Social networks and influences on activity-travel behavior

Recent studies from the Eindhoven group

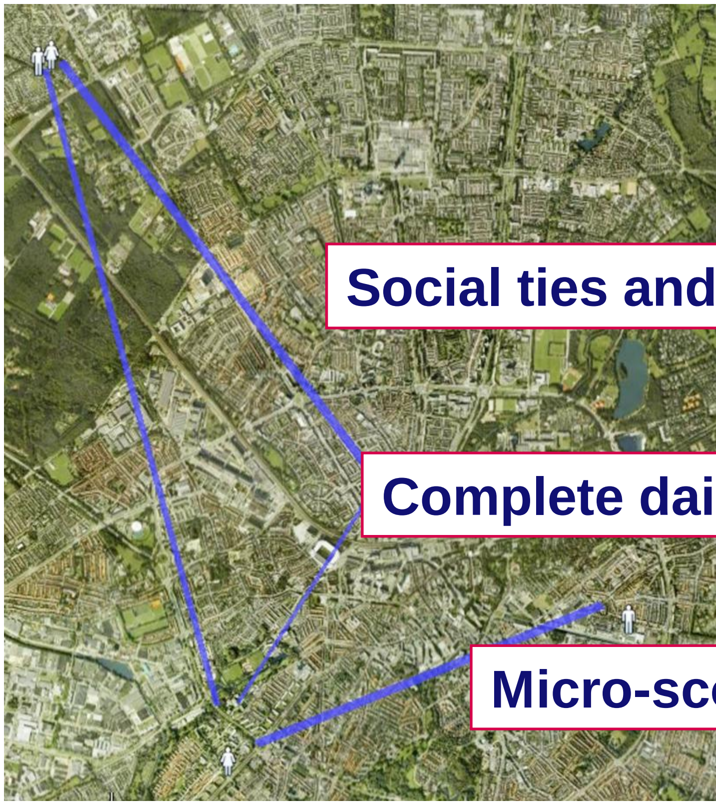
Renni Anggraini, Theo Arentze, Pauline van den Berg, Nicole Ronald, Harry Timmermans

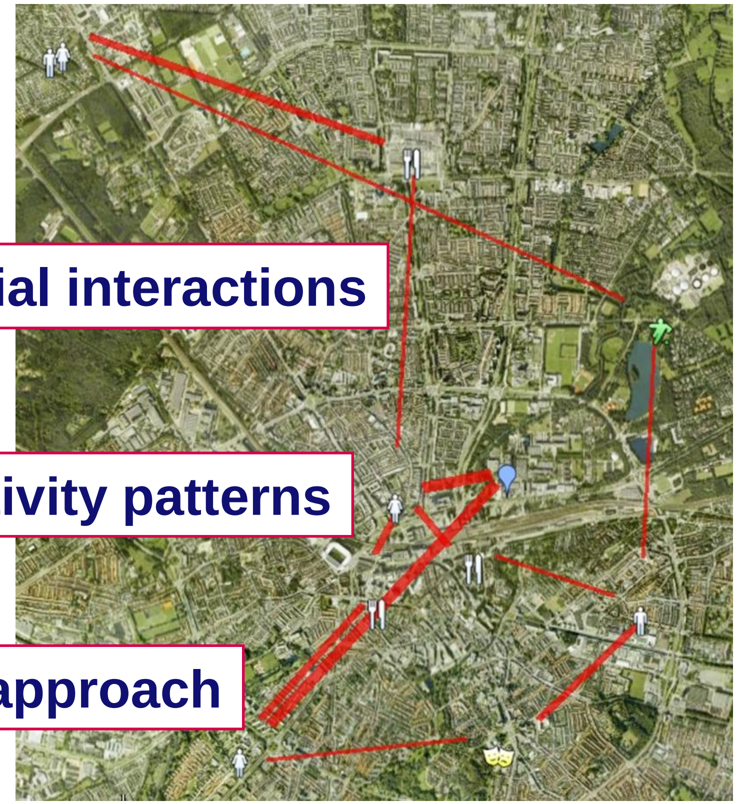
TU/e Technische Universiteit Eindhoven University of Technology

Purpose of the talk

- Study of social networks and social interactions has a long history in core disciplines
- It is a relatively new but rapidly emerging field in transportation research
- Purpose of my talk
  - Discuss some basic theories and concepts and highlight relevance for travel behavior
  - Show approaches and results from our research program in Eindhoven

What we are studying and modeling





Social ties and spatial interactions

Complete daily activity patterns

Micro-scopic approach

Three topic areas related to travel behavior

1. **Activity and trip generation**
2. **Social selection and network dynamics**
3. **Social influence**

Activity and trip generation

- **Travel for social purposes accounts for a large proportion of travel demand**

- **Activity scheduling decisions**
  - **Joint activity participation – coupling constraints**
  - **Support and task allocation**
  - **Sharing of and competition for resources**

Social selection

- **Homophily – more similar is more attractive**
- **Common friends – increases probability of a tie**
- **Centrality of position in social network**
- **Geographic distance**

Social network dynamics

- **Social networks are not static**
- **Influenced by life-cycle events**
- **Influences long-term mobility choices of households**

Social influence

- **Individuals make decisions in a social environment**
  - **Mimicking – copying behavior**
  - **Information exchange – learn about choice alternatives**
  - **Adapting aspirations and norms**

➢ **Adoption of new technologies / modes of behavior**

➢ **Non-linear effects (hypes, herd behavior, etc.)**

Studies in Eindhoven on the topic

1. **Social interactions, ICT and travel behavior – data collection and analysis (Pauline van den Berg)**

2. **Simulating social networks in geographic space (Theo Arentze)**

3. **An agent-based framework for modeling social activities and travel (Nicole Ronald)**

4. **Dynamics of social networks and long-term mobility choices (Fariya Sharmeen)**

Outline of the remainder of the talk

- **Brief introduction activity-based approach**
  - **Within household interactions**
- **Study 1**
  - **Data collection and analysis**
- **Study 2**
  - **Simulating social networks**
- **Study 3**
  - **Agent-based model**

Technische Universiteit Eindhoven University of Technology

Brief introduction activity-based approach

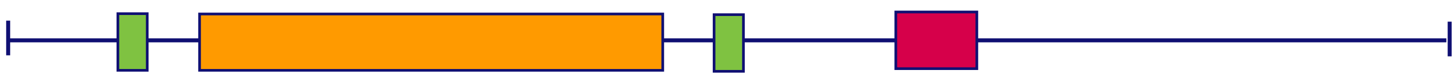
- **Within household interactions**

The activity-based approach

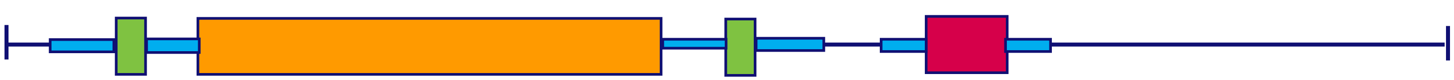
- **Activity-travel choices in a scheduling process**

1. **Activity generation**: Which activities for how long?

