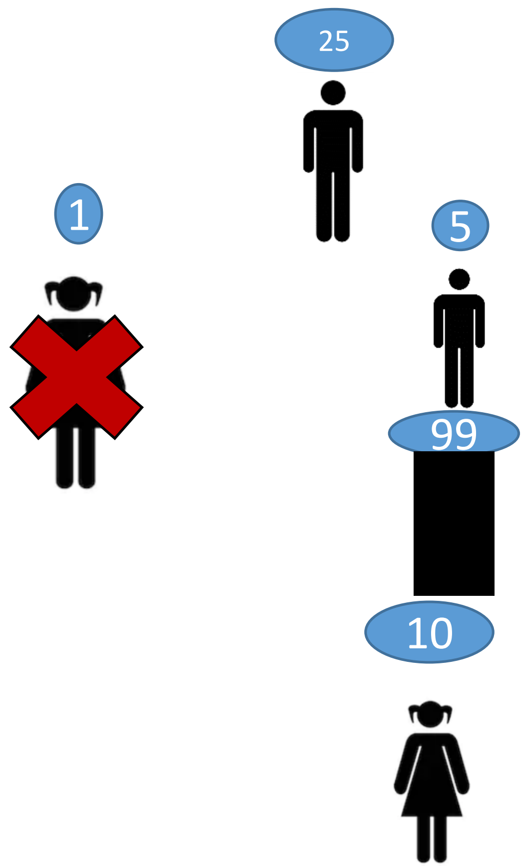
How are people coming and going?

Day 26: Person 1 dies



How are people coming and going?

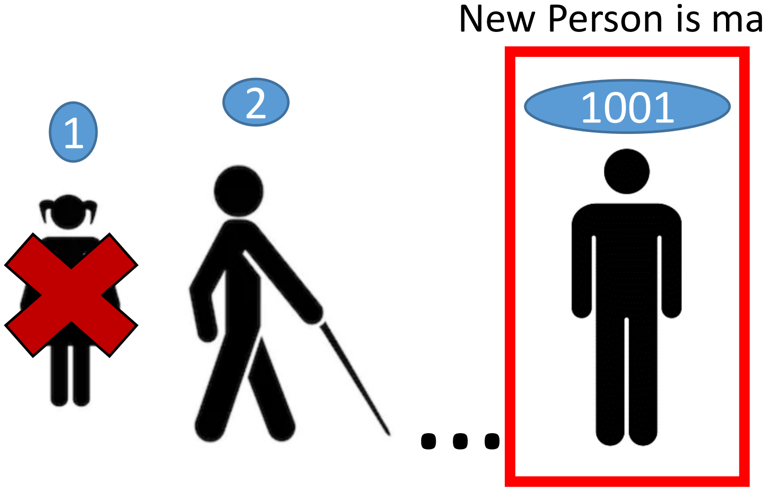
Day 26: Person 1 dies



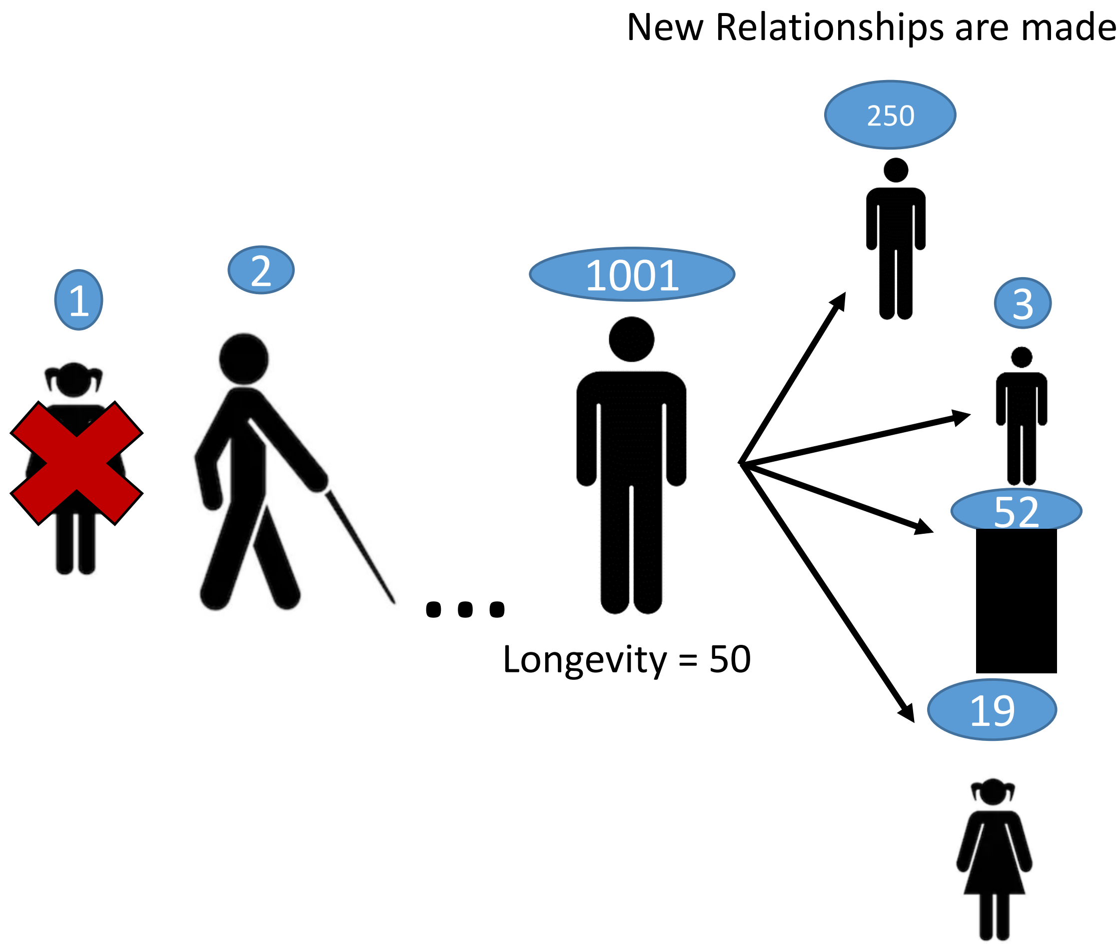
Now all these people have
Actual degree < Ideal degree…

