

# Simulation and mining of social networks in a pandemic context

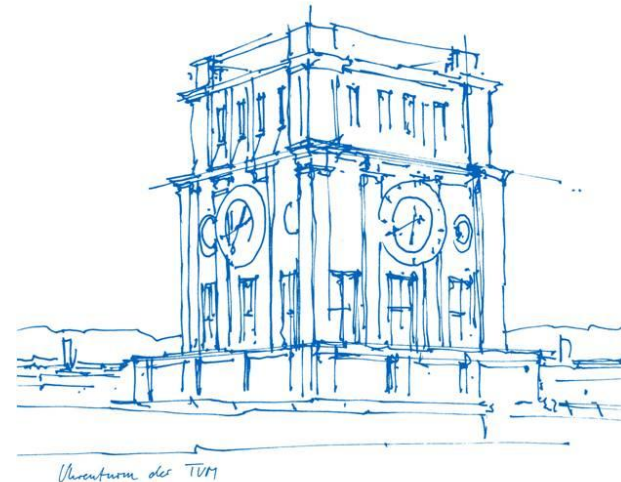
Aldi Topalli

Technische Universität München

Faculty of Informatics

Chair of Social Computing

Garching, 10.05.2021



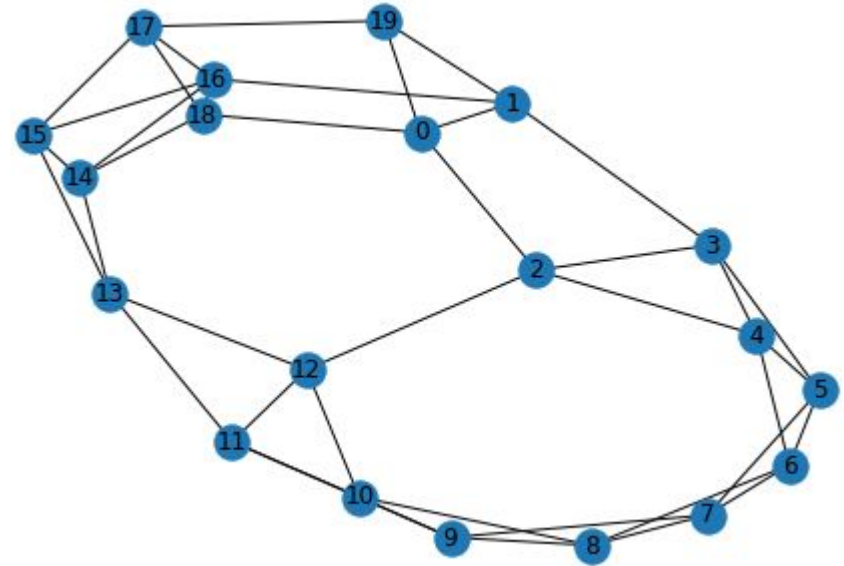
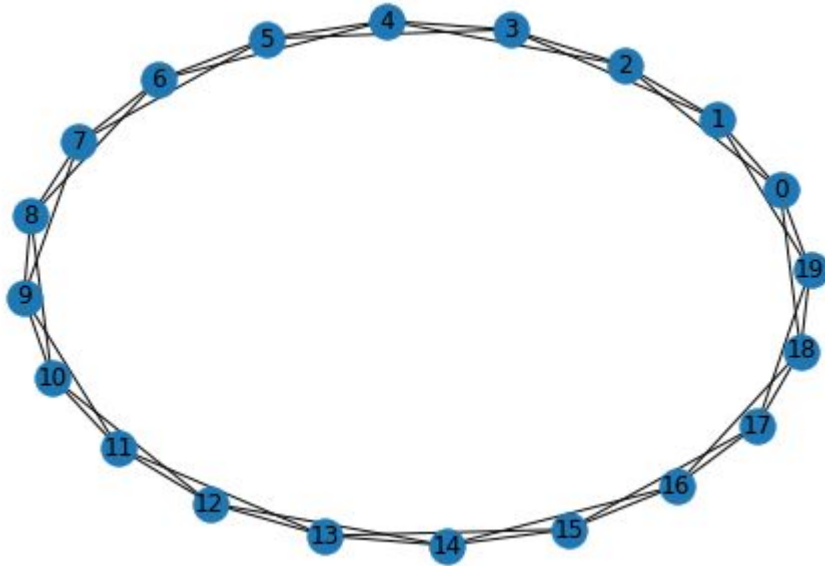
# Outline