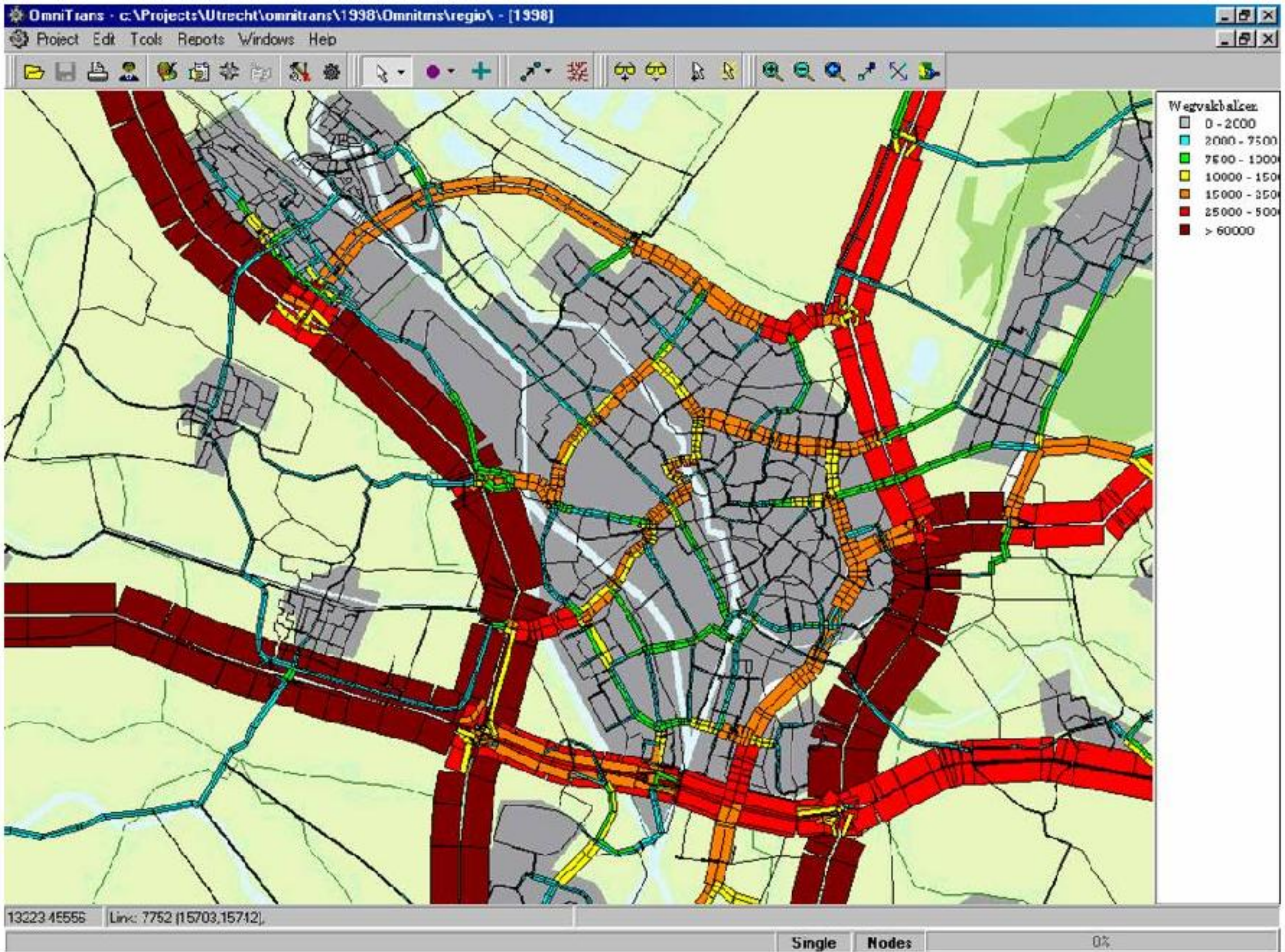
2. **Activity scheduling**: Sequence and timing of the activities?



3. **Travel choices**: Locations, trip-chaining and transport mode?

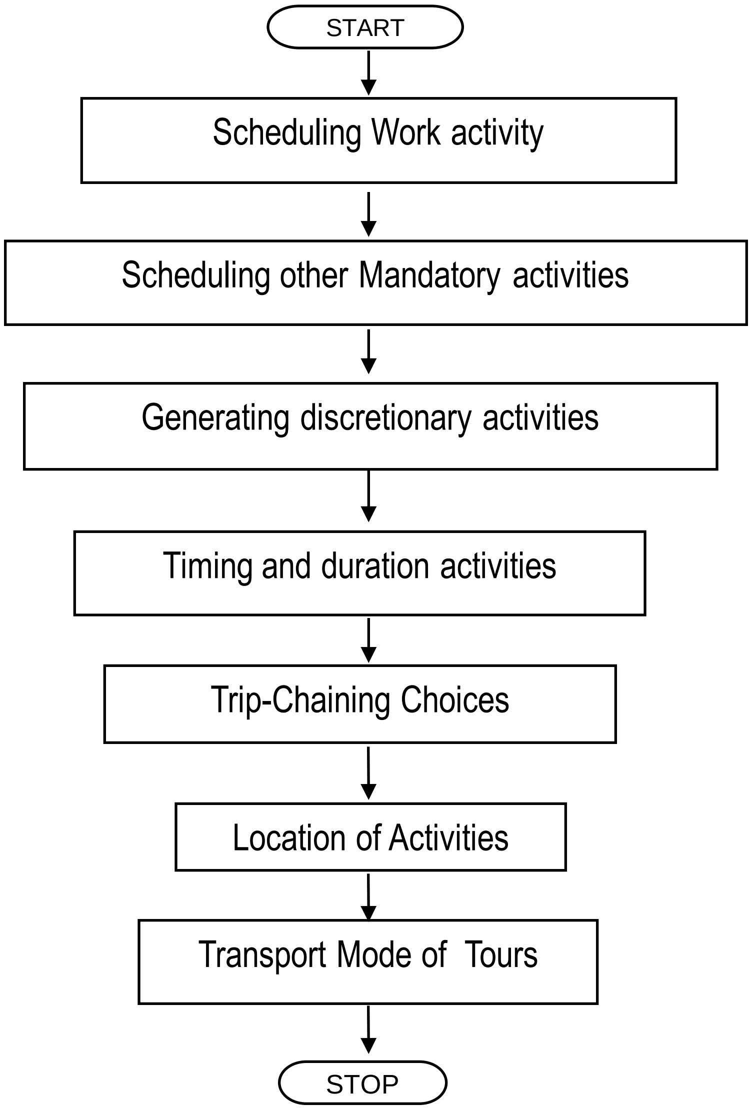


Activity patterns can be mapped to predict traffic flows throughout a day



OmniTrans ltd.