… and the population shrank.

How are people coming and going?



How are people coming and going?



Outline of MABUSE Model

INITIALIZATION

1. Nodes are made (including some infected)
2. Edges are made

SIMULATION RUN

3. Edges change over time
4. Risk acts happen with neighbors
5. Nodes come and go

Now, putting it all together…

One layer, one pathogen scenario

Node Longevity

Minimum (days)	50
Maximum (days)	100

Gender

Male	0.7
Female	0.3

Age

Young	0.4
Old	0.6

Network Layer

Drug layer

Risk Behavior (Per Week)

Low	100
Medium	100
High	100

Ideal Degree (Partners)

Low	0
Medium	1
High	0

Edge Longevity (Years)

Short Max	30
Short Min	1
Long Max	1825
Long Min	30

Transmission probability #1

Multi-layer scenario

Node Longevity

Minimum (days)	50
Maximum (days)	100

Gender

Male	0.7
Female	0.3

Age

Young	0.4
Old	0.6

Network Layer #1 Drug layer

Risk Behavior (Per Week)

Low	100
Medium	100
High	100

Ideal Degree (Partners)

Low	0
Medium	1
High	0

Edge Longevity (Years)

Short Max	30
Short Min	1
Long Max	1825
Long Min	30

Transmission probability #1

Network Layer #2 Sex layer

Risk Behavior (Per Week)

Low	100
Medium	100
High	100

Ideal Degree (Partners)

Low	0
Medium	1
High	0

Edge Longevity (Years)

Short Max	30
Short Min	1
Long Max	1825
Long Min	30

Transmission probability #1

Multi-pathogen scenario

Node Longevity

Minimum (days)	50
Maximum (days)	100

Gender

Male	0.7
Female	0.3

Age

Young	0.4
Old	0.6

Network Layer

Risk Behavior (Per Week)

Low	100
Medium	100
High	100

Ideal Degree (Partners)

Low	0
Medium	1
High	0

Drug layer

Edge Longevity (Years)

Short Max	30
Short Min	1
Long Max	1825
Long Min	30

Transmission probability #1

Transmission probability #1

Multi-pathogen Multi-layer scenario

Node Longevity

Minimum (days)	50
Maximum (days)	100

Gender

Male	0.7
Female	0.3

Age

Young	0.4
Old	0.6

Network Layer #1 Drug layer

Risk Behavior (Per Week)

Low	100
Medium	100
High	100

Ideal Degree (Partners)

Low	0
Medium	1
High	0

Edge Longevity (Years)

Short Max	30
Short Min	1
Long Max	1825
Long Min	30

Transmission probability #1

Transmission probability #1

Network Layer #2 Sex layer

Risk Behavior (Per Week)

Low	100
Medium	100
High	100

Ideal Degree (Partners)

Low	0
Medium	1
High	0

Edge Longevity (Years)

Short Max	30
Short Min	1
Long Max	1825
Long Min	30

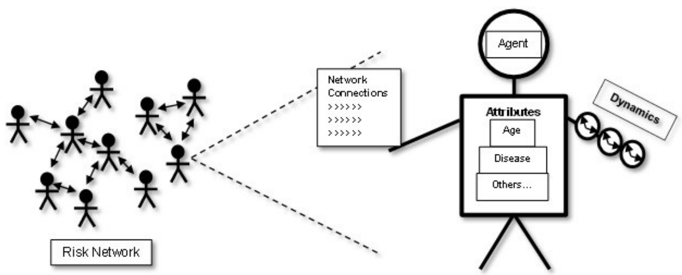
Transmission probability #1

Transmission probability #1

Design of MABUSE

Key features

- Dynamic network simulation
  - Multiple (dynamic) node attributes
  - Multiple (dynamic) edge layers
  - Multiple simultaneous pathogens



- Stochastically defined node agency
  - E.g. risk activity along network edges on different layers is what drives the flow of multiple pathogen types

How is the implementation going to be validated?

