# Comparing Networks by their Group Structure

## with an application to acculturation networks

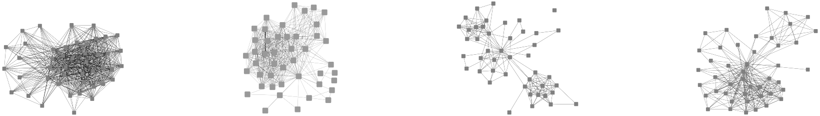
Jürgen Lerner    Ulrik Brandes

(University of Konstanz)

XXVII Sunbelt'07    Corfu, Greece    May 1–6, 2007

# Short Outline

suppose we have several personal networks ...



... and want to determine their differences and similarities

- number of actors, ties, or connected components, degree distribution, diameter
- reciprocity, transitivity, clustering coefficient
- isomorphic, edit distance, common subgraphs

this talk: compare networks by their group structure

# Overview

## Comparing Networks by their Group Structure

1. **example application:**
   **towards a network measure for acculturation**
2. defining class-level networks
3. average class-level networks
4. conclusion and future work

# Acculturation

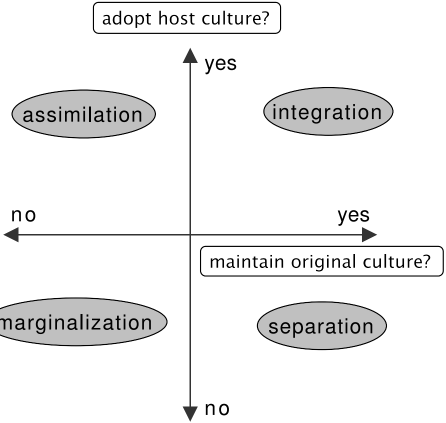
traditional usage: outcome of cultures coming into contact

recent usage: measuring the level of integration of **migrants** into a host culture

# Traditional Acculturation Scale (ARSMA II)

[Cuéllar/Arnold/Maldonado'95] modes of acculturation [Berry'97]

- I speak English
- I speak Spanish
- I associate with Anglos
- I associate with Mexicans
- I enjoy English language TV
- I enjoy Spanish language TV
- My friends now are of Anglo origin
- My friends now are of Mexican origin
- . . .

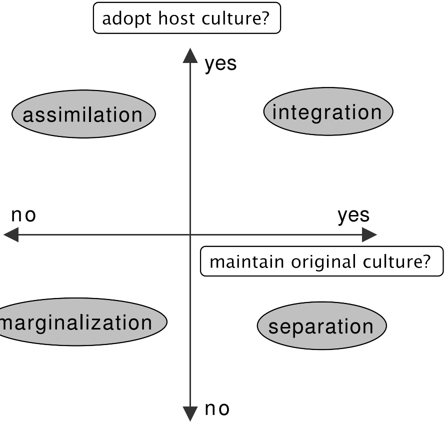



▸ influence of **personal networks**?

# Traditional Acculturation Scale (ARSMA II)

[Cuéllar/Arnold/Maldonado'95] modes of acculturation [Berry'97]

- I speak English
- I speak Spanish
- **I associate with Anglos**
- **I associate with Mexicans**
- I enjoy English language TV
- I enjoy Spanish language TV
- **My friends now are of Anglo origin**
- **My friends now are of Mexican origin**
- . . .




- influence of **personal networks**?

# Data Set `www.egoredes.net`

approximately 500 respondents (**ego**s):

1. **(questions about ego)**
   age, years of residence, health, ARSMA II, . . .
2. **(name generator)** list of 45 alters
3. **(questions about alters)** born, lives, type of relation, . . .
4. **(ties)** which alters know each other

**this talk**: differences and similarities between these networks (comparison on the **class level**)

special thanks for providing the data goes to the project
**Acculturation & Personal Networks across cultures**
$\Longrightarrow$ see talks by Javier Avila, Miranda Lubbers, Chris McCarty, and José Luis Molina on Saturday

# Overview

## Comparing Networks by their Group Structure

1. example application:
   towards a network measure for acculturation
2. **defining class-level networks**
3. average class-level networks
4. conclusion and future work

# Definition of Class-Level Network

two steps

1. defining actor classes, dependent on
   - network structure (e. g., structural/regular equivalence)
   - **actor attributes** (yields labeled classes)

   classes become nodes

2. defining inter-class ties
   (how strongly are two classes connected?)

# Definition of Class-Level Network

two steps

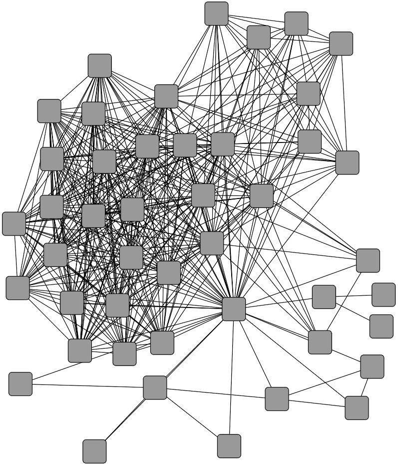
1. defining actor classes, dependent on
   - network structure (e. g., structural/regular equivalence)
   - **actor attributes** (yields labeled classes)

   classes become nodes

2. defining inter-class ties
   (how strongly are two classes connected?)