Modeling decisions - Albatross



Rules for making decisions

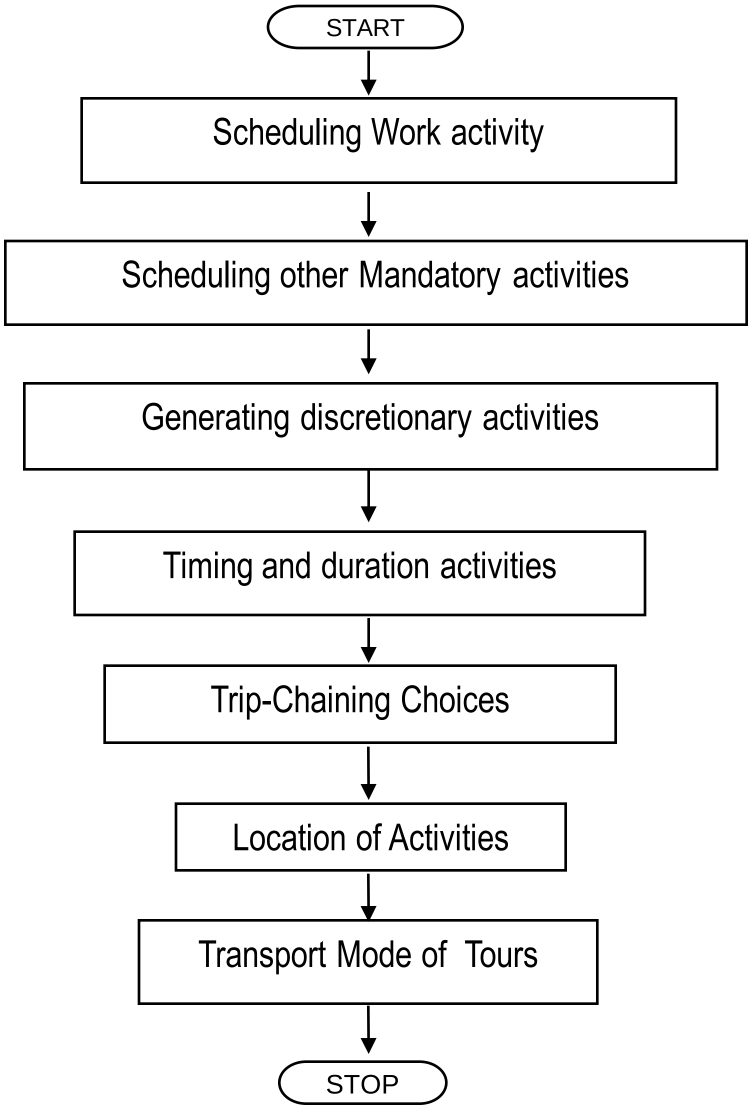
IF (condition = X)
THEN (perform action Y)

Rules are formatted as decision trees

Decision trees are derived from choice observations

Thus, decisions are based on heuristics

Modeling decisions - Albatross

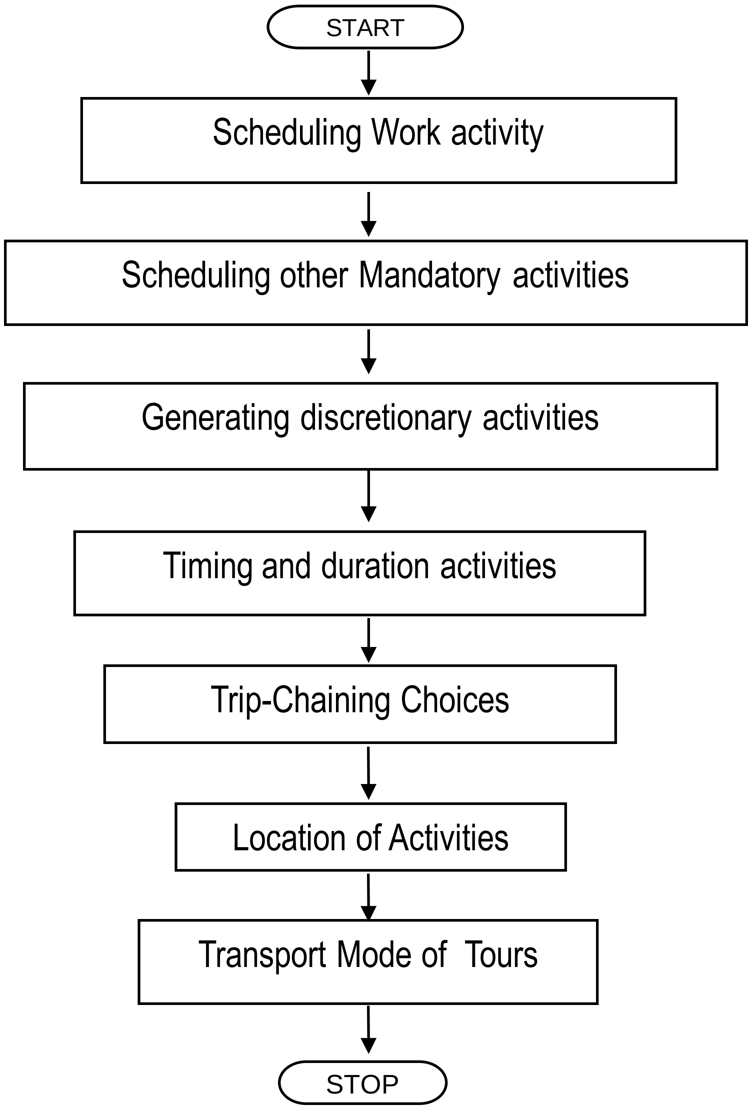


Data of the environment

Travel times by transport mode
- **Road network**
- **Bus and train services**
- **Parking spaces**

Available locations for activities
- **Land-use data**
- **Employment data**
- **Opening hours data**