- Validate Network Dynamism
- Validate Pathogen Flow

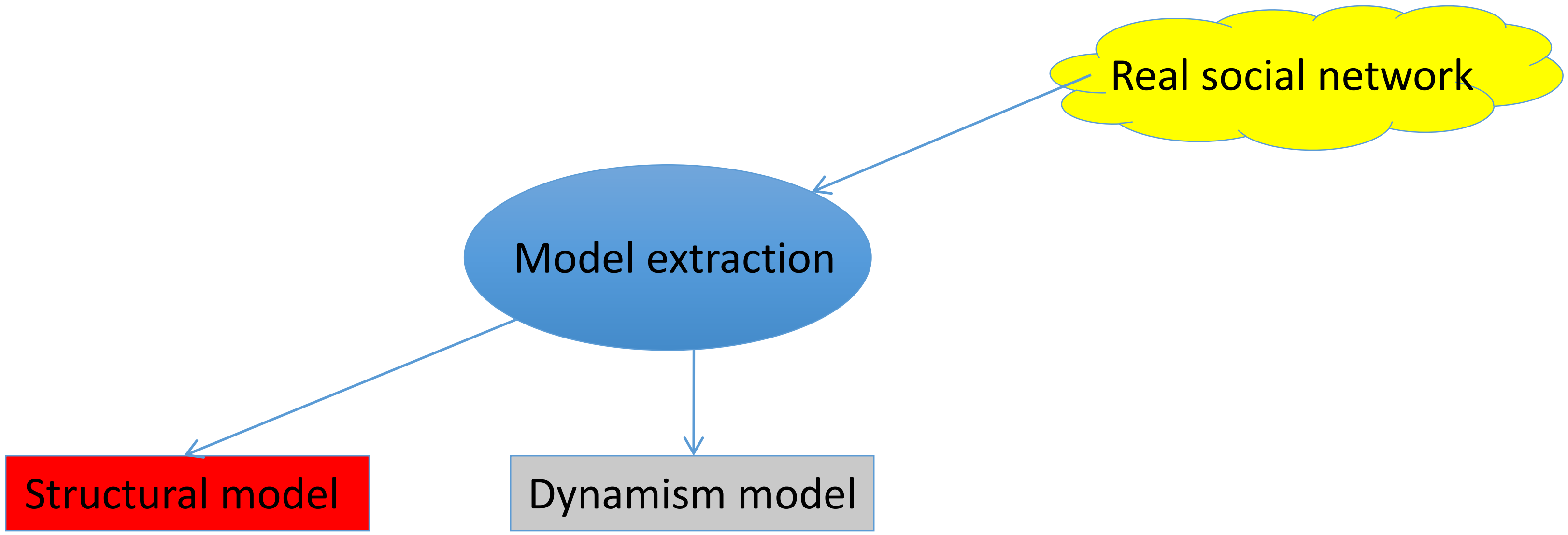
Validation of Network Simulation

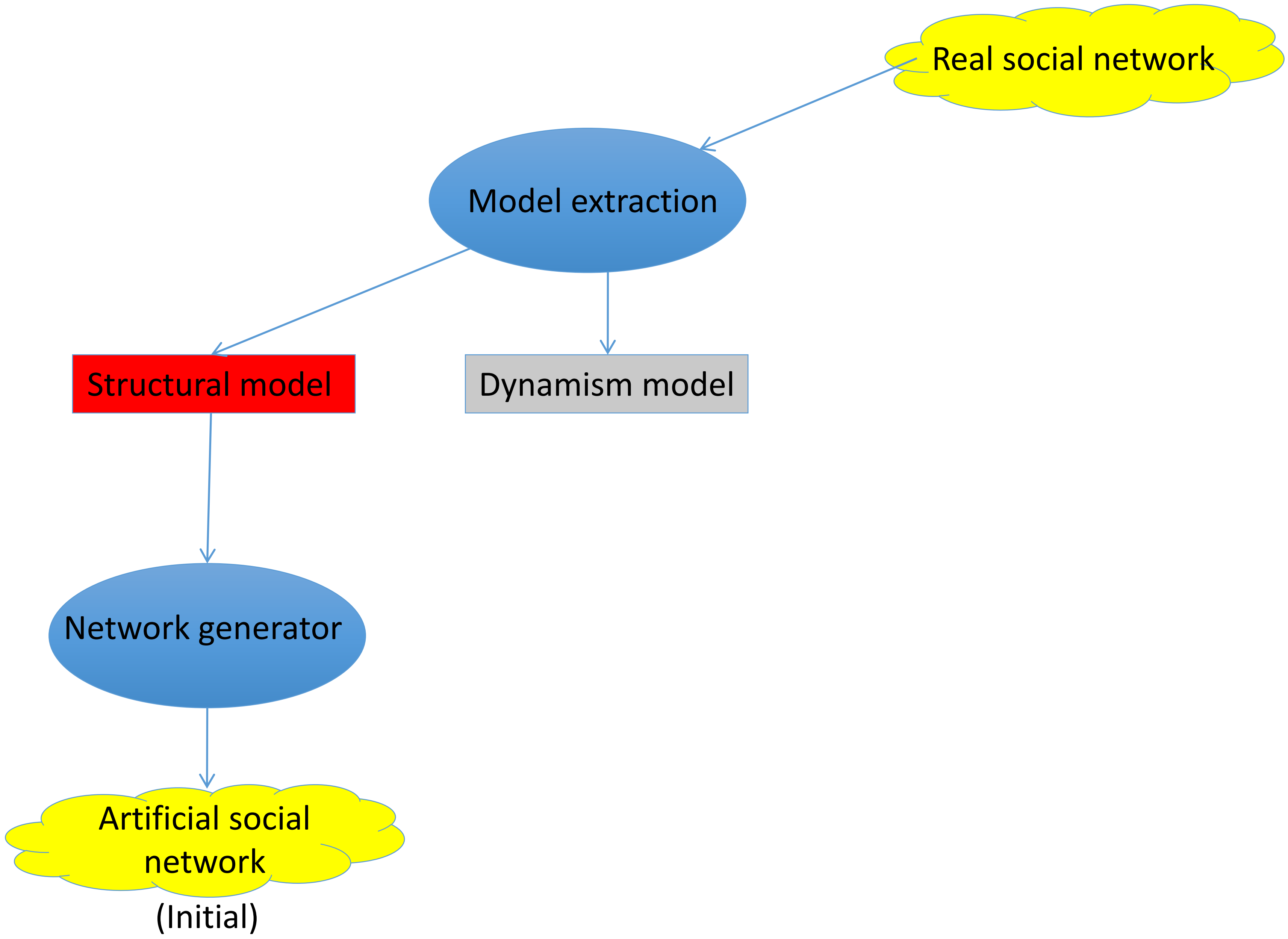
1. Take a real network. Compute statistics
   A. Structure (univariates and bivariates)
   B. Dynamism (ideal degrees, longevity)
2. Use 1A+1B to do a 10 year simulation.
   At the end, Compute statistics
   A. Structure (univariates and bivariates)
   B. Dynamism (ideal degrees, longevity)