# defining actor classes

## argentinean woman living in spain



considering *country of origin* and *country of residence* (4 selected classes)

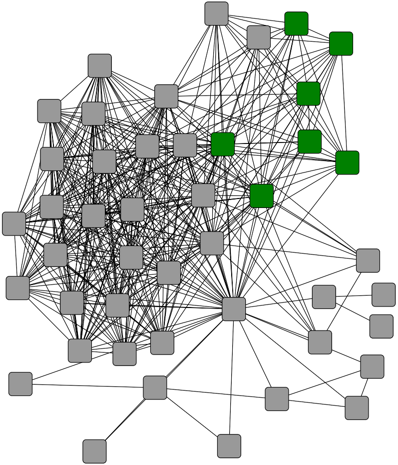
**(origin)** Argentineans living in Argentina

**(fellows)** Argentineans living in Spain

**(host)** Spaniards living in Spain

# defining actor classes

## argentinean woman living in spain



considering *country of origin* and *country of residence* (4 selected classes)

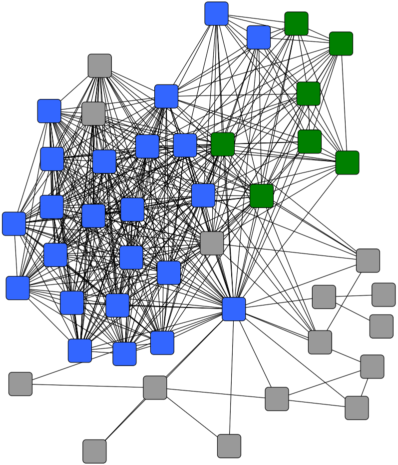
**(origin)** Argentineans living in Argentina

**(fellows)** Argentineans living in Spain

**(host)** Spaniards living in Spain

# defining actor classes

## argentinean woman living in spain



considering *country of origin* and *country of residence* (4 selected classes)

**(origin)** Argentineans living in Argentina

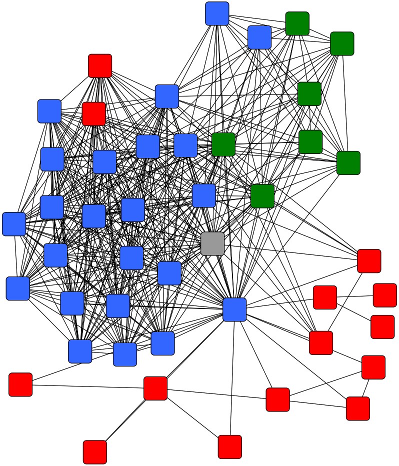
**(fellows)** Argentineans living in Spain

**(host)** Spaniards living in Spain

**(transnationals)** all other cases

# defining actor classes

## argentinean woman living in spain



considering *country of origin* and *country of residence* (4 selected classes)

**(origin)** Argentineans living in Argentina

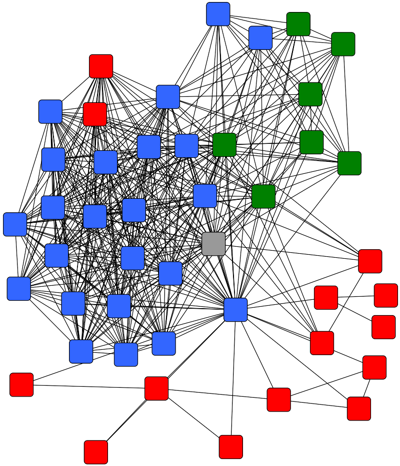
**(fellows)** Argentineans living in Spain

**(host)** Spaniards living in Spain

**(transnationals)** all other cases

# defining actor classes

## argentinean woman living in spain



considering *country of origin* and *country of residence* (4 selected classes)

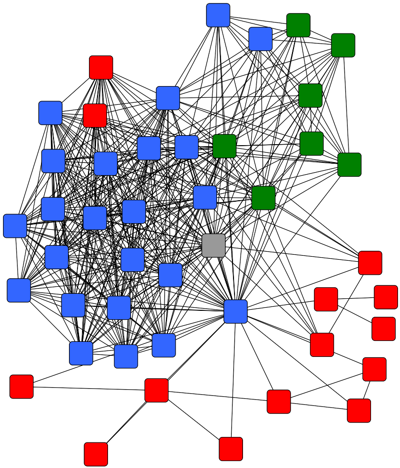
**(origin)** Argentineans living in Argentina

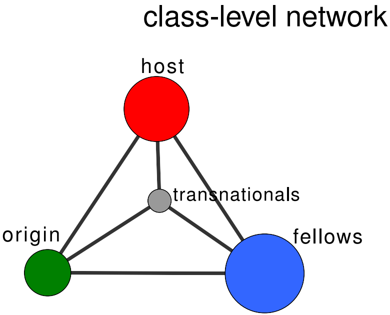
**(fellows)** Argentineans living in Spain

**(host)** Spaniards living in Spain

**(transnationals)** all other cases

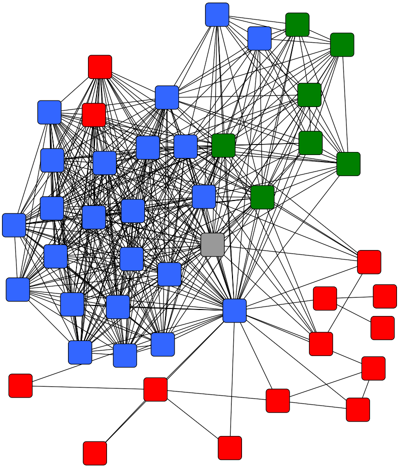
# actor classes become nodes



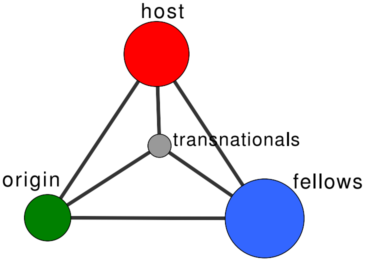



edge weights?