Modeling decisions - Albatross

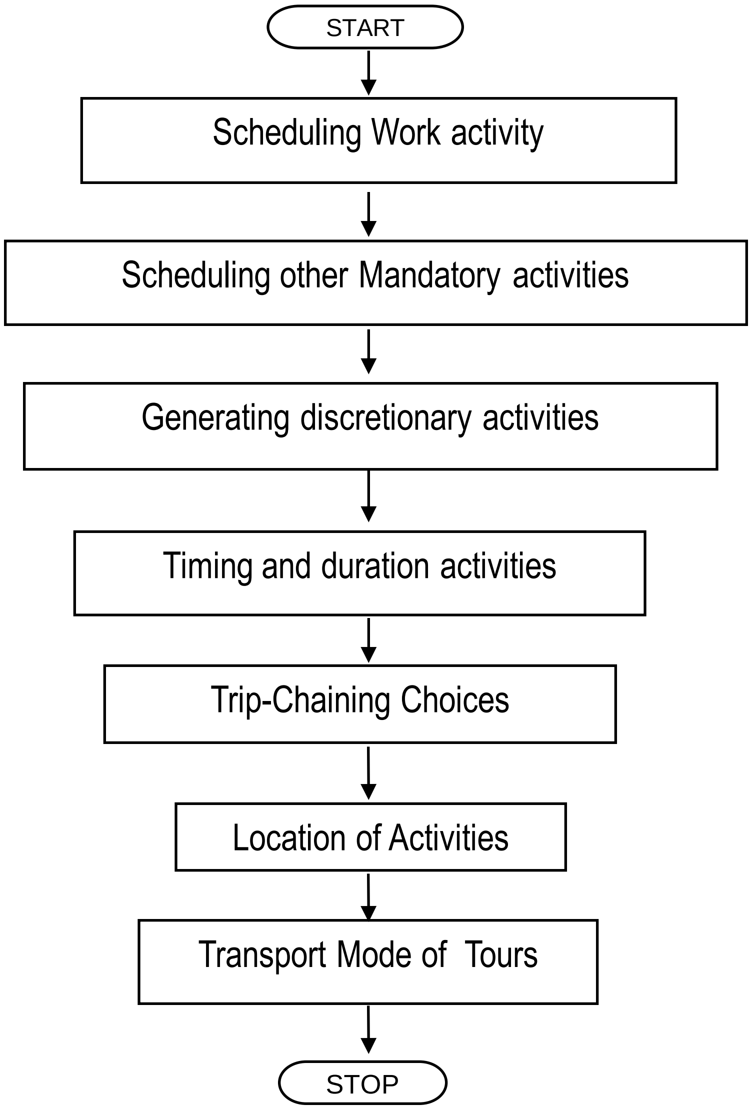


Constraints on decisions

Space-time constraints

Resource constraints

Modeling decisions – Albatross



Within household interaction

Car allocation
Who gets the car for which activity?

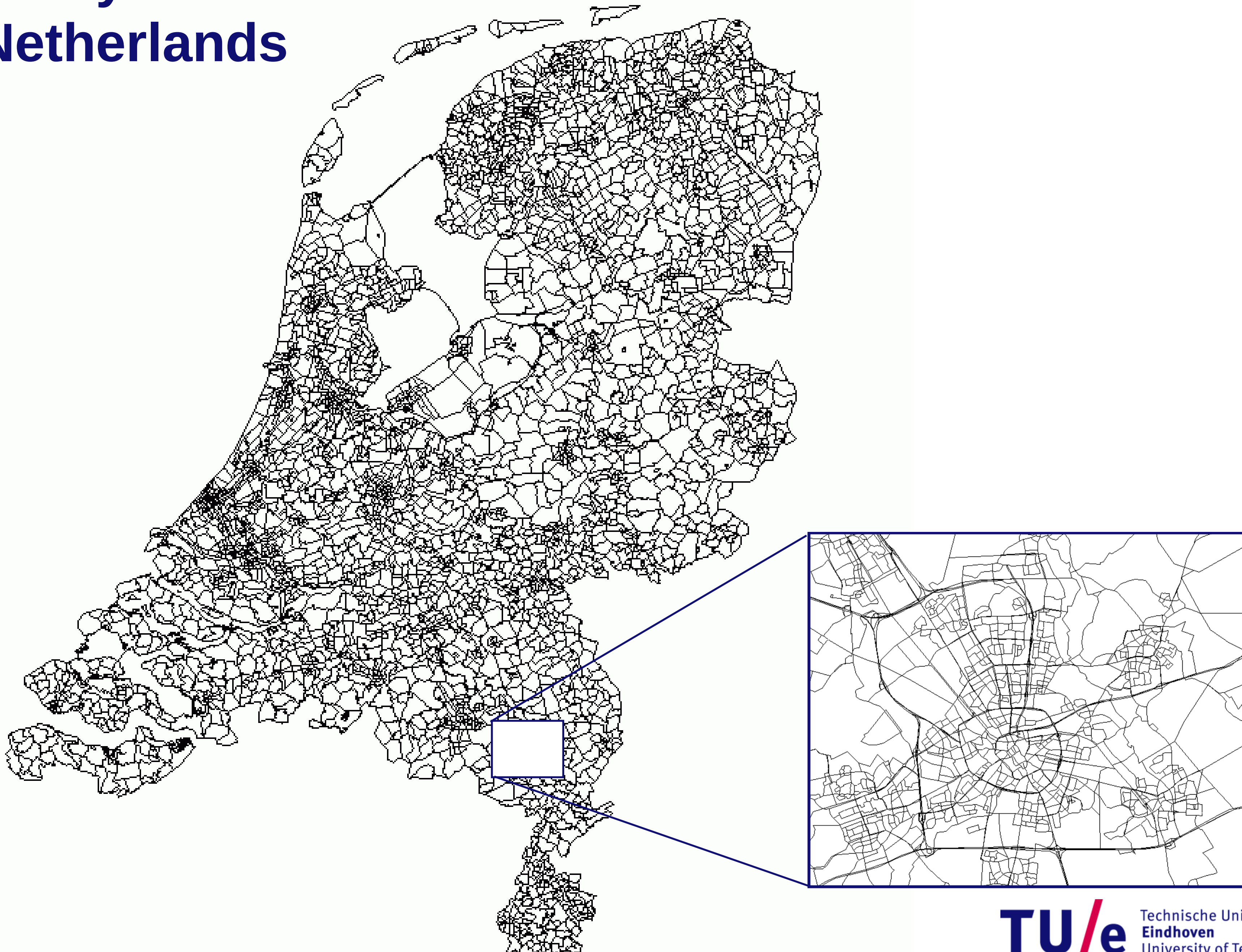
Task-allocation
Who conducts which household activity?