- Generation of long-term social network
- Action-based risk prediction

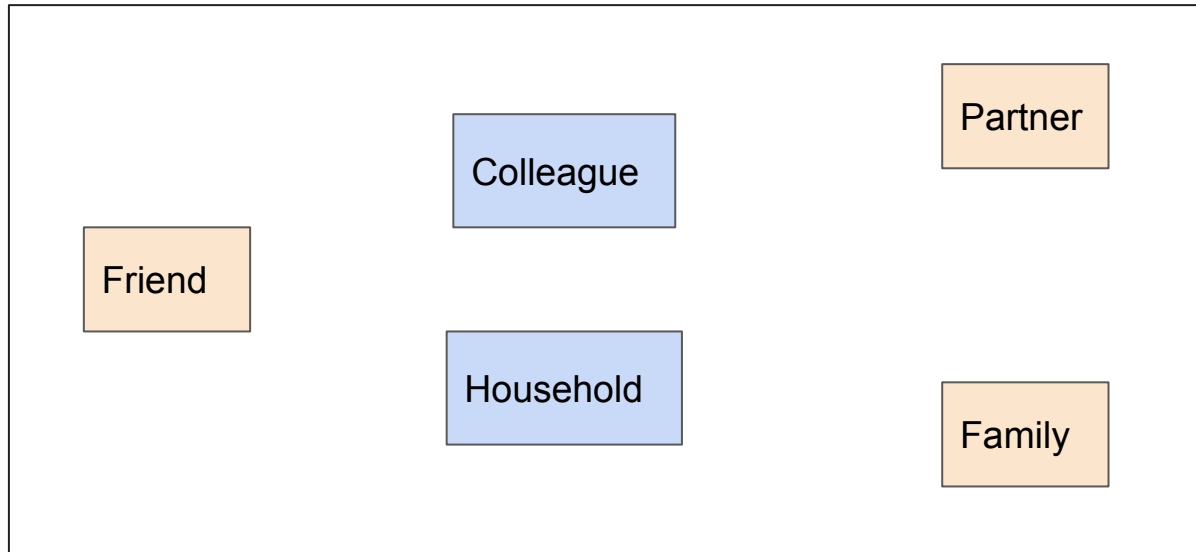
# Objective

- Generate a real-world graph
- Model long-term relationships
- Store work and home facility