# actor classes become nodes



class-level network




**edge weights?**

# Normalization of Edge-Weights

## how strongly are two classes connected?

given: network $G = (V, E)$ and two actor classes $A$ and $B$

1. **(un-normalized count)** $\#\{(a, b) \in E\,;\ a \in A,\ b \in B\}$
   $\Rightarrow$ *larger classes will be stronger connected*
2. **(density)** $\frac{\#\{(a,b) \in E\,;\ a \in A\ b \in B\}}{|A| \cdot |B|}$
   $\Rightarrow$ *tends to zero when class sizes increase (sparsity)*
3. **(avg. number of $B$-neighbors)** $\frac{\#\{(a,b) \in E\,;\ a \in A\ b \in B\}}{|A|}$
   $\Rightarrow$ *asymmetric class-level network*
4. **(symmetric normalization)** $\frac{\#\{(a,b) \in E\,;\ a \in A\ b \in B\}}{\sqrt{|A| \cdot |B|}}$
   $\Rightarrow$ this is what we take

# Normalization of Edge-Weights

## how strongly are two classes connected?

given: network $G = (V, E)$ and two actor classes $A$ and $B$

1. **(un-normalized count)** $\#\{(a, b) \in E\,;\; a \in A,\ b \in B\}$
   $\Rightarrow$ *larger classes will be stronger connected*
2. **(density)** $\frac{\#\{(a,b)\in E\,;\; a\in A\, b\in B\}}{|A|\cdot|B|}$
   $\Rightarrow$ *tends to zero when class sizes increase (sparsity)*
3. **(avg. number of $B$-neighbors)** $\frac{\#\{(a,b)\in E\,;\; a\in A\, b\in B\}}{|A|}$
   $\Rightarrow$ *asymmetric class-level network*
4. **(symmetric normalization)** $\frac{\#\{(a,b)\in E\,;\; a\in A\, b\in B\}}{\sqrt{|A|\cdot|B|}}$

   $\Rightarrow$ this is what we take

# Normalization of Edge-Weights

## how strongly are two classes connected?

given: network $G = (V, E)$ and two actor classes $A$ and $B$

1. **(un-normalized count)** $\#\{(a, b) \in E\,;\; a \in A,\; b \in B\}$
   $\Rightarrow$ *larger classes will be stronger connected*
2. **(density)** $\frac{\#\{(a,b) \in E\,;\; a \in A\; b \in B\}}{|A| \cdot |B|}$
   $\Rightarrow$ *tends to zero when class sizes increase (sparsity)*
3. **(avg. number of $B$-neighbors)** $\frac{\#\{(a,b) \in E\,;\; a \in A\; b \in B\}}{|A|}$
   $\Rightarrow$ *asymmetric class-level network*
4. **(symmetric normalization)** $\frac{\#\{(a,b) \in E\,;\; a \in A\; b \in B\}}{\sqrt{|A| \cdot |B|}}$
   $\Rightarrow$ this is what we take

# Normalization of Edge-Weights

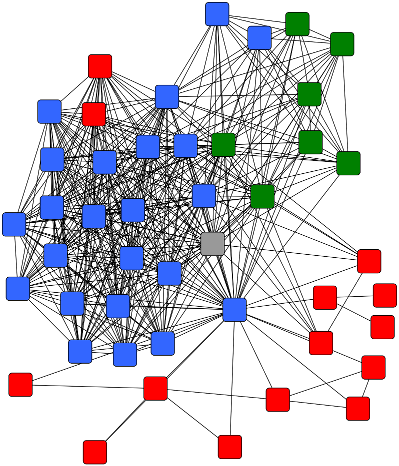
## how strongly are two classes connected?

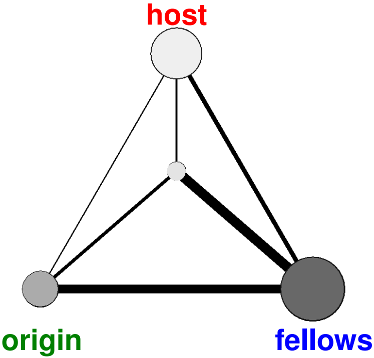
given: network $G = (V, E)$ and two actor classes $A$ and $B$

1. **(un-normalized count)** $\#\{(a, b) \in E\,;\ a \in A,\ b \in B\}$
   $\Rightarrow$ *larger classes will be stronger connected*
2. **(density)** $\frac{\#\{(a,b) \in E\,;\ a \in A\, b \in B\}}{|A| \cdot |B|}$
   $\Rightarrow$ *tends to zero when class sizes increase (sparsity)*
3. **(avg. number of $B$-neighbors)** $\frac{\#\{(a,b) \in E\,;\ a \in A\, b \in B\}}{|A|}$
   $\Rightarrow$ *asymmetric class-level network*
4. **(symmetric normalization)** $\frac{\#\{(a,b) \in E\,;\ a \in A\, b \in B\}}{\sqrt{|A| \cdot |B|}}$
   $\Rightarrow$ this is what we take

# Class-Level Network

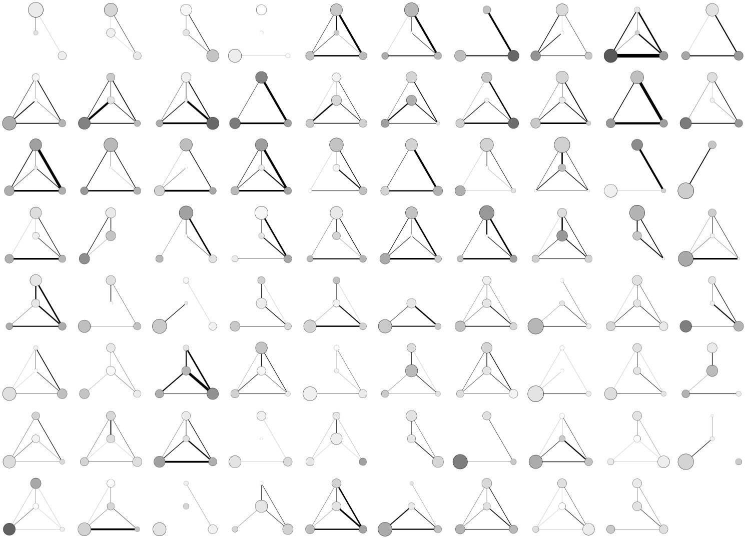
## argentinean woman living in spain






node size = class size
darkness = intra-class ties

# 79 Argentinean Migrants to Spain



# Overview

## Comparing Networks by their Group Structure

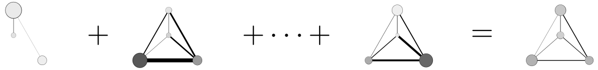
1. example application:
   towards a network measure for acculturation
2. defining class-level networks
3. **average class-level networks**
4. conclusion and future work

# Averaging Class-Level Networks

do Argentineans have different networks than Moroccans?