Joint activity participation

Additional coupling constraints

Technische Universiteit
Eindhoven
University of Technology

Study area is the Netherlands



Technische Universiteit
Eindhoven
University of Technology

Deriving rules from data

National travel diary data-set

- 29,221 households, 66,482 persons, 231,899 trip records
- One day, all members of the households
- Balanced days of the week, seasons of the year
- Balanced regions within Netherlands

Validation of the model

- **Choice facet level**
  - **Goodness-of-fit of decision trees on hold-out sets**

- **Distributions of choice facets in the sample**

- **Sensitivity of the model**

Conclusions

- **Extended Albatross model takes into account**
  - **Joint activities**
  - **Task allocation**
  - **Car allocation**
  - **Coupling constraints**

- **Model is estimated on a national household travel survey and operational for large-scale applications**

- **Comparison indicates that incorporating within household interactions makes a difference in predictions**

Study 1

Social networks, ICT use and social activity-travel choice

Pauline van den Berg, Theo Arentze, Harry Timmermans

Objectives

- **To obtain Insight into social networks, frequency of social interaction, their purpose, with whom, distance and communication mode**

- **To obtain insight in influence of ICT on social travel**
  - **Substitution?**
  - **Complementarity?**
  - **Facilitation?**
  - **Neutral?**

- **To obtain insight in effects of spatial factors**

Data collection

- **January - June 2008**
- **Eindhoven region**

- **Social interaction diary**
  - **747 respondents**
  - **8074 social interactions (4177 face-to-face)**

- **Social networks**
  - **116 respondents**

Data collection – diary

- **For each diary day (N=1494):**
  - the number of social interactions with one person per day

- **For each social interaction (N=8074, 5.4 a day):**
  - main purpose
  - who initiated (N = 5291 initiated by ego)
  - social category of the contacted person
  - distance between ego and alter
  - communication mode choice

TU/e Technische Universiteit Eindhoven University of Technology

Data collection – social networks

- **Name generator**

1. Think about the people you feel very close to. They are:

  - people with whom you discuss important matters,
  - or regularly keep in touch with,
  - or that are there for you if you need help

  They can be household members, relatives, colleagues or fellow students, neighbors, club members and (other) friends

2. **Think about the people you feel somewhat close to. They are**

  - people that are more than just casual acquaintances, but not very close

  They can be household members, relatives, colleagues or fellow students, neighbors, club members and (other) friends

Data collection – social networks

- **Name interpreter - information of each alter**
  - **Gender**
  - **Age**
  - **Distance between residences**
  - **Type of relationship**
  - **How long they know each other**
  - **Frequency of contact by mode**
    - **Face-to-face**
    - **Telephone**
    - **Sms**
    - **Msn / IM**

Technische Universiteit **Eindhoven** University of Technology

Explanatory variables

- **Personal and household characteristics**

  (Gender, age, education, income, work and/or school hours, partner, children)

- **Mobility characteristics**

  (Car ownership, urban density)

- **Social characteristics**

  (Social network size, number of clubs/associations)

- **Day of the week**

Data analysis - diaries

- **Linked set of models to predict travel generation**
  - **Poisson regression model – number of interactions**
  - **MNL – purpose of social activities**
  - **Regression model – distance traveled**
  - **MNL – transport mode choice**

- **Structural equation model to test causal network structures among variables**

Data analysis – diaries - example

- **Model 1: Poisson regression model – number of interactions**

	B
Male	-0.100
Partner	-0.099
Children under 18	0.240
Primary education	-0.212
Involved in 2 or more clubs	0.163
# very strong network ties	0.007
# reasonable strong ties	0.003
Sunday	-0.412

Data analysis – diaries - example

- **Model 2: MNL – purpose of social activities**

	N	%
Joint activity	835	20%
Visit/host	592	14%
Talk/chat	1185	28%
Short question/message	353	8%
Info/advice/discussion	535	13%
Other	656	16%
	4177	100%

Data analysis – diaries - example

Model 2: Purpose of social activities – estimation results

	Joint activity	Visit / host	Talk / chat	Question / mess.	Info / discus
Male		-0.392	-0.418		
<40	0.545	0.553			
>60			-0.470		
Partner	-0.538		-0.472	-0.631	
Children			-0.402		
Low education					-0.494
No work / school			-0.535	-0.763	
Full time work / school	0.564		0.464		0.632
No car	-0.590		-0.664	-0.936	-0.608
High urban density	0.761	0.577	0.405	0.540	0.603
No clubs				0.484	
2 or more clubs	0.593			0.625	
# very strong ties		0.022			
# reasonable strong		-0.010	-0.012		
Monday			-0.459		-0.678
Wednesday				-0.603	
Saturday		0.861			
Sunday		1.198			
# F2F interactions	-0.068	-0.085		0.079	

Some conclusions - diaries

- **Relatively large sample of households provided extensive information on**
  - **Social interactions (2 days diary)**
  - **Social networks**

- **First analysis results indicate**
  - **Many significant effects**
  - **In particular gender, age, the presence of children in the household, level of education and the day of the week**