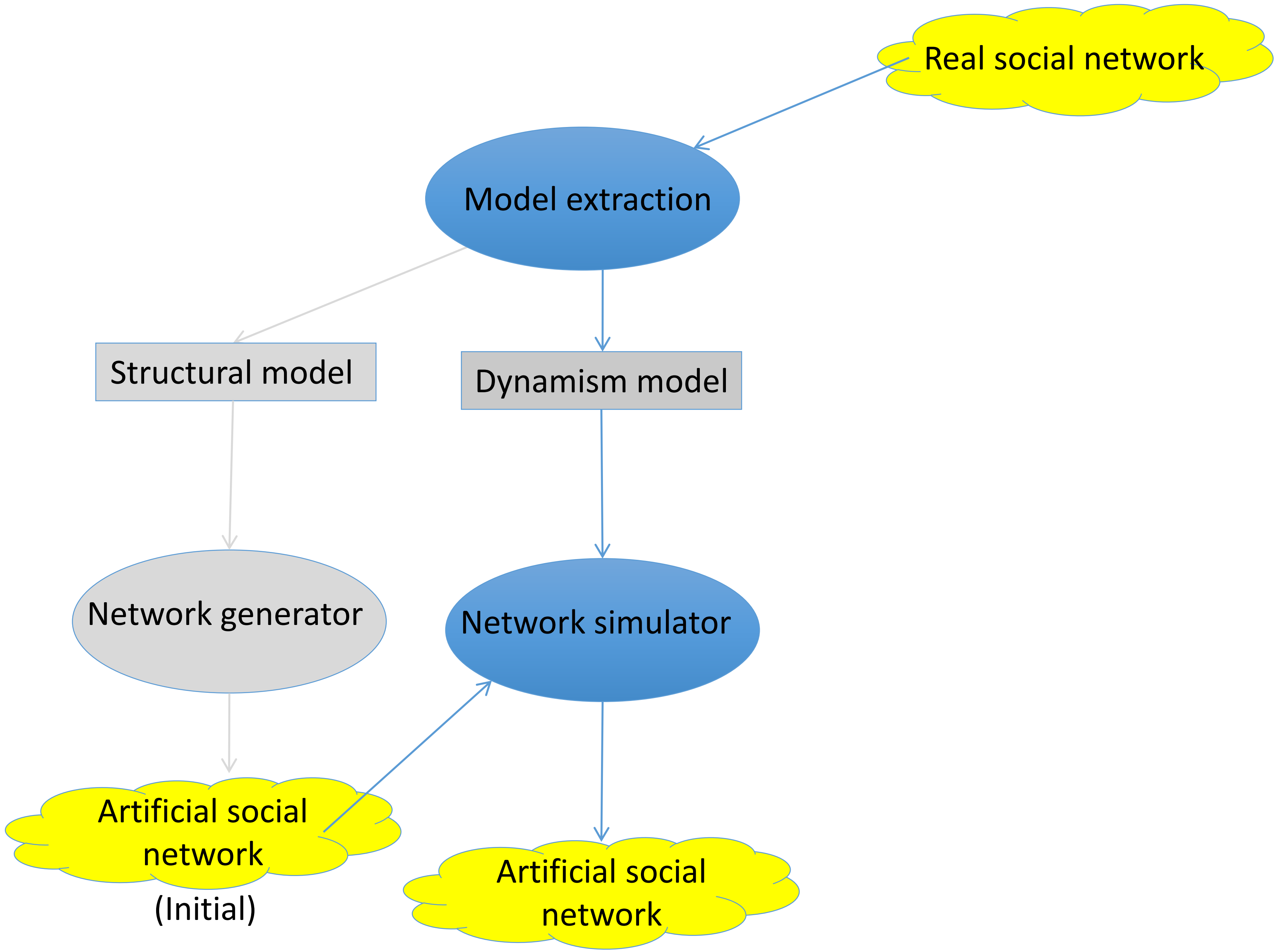
Is 1A "close" to 2A?