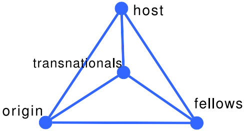
**arithmetic mean** $\overline{X} = (X_1 + X_2 + \cdots + X_N)/N$

how to add networks?



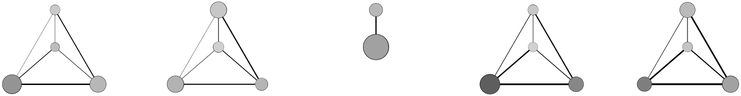
- add class-sizes
- add un-normalized edge counts
- normalize at the end

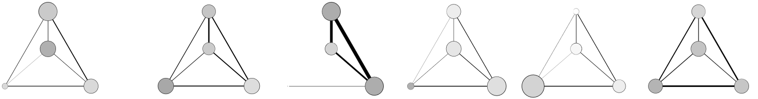
# Average Networks of . . .




## . . . migrants to Spain

Dominican Republic Argentina Equatorial Guinea Senegal/Gambia Morocco

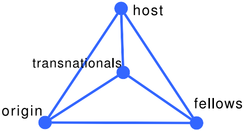




Dominican Republic Colombia Cuba Haiti Mexico Puerto Rico

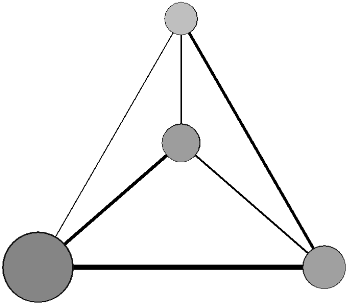
## . . . migrants to the USA

# Dependence on Time of Residence

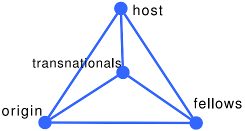



averaging over all migrants that ...

...moved to host country **less than 1** year ago

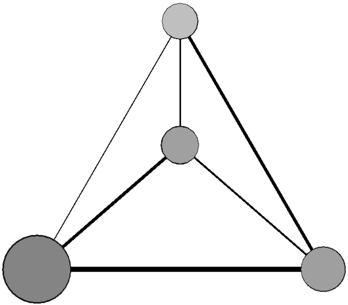


# Dependence on Time of Residence

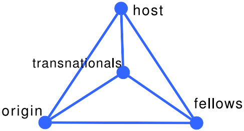



averaging over all migrants that ...

... moved to host country **1** year ago

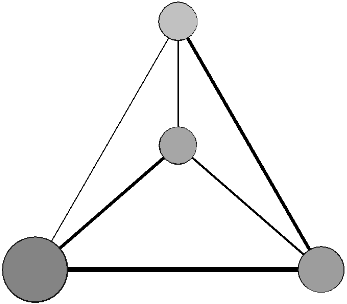


# Dependence on Time of Residence

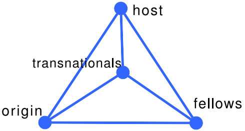



averaging over all migrants that . . .

. . . moved to host country **2** years ago

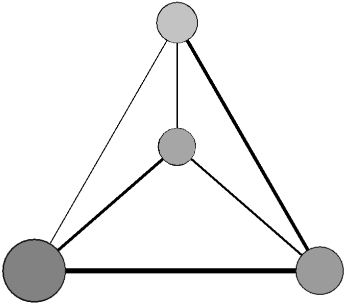


# Dependence on Time of Residence

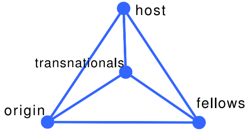



averaging over all migrants that . . .

. . . moved to host country **3** years ago

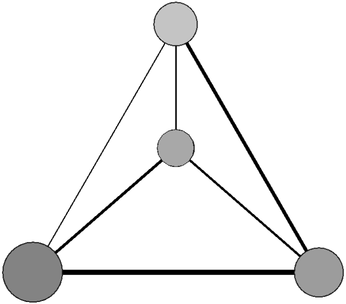


# Dependence on Time of Residence

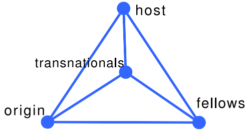



averaging over all migrants that . . .

. . . moved to host country **4** years ago

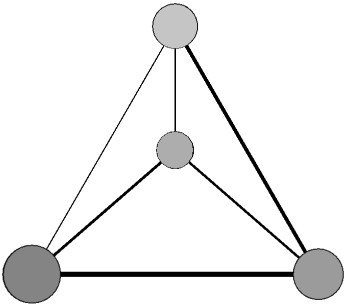


# Dependence on Time of Residence

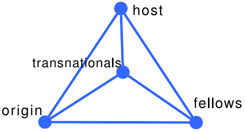



averaging over all migrants that . . .

. . . moved to host country **5** years ago

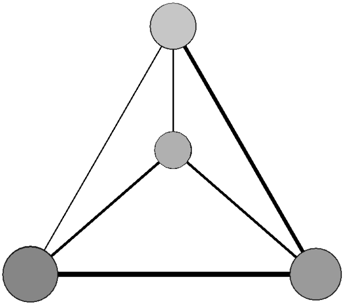


# Dependence on Time of Residence

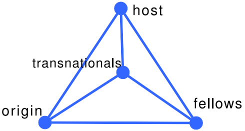



averaging over all migrants that . . .