- **Further analysis of the data is ongoing**

Study 2

Simulating social networks in geographic space

Theo Arentze, Pauline van den Berg, Harry Timmermans

Objective

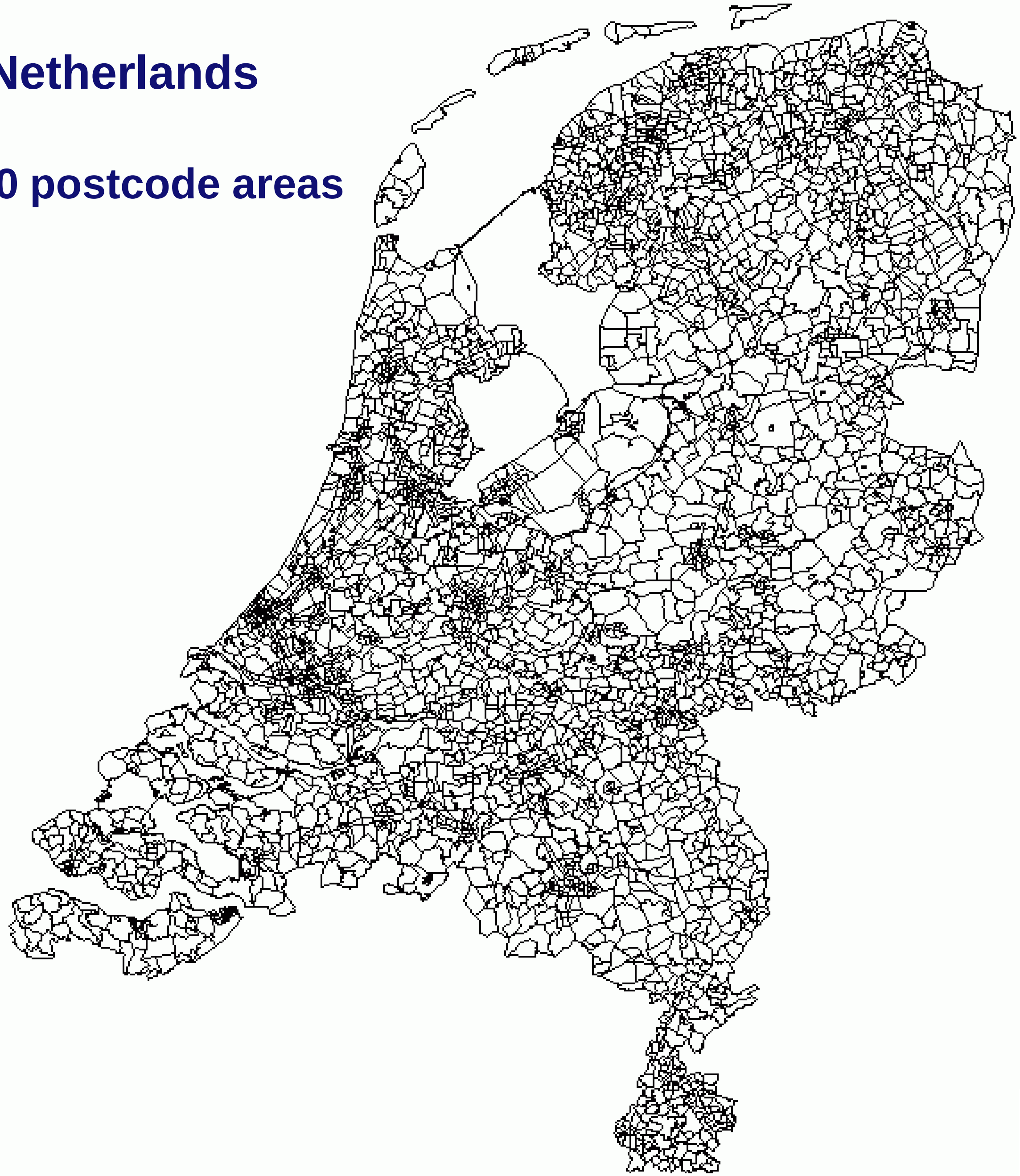
- **To micro-simulate the whole social network of an entire population**

- **Requirements**
  - **The synthesized network has the same statistical characteristics as the real network**
  - **Link-selection decisions are consistent with existing behavioral theories**
  - **Links are reciprocal (A – B ⇔ B – A)**

➢ **A synthetic social network can be used to micro-simulate social interactions in space**