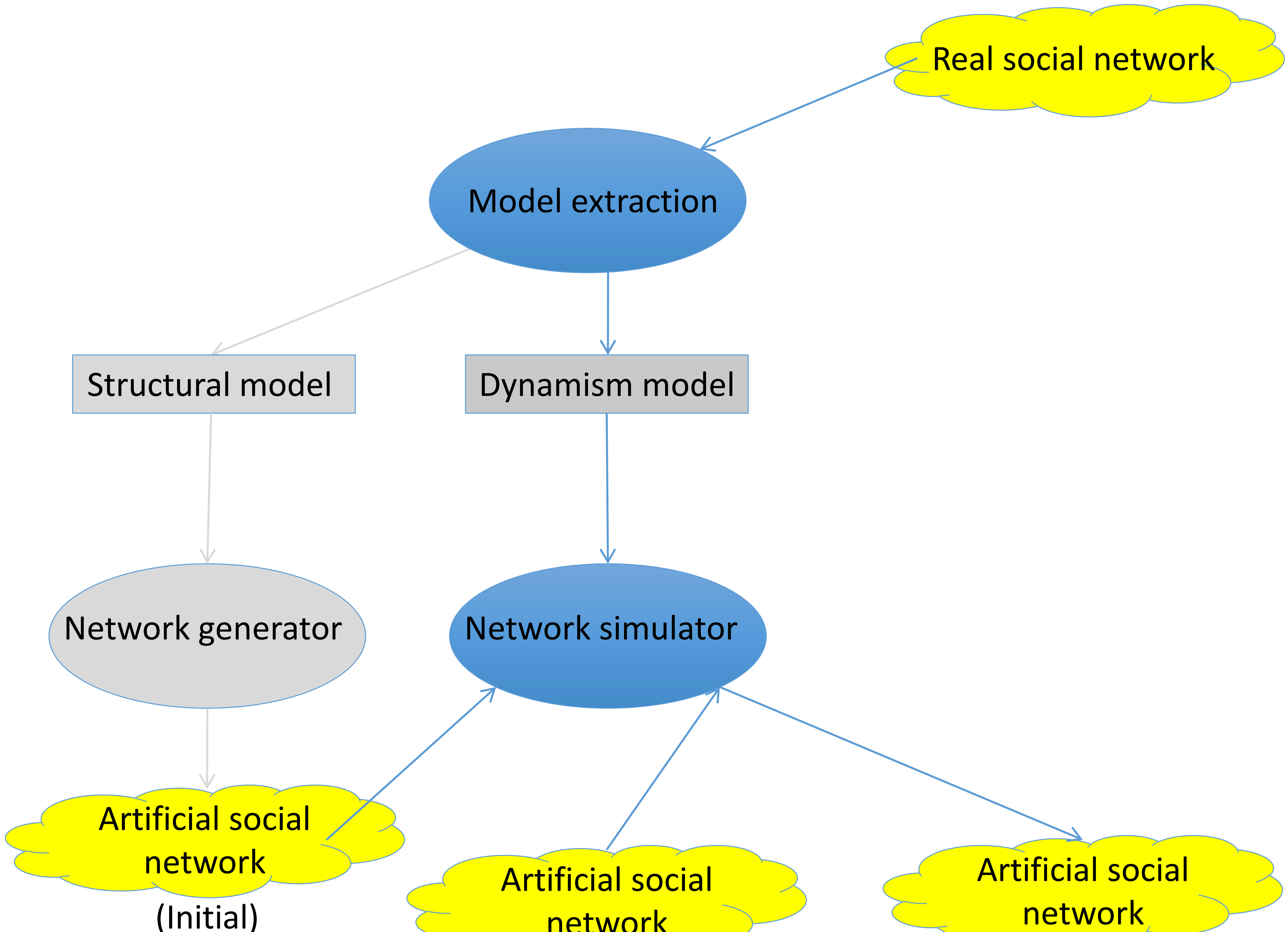
Is 1B "close" to 2B?