# Watts-Strogatz model



# Edge types

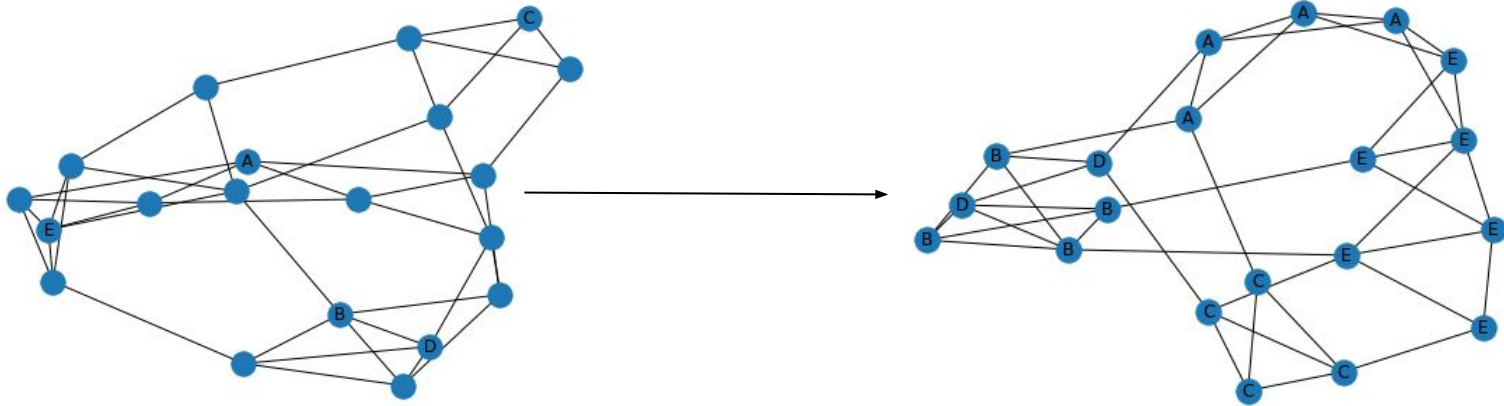


Based on node  
attributes

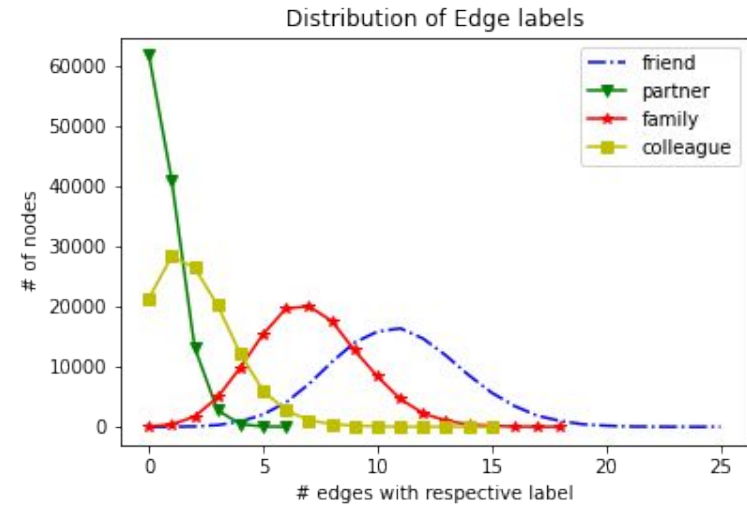
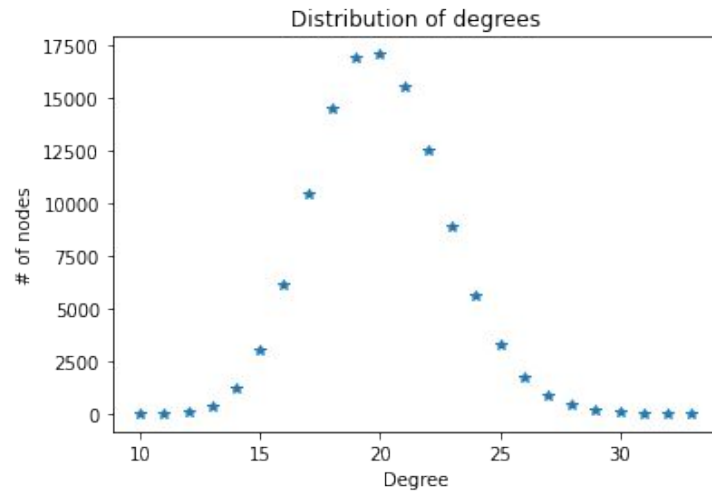
Random assignment

# Node classification [1]

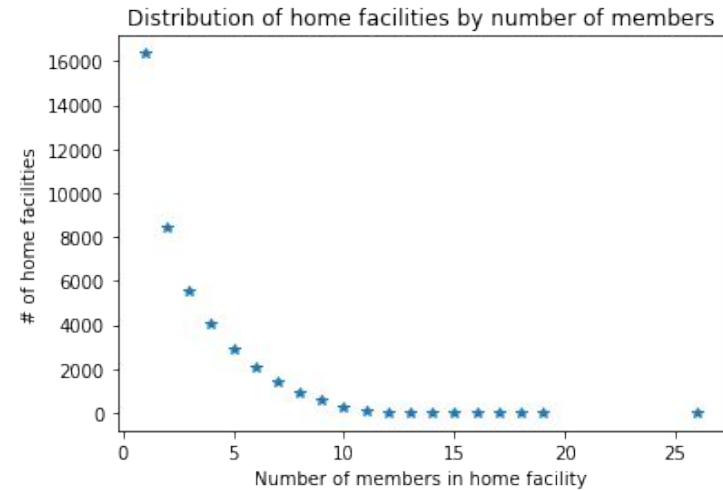
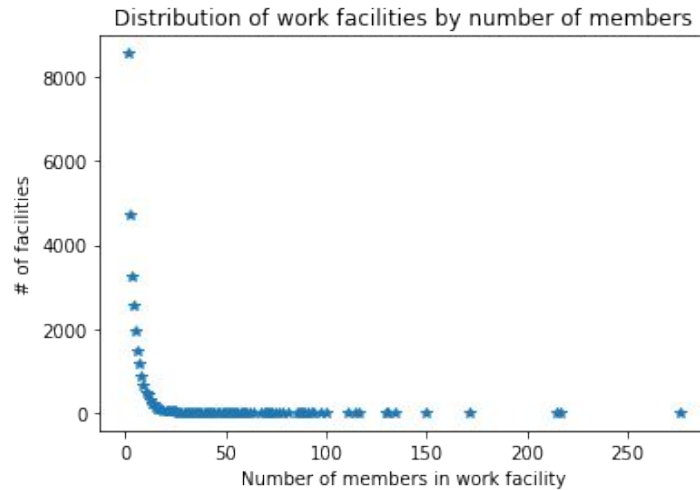
- Assign given facility to a random node
- Propagate attribute: Close nodes should have similar attributes