The Netherlands

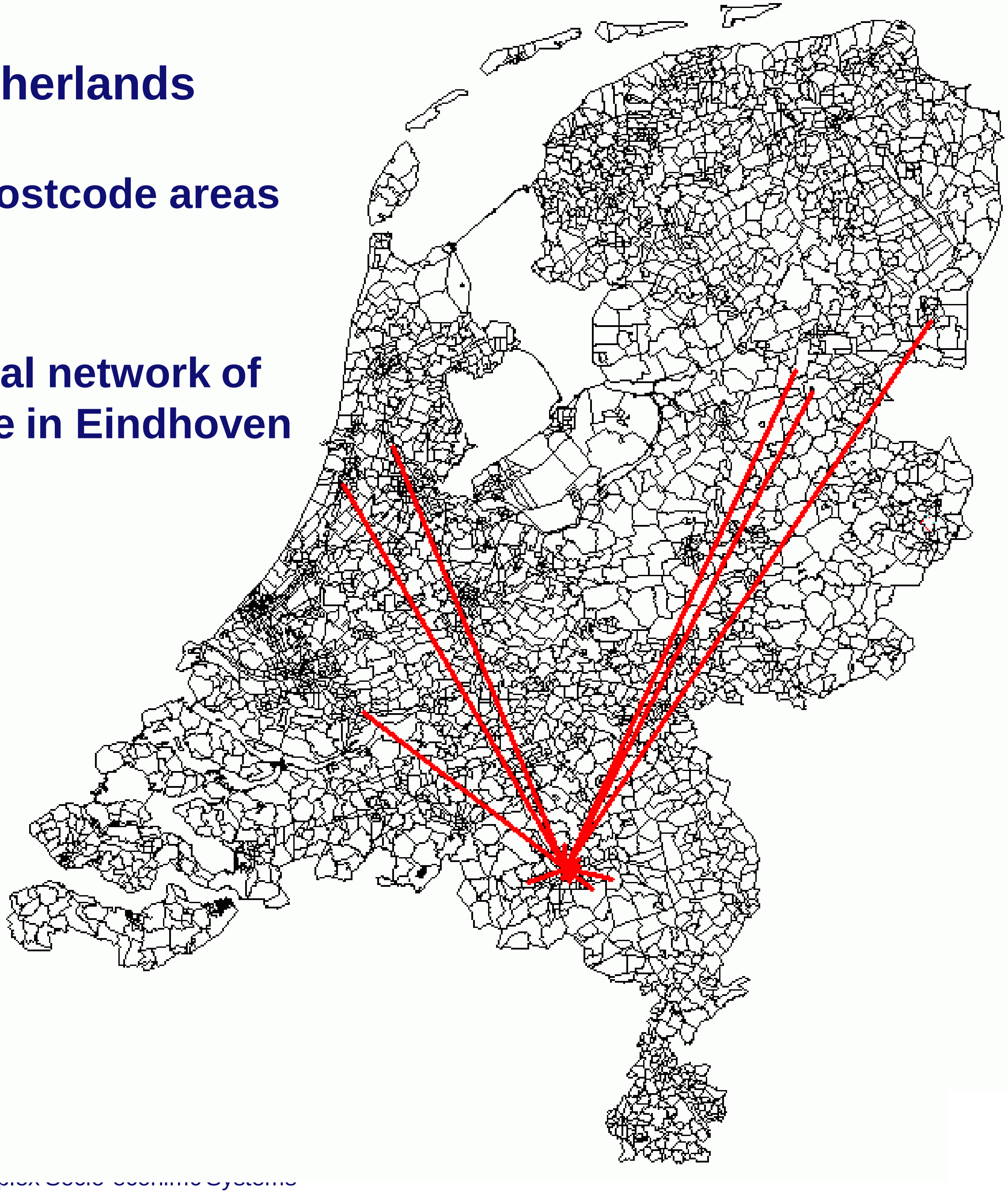
± 4000 postcode areas



The Netherlands

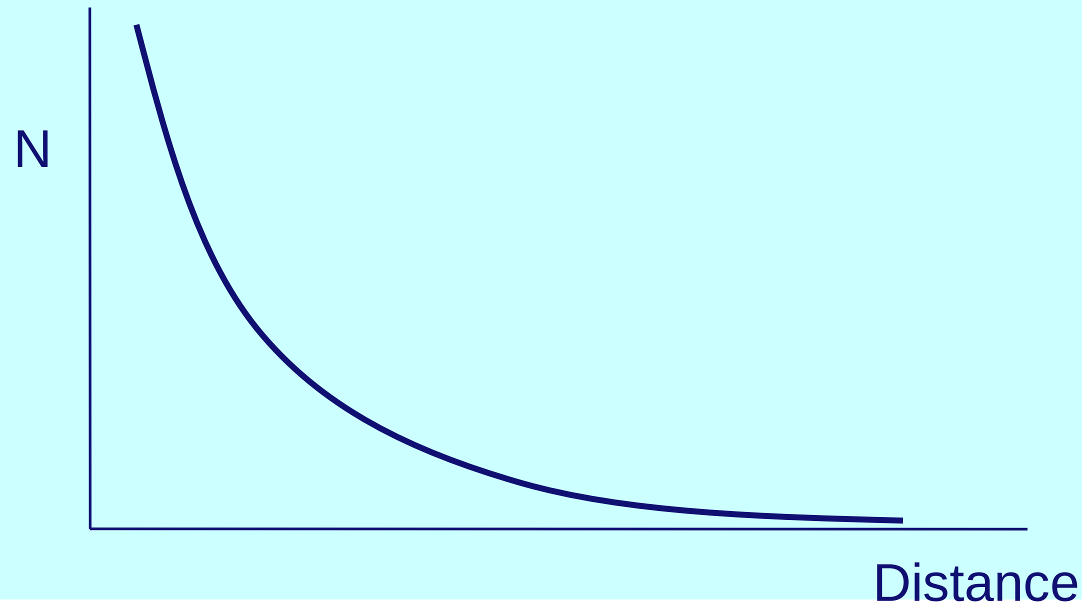
± 4000 postcode areas

The social network of someone in Eindhoven

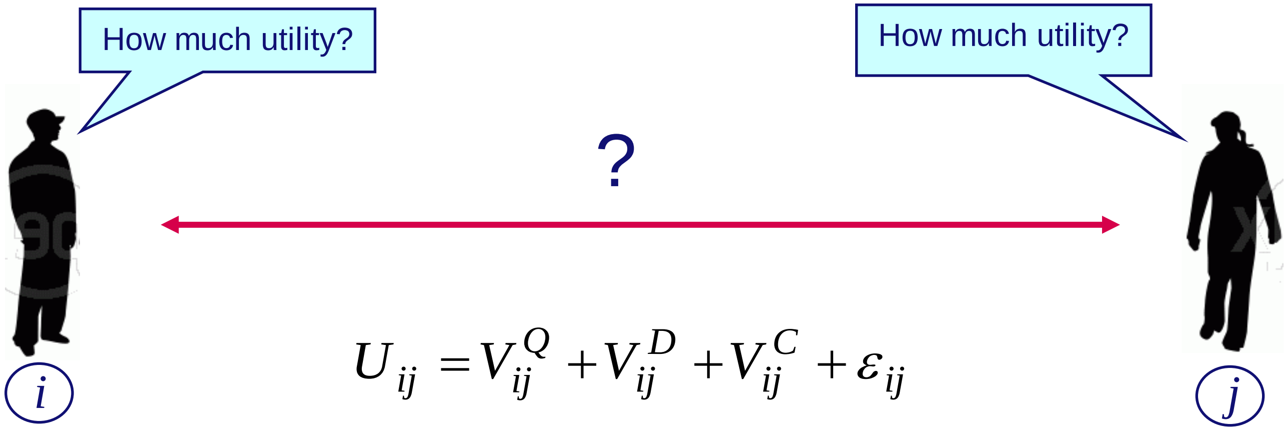


Average size of a personal social network = 24.6 links

Distribution of links across distance



The link selection model

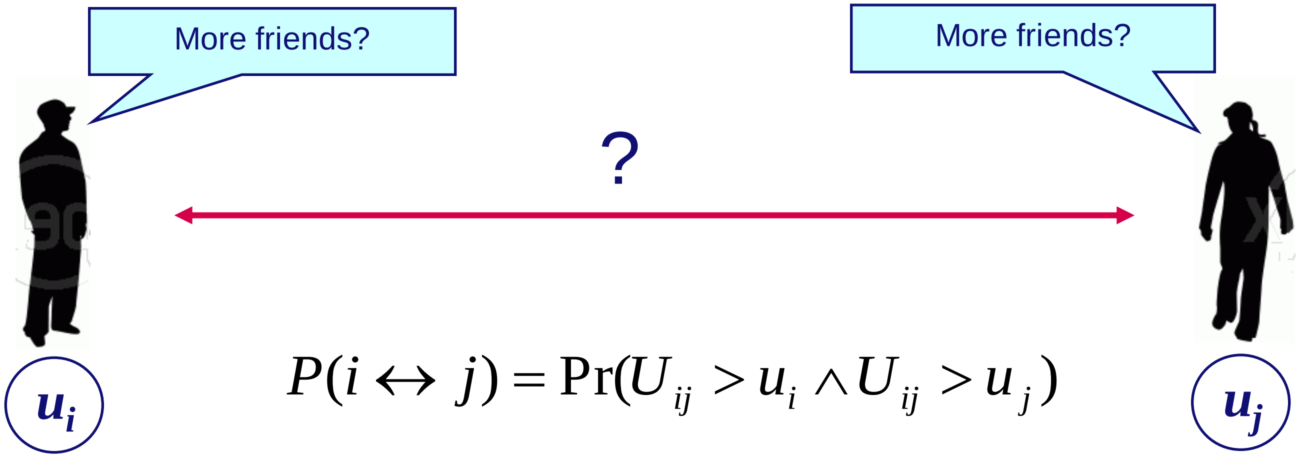


$$U_{ij} = V_{ij}^{Q} + V_{ij}^{D} + V_{ij}^{C} + \varepsilon_{ij}$$

Utility of a link depends on

- Similarity in attributes (homophily) (V_{ij}^{Q})
- Geographical distance (V_{ij}^{D})
- Having common friends (V_{ij}^{C})
- Randomness (unobserved factors) (ε_{ij})

The link selection model



$$P(i \leftrightarrow j) = \Pr(U_{ij} > u_i \wedge U_{ij} > u_j)$$

$$P(i \leftrightarrow j) = \Pr(U_{ij} > \max[u_i, u_j]) = \frac{\exp(V_{ij} - \max[u_i, u_j])}{1 + \exp(V_{ij} - \max[u_i, u_j])}$$