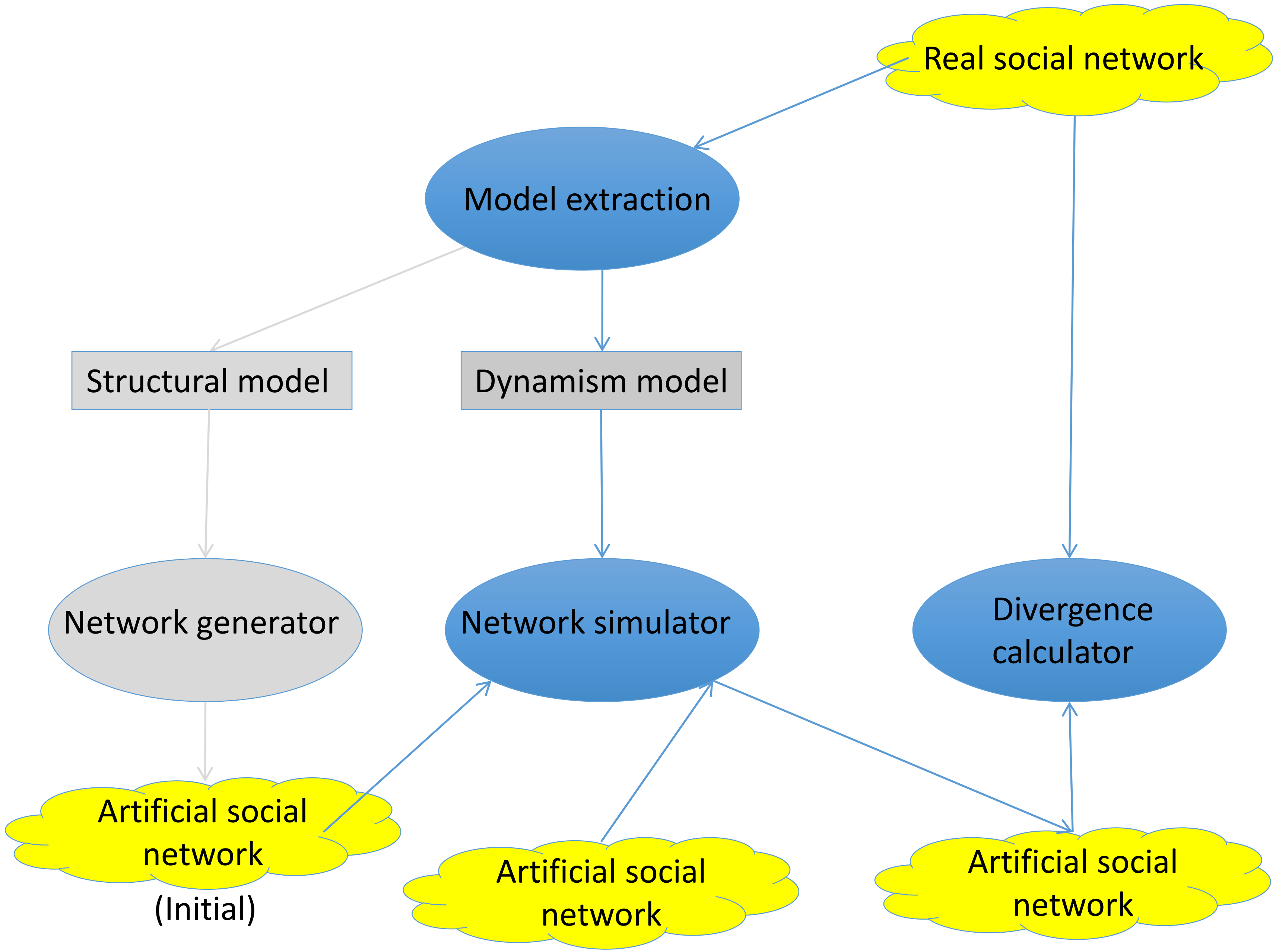
Use Shannon-Jensen divergence to measure the "distance" between distributions