. . . moved to host country **6** years ago

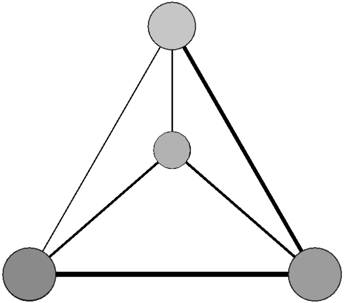


# Dependence on Time of Residence

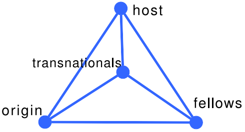



averaging over all migrants that . . .

. . . moved to host country **7** years ago

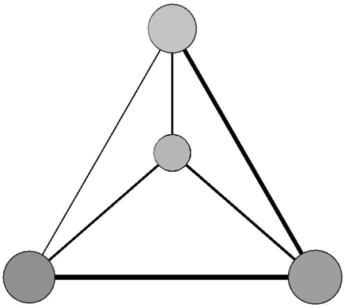


# Dependence on Time of Residence

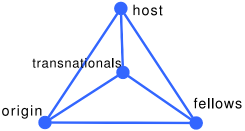



averaging over all migrants that . . .

. . . moved to host country **8** years ago

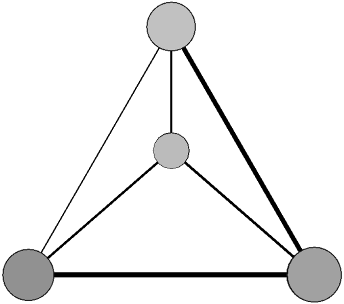


# Dependence on Time of Residence

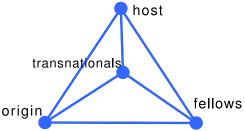



averaging over all migrants that . . .

. . . moved to host country **9** years ago

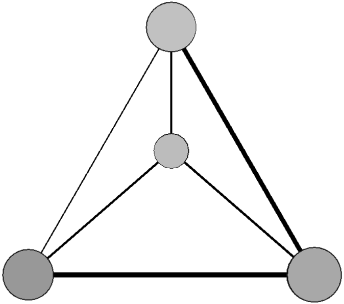


# Dependence on Time of Residence

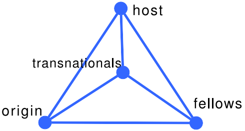



averaging over all migrants that . . .

. . . moved to host country **10** years ago

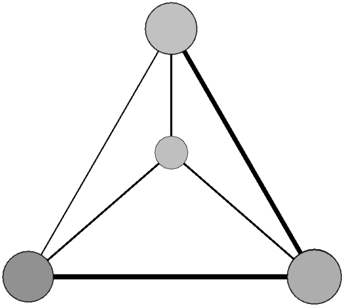


# Dependence on Time of Residence

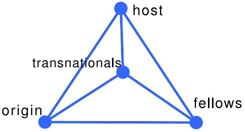



averaging over all migrants that . . .

. . . moved to host country **11** years ago

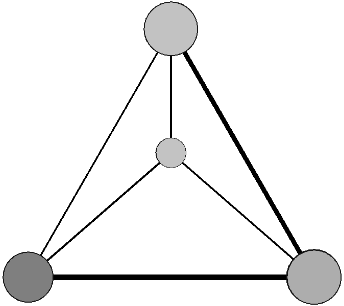


# Dependence on Time of Residence

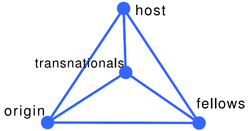



averaging over all migrants that . . .

. . . moved to host country **12** years ago

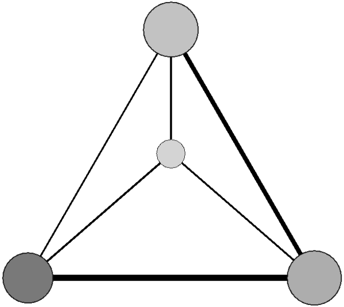


# Dependence on Time of Residence

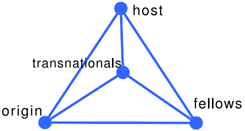



averaging over all migrants that . . .

. . . moved to host country **13** years ago

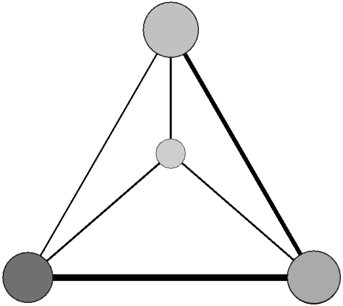


# Dependence on Time of Residence

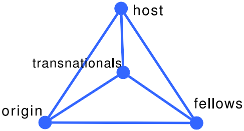



averaging over all migrants that ...

...moved to host country **14** years ago

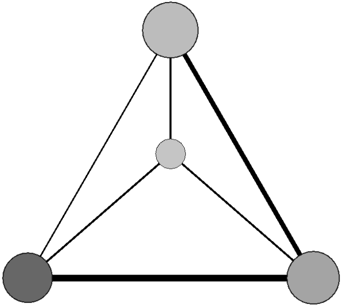


# Dependence on Time of Residence

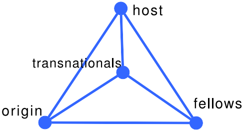



averaging over all migrants that ...

... moved to host country **15** years ago

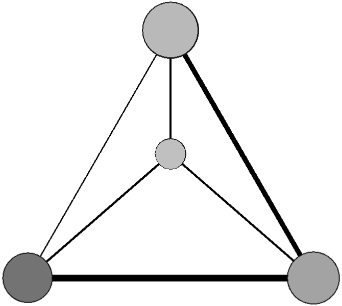


# Dependence on Time of Residence

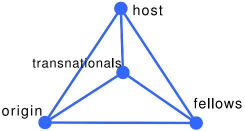



averaging over all migrants that . . .