Social network data of the sample

Category	N	%	N / person
Partner	92	3.2	0.79
Father/mother	111	3.9	0.96
Child	192	6.7	1.66
Brother/sister	254	8.9	2.19
Other relative	670	23.5	5.78
Household member	10	0.4	0.09
Neigborhood member	208	7.3	1.79
Colleague	150	5.3	1.29
Fellow student	48	1.7	0.41
Union member	170	6.0	1.47
Other friend	740	25.9	6.38
Other acquaintance	200	7.0	1.72
Other	4	0.1	0.03
Total	2849	100.0	24.56

N / person = 11.4

N / person = 13.1

Social categories considered in the present analysis

Micro-simulation procedure

1. **Synthesize the entire Dutch population**

2. **Calculate a-priori probabilities $P(X, d)$**

3. **Estimate the link-selection model on link data**

4. **Use the model to create social links between individuals**

- **Test case: ego's from Eindhoven**

Link-selection model estimation results

Parameter	Estimate
Distance (delta)	-1.606
Gender similarity (beta)	0.713
Age similarity (beta)	0.888
Utility threshold base value (Z)	6.835
Utility threshold effect Male (Z)	0.512
Utility threshold effect Age < 40 yr (Z)	0.177
Utility threshold effect Age 60 -< 70 yr (Z)	-0.608
Utility threshold effect Age 70+ yr (Z)	-0.228
LL null model	-8832
LL final model	-4699
Rho square	0.468

The correct scale of the threshold values depend on the size of the population – needs to be calibrated -> 0.65

Results – social network statistics

Social network characteristics

	Observed	Predicted
N egos	105	152
Same gender ratio	0.715	0.678
Same age class ratio	0.550	0.592
Distance, mean (km)	26.1	27.4
Distance, std. dev. (km)	42.2	42.4
Network size, mean	12.8	12.7
Network size, std. dev.	9.6	5.3

Close match is result of fitting threshold values

Variance in network size is underestimated

➢ **Threshold values should be a random parameter**

Conclusions

- **We introduced an approach to micro-simulate population-wide social networks**

- **The proposed link-selection model is consistent with existing utility-based theories of link selection behavior**

- **A new method was developed to estimate the model on link data, which takes spatial characteristics of the population into account**

- **The synthesized network can be used to micro-simulate the social interactions in an entire population in geographical space**

TU/e Technische Universiteit Eindhoven University of Technology

Future research

- **Straightforward refinements of the utility function**
  - **incorporate a more exhaustive set of person attributes**
  - **allow asymmetric utility functions**
  - **elaborate the threshold function**

- **The model needs to be extended to take the Common-friends factor into account. This has implications for the larger system**

An agent-based framework for modeling social activities and travel

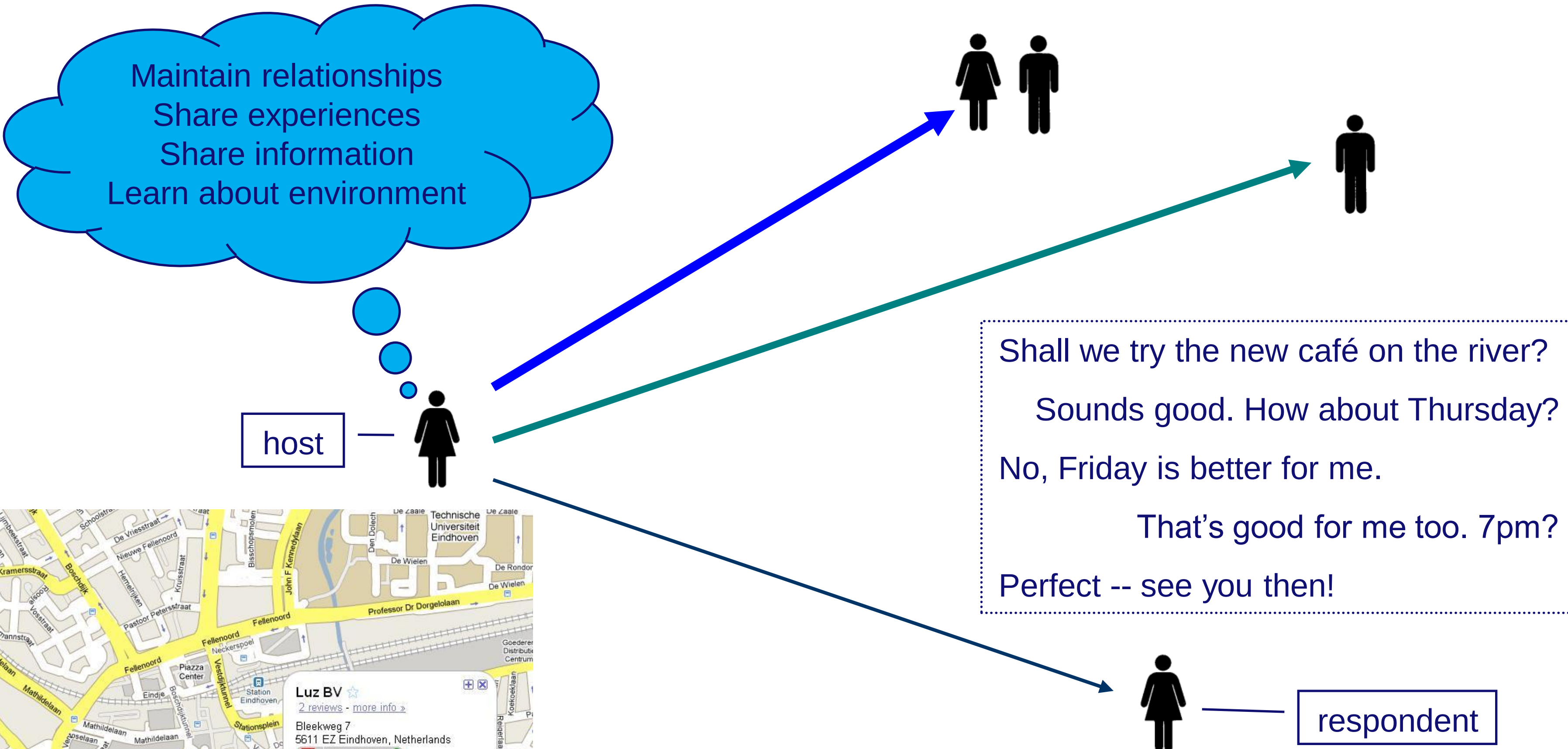
Nicole Ronald, Theo Arentze, Harry Timmermans