Real social network
Model extraction
Structural model
Dynamism model


Real social network
Model extraction
Structural model
Dynamism model
Network generator
Artificial social network
(Initial)


Real social network
Model extraction
Structural model
Dynamism model
Network generator
Network simulator
Artificial social network
(Initial)
Artificial social network

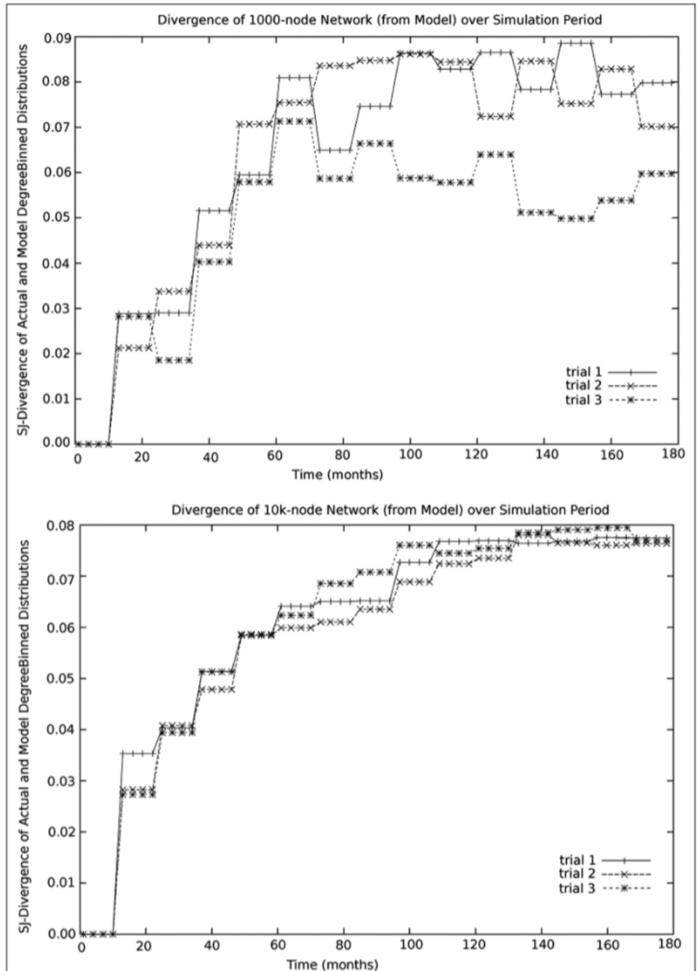

Real social network
Model extraction
Structural model
Dynamism model
Network generator
Network simulator
Artificial social network
(Initial)
Artificial social network
Artificial social network


Real social network
Model extraction
Structural model
Dynamism model
Network generator
Network simulator
Divergence calculator
Artificial social network
(Initial)
Artificial social network
Artificial social network

Results

Shannon-Jensen divergence

Shannon-Jensen divergence



Time

Validation of Pathogen

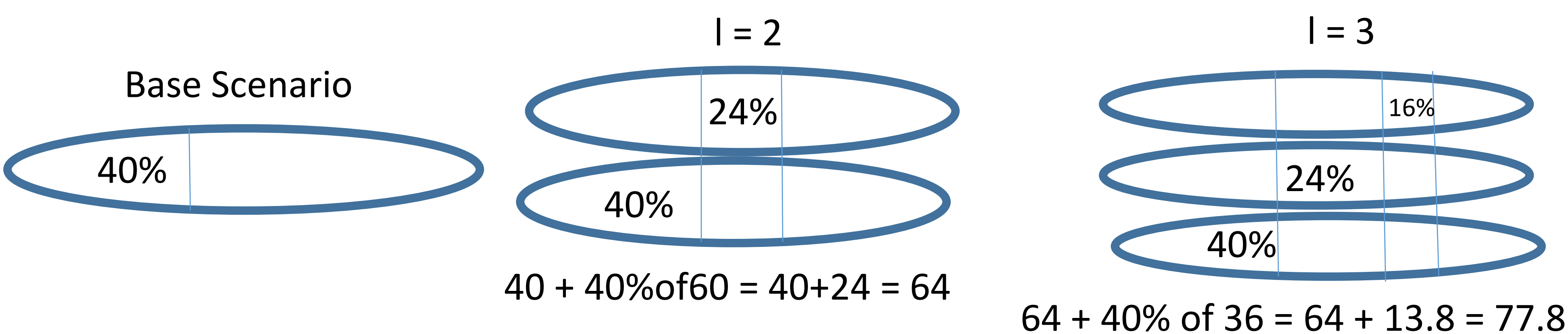
- Take Base Scenario
  - Single Layer
  - Single Pathogen

- Compare Base Scenario with the following Artificial Scenarios

Artificial Scenario 1

Table 5. Scenario 1: Multiple layers, one pathogen.