. . . moved to host country **16** years ago

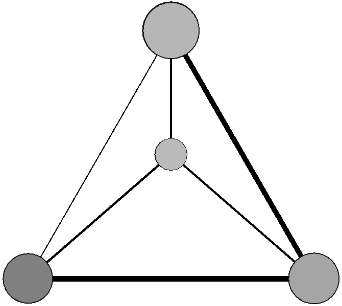


# Dependence on Time of Residence

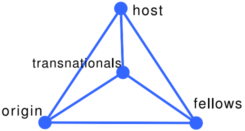



averaging over all migrants that ...

... moved to host country **17** years ago

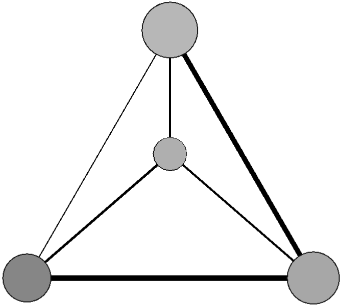


# Dependence on Time of Residence

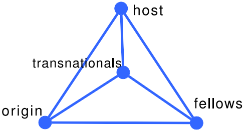



averaging over all migrants that . . .

. . . moved to host country **18** years ago

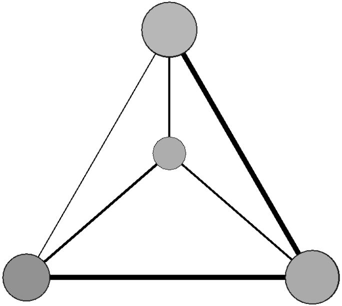


# Dependence on Time of Residence

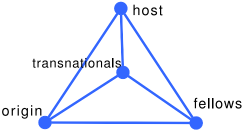



averaging over all migrants that . . .

. . . moved to host country **19** years ago

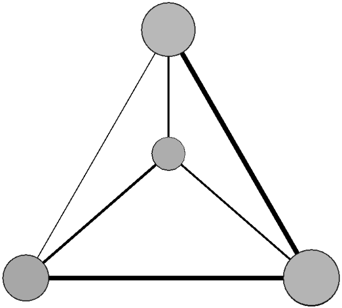


# Dependence on Time of Residence

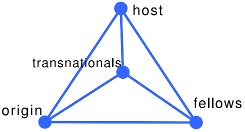



averaging over all migrants that . . .

. . . moved to host country **20** years ago

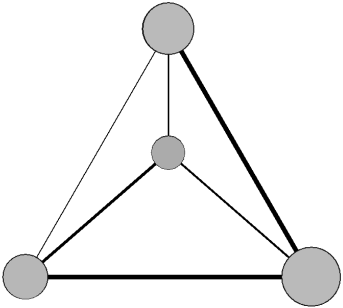


# Dependence on Time of Residence

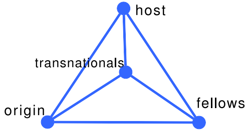



averaging over all migrants that . . .

. . . moved to host country **21** years ago

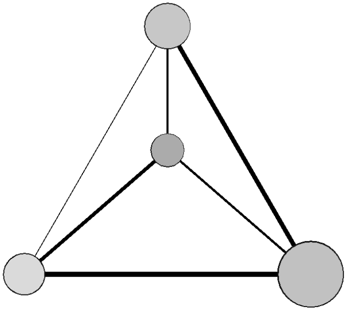


# Dependence on Time of Residence

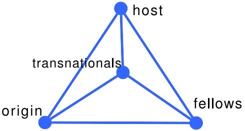



averaging over all migrants that . . .

. . . moved to host country **22** years ago

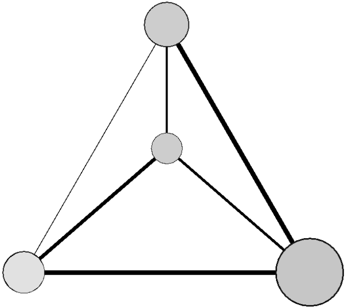


# Dependence on Time of Residence

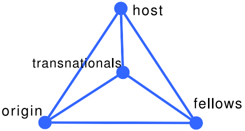



averaging over all migrants that ...

...moved to host country **23** years ago

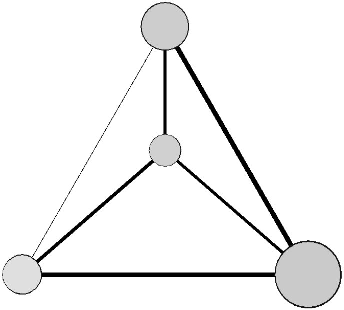


# Dependence on Time of Residence

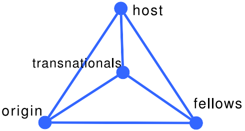



averaging over all migrants that ...

... moved to host country **24** years ago

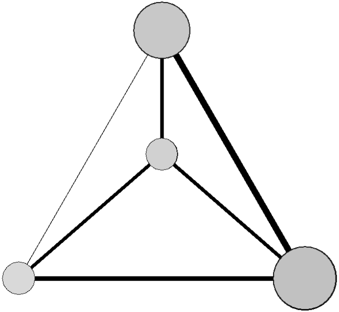


# Dependence on Time of Residence

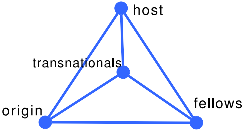



averaging over all migrants that ...

... moved to host country **25** years ago

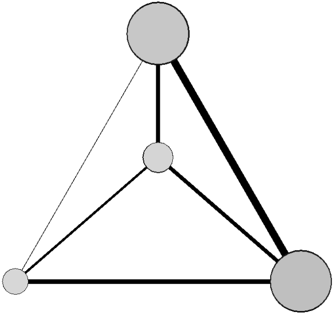


# Dependence on Time of Residence

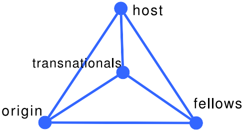



averaging over all migrants that ...

...moved to host country **26** years ago

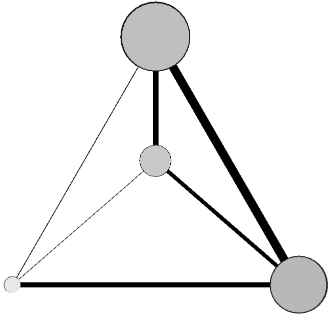


# Dependence on Time of Residence

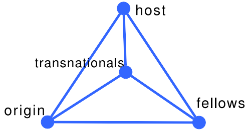



averaging over all migrants that . . .

. . . moved to host country **27** years ago

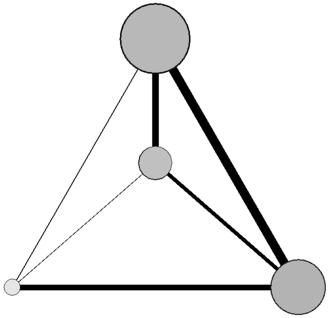


# Dependence on Time of Residence

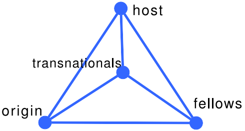



averaging over all migrants that . . .

. . . moved to host country **28** years ago

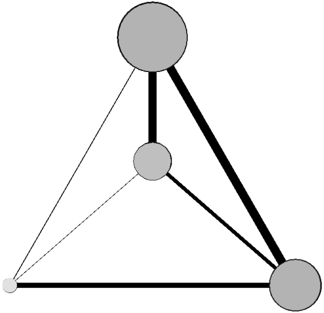


# Dependence on Time of Residence

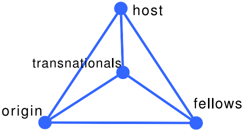



averaging over all migrants that . . .

. . . moved to host country **29** years ago

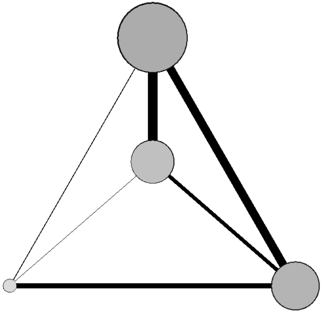


# Dependence on Time of Residence

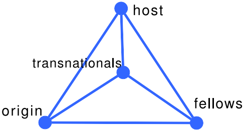



averaging over all migrants that ...

... moved to host country **30** years ago

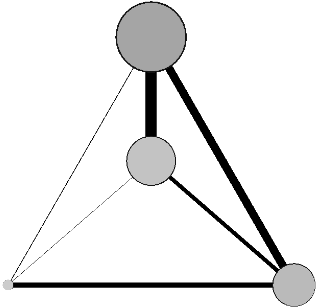


# Dependence on Time of Residence

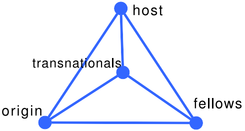



averaging over all migrants that . . .

. . . moved to host country **31** years ago

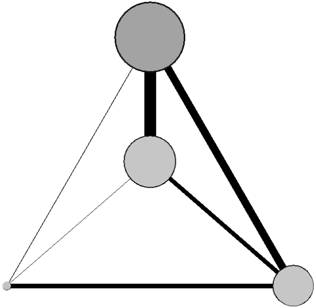


# Dependence on Time of Residence

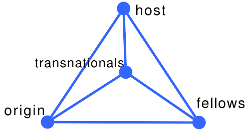



averaging over all migrants that . . .

. . . moved to host country **32** years ago

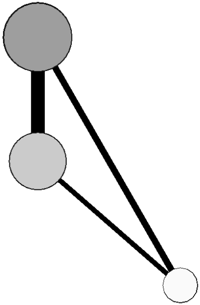


# Dependence on Time of Residence

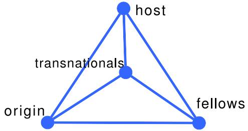



averaging over all migrants that . . .

. . . moved to host country **33** years ago

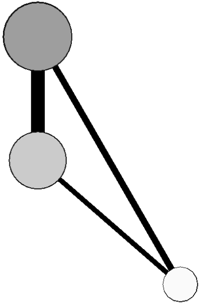


# Dependence on Time of Residence

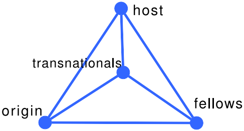



averaging over all migrants that . . .

. . . moved to host country **34** years ago

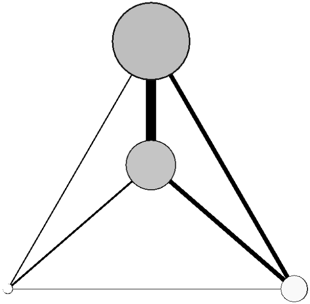


# Dependence on Time of Residence

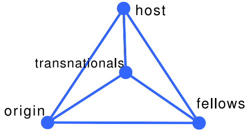



averaging over all migrants that . . .

. . . moved to host country **35** years ago

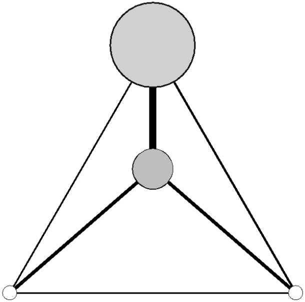


# Dependence on Time of Residence

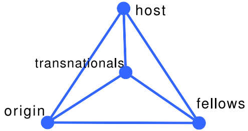



averaging over all migrants that . . .

. . . moved to host country **36** years ago

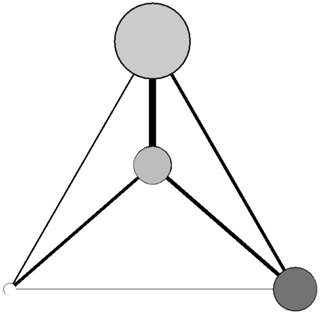


# Dependence on Time of Residence

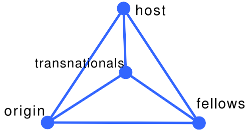



averaging over all migrants that . . .

. . . moved to host country **37** years ago

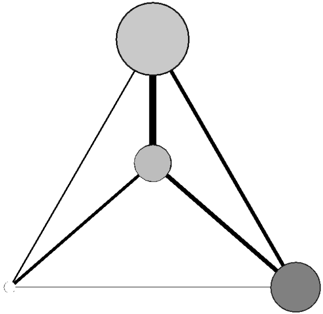


# Dependence on Time of Residence

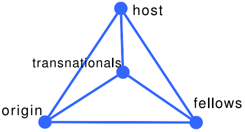



averaging over all migrants that . . .

. . . moved to host country **38** years ago

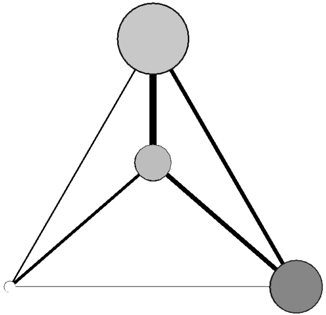


# Dependence on Time of Residence

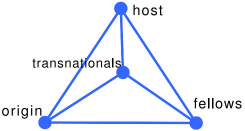



averaging over all migrants that . . .

. . . moved to host country **39** years ago

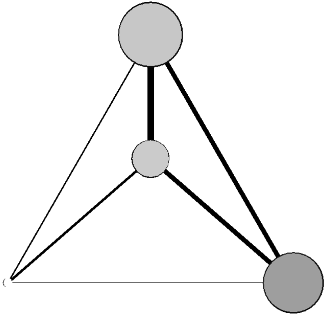


# Dependence on Time of Residence

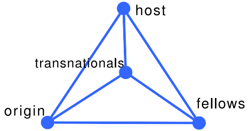



averaging over all migrants that ...

... moved to host country **40** years ago

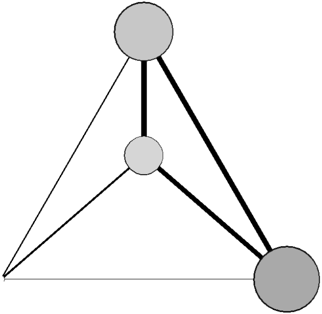


# Dependence on Time of Residence

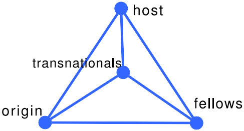



averaging over all migrants that . . .

. . . moved to host country **41** years ago

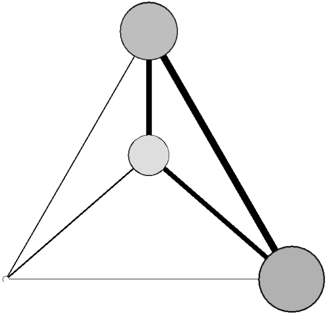


# Dependence on Time of Residence

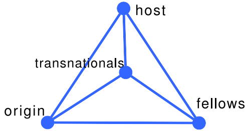



averaging over all migrants that . . .

. . . moved to host country **42** years ago

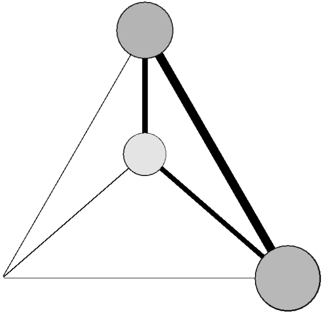


# Dependence on Time of Residence

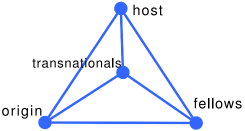



averaging over all migrants that . . .

. . . moved to host country **43** years ago

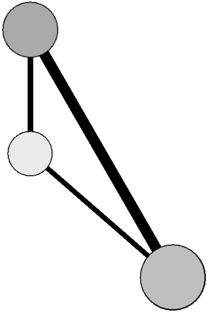


# Dependence on Time of Residence

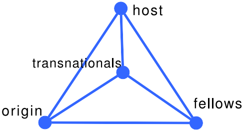



averaging over all migrants that ...

...moved to host country **44** years ago

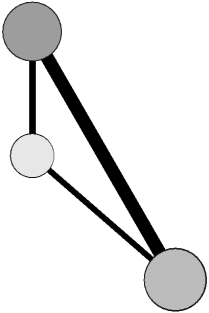


# Dependence on Time of Residence

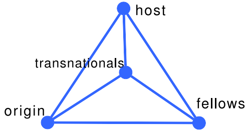



averaging over all migrants that ...

... moved to host country **45** years ago

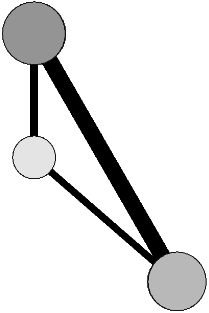


# Dependence on Time of Residence




averaging over all migrants that . . .

. . . moved to host country **46** years ago



# Conclusion and Future Work

- comparing networks by their group structure (actor classes defined by attributes)
  - very simple and efficient
  - allows averaging over sets of networks
  - needs meaningful definition of actor classes
- **future work**
  - acculturation mode $\Leftrightarrow$ network structure
  - consider classes defined by attributes and ties (e. g., relative regular equivalence [Boyd/Everett'99])
  - determine meaningful combinations of attributes