# Results



# Results





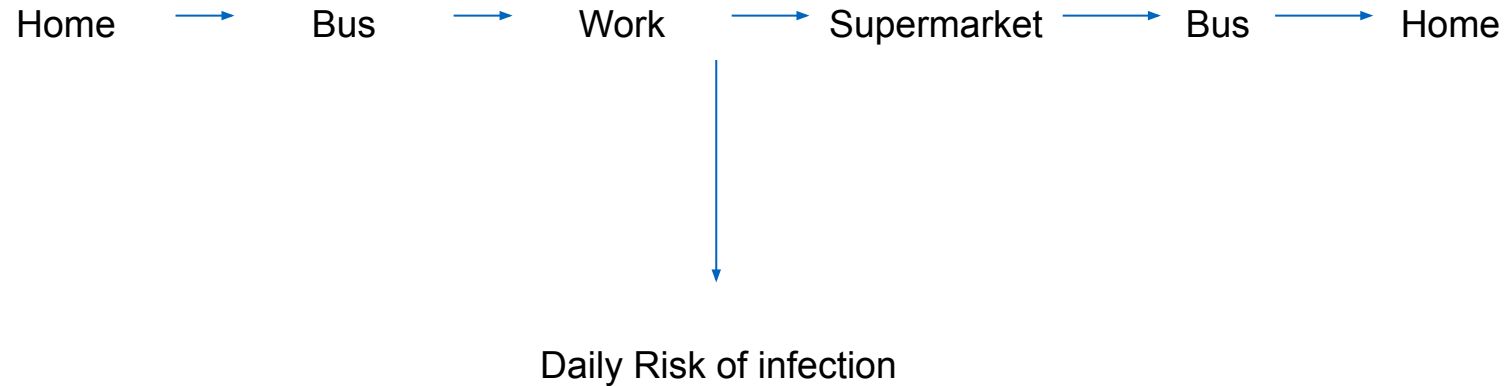
# Outline

- Generation of long-term social network
- Action-based risk prediction

# Daily “sentence”

Home → Bus → Work → Supermarket → Bus → Home

# Daily “sentence”



# Experiment

- Train concept embeddings
- Train a bidirectional LSTM on the concept embeddings
- Cluster the embedding space
- Discover risk clusters

# Results

- Results from embeddings → not good
- Second approach necessary
- Idea
  - Distribute the risk based on each person doing the action
  - Normalize for each person to lower influence of highly active people

# Bag of actions

Create a timeline of actions per person (14 days)

- Distribute risk based on:
  - Group size
  - Duration
- Normalize (Softmax) actions per person
- Sum all same actions to get the risk for the specific action
- Rank the actions in the timeline

# More detail

- $Risk_{jt} = \text{group\_size} + \text{duration}$  (j-th action, t-th person)
- $Risk'_{jt} = \text{Softmax}_j(Risk_t)$
- $Risk_t = \sum_j Risk'_{j,t}$
- Riskiest action for person t = maximum  $Risk'_j$  of the timeline

# More detail

- $Risk_{jt} = \text{group\_size} + \text{duration}$  (j-th action, t-th person)
- $Risk'_{jt} = \text{Softmax}(Risk_t)$
- $Risk_j = \sum_t Risk'_{j,t}$
- Riskiest action for person t = maximum  $Risk'_j$  of the timeline

Result ~ 40 % accuracy



# Discussion and interpretation

- ML in the context of data
- Fine grained information (ML - related, Discriminative model)
- Patterns should be visible at an earlier stage
- Enable god-like mode

End