Number of Layers	Pathogen 1 prevalence at 60 months (%)
$l = 2$ (and $p = 1$)	70
$l = 3$ (and $p = 1$)	81
$l = 4$ (and $p = 1$)	89
Base scenario ($p = 1$ and $l = 1$)	42



Artificial Scenario 2

Table 6. Scenario 2: Multiple pathogens, one layer.

Number of pathogens	Average prevalence at 60 months (across all p pathogens) (%)	Ave. pairwise correlation of pathogen occurrence
$p = 2$ (and $l = 1$)	42	0.86
$p = 3$ (and $l = 1$)	43	0.76
$p = 4$ (and $l = 1$)	44	0.72
Base scenario ($p = 1$ and $l = 1$)	42	

Artificial Scenario 3

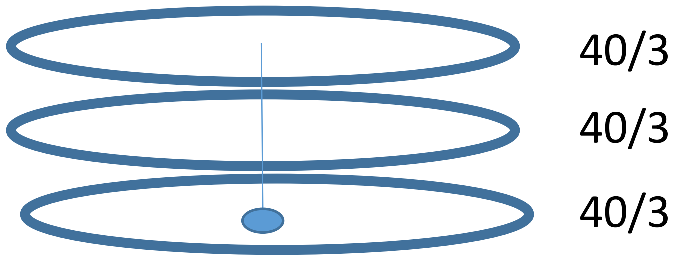
Table 7. Scenario 3: Multiple layers, multiple non-interacting pathogens.

Number of pathogens p Number of layers l	Average prevalence at 60 months (across all p pathogens) (%)	Ave. pairwise correlation of pathogen occurrence
$p = l = 2$	42	0.12
$p = l = 3$	43	−0.11
$p = l = 4$	42	0.08
Base scenario ($p = l = 1$)	42	

Artificial Scenario 4

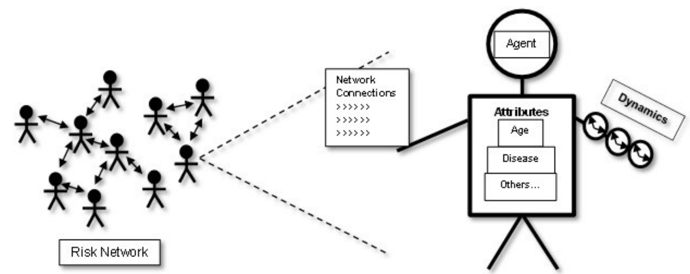
Table 8. Scenario 4: Multiple layers, multiple interacting pathogens.

Number of Pathogens p Number of Layers l	Average prevalence at 60 months (across all p pathogens)(%)
$p = l = 2$	22
$p = l = 3$	15
$p = l = 4$	12
Base Scenario ($p = l = 1$)	42



Multi Actor-Based Universal Simulation Engine

- MABUSE is stochastic, agent-based, discrete event simulator for epidemics in dynamic networks.



- The simulator runs on specialized high-performance hardware and is capable of running simulations of dynamic networks having 1,000,000+ nodes, and handling approximately 800 million discrete events per second.

Conclusion

- MABUSE allows us to simulate
  - Multiple (dynamic) node attributes
  - Multiple (dynamic) edge layers
  - Multiple simultaneous pathogens
- A realistic dynamic network
  - From the POV of an agent (micro, ABM-style)
  - From POV of the network (macro, ERGM-style)
- Model where each agent takes into account its own properties and the properties of other agents to make connections, and act on those connections

Thank you

Any Questions?

Jensen-Shannon Divergence

"Real"

D_1:	M	55%
	F	45%

"Simulation @ 30"

D_1:	M	75%
	F	25%

Mid:	M	65%
	F	35%

$$SJ(D1, D2) = \frac{1}{2}[KL(D1, M) + KL(D2, M)]$$

$$\text{KL}(D1, M) = \sum_{i=M,F} D1(i) * \log\left(\frac{D1(i)}{Mid(i)}\right) = 0.55 * \log\left(\frac{0.55}{0.65}\right) * 0.45 * \log\left(\frac{0.44}{0.35}\right)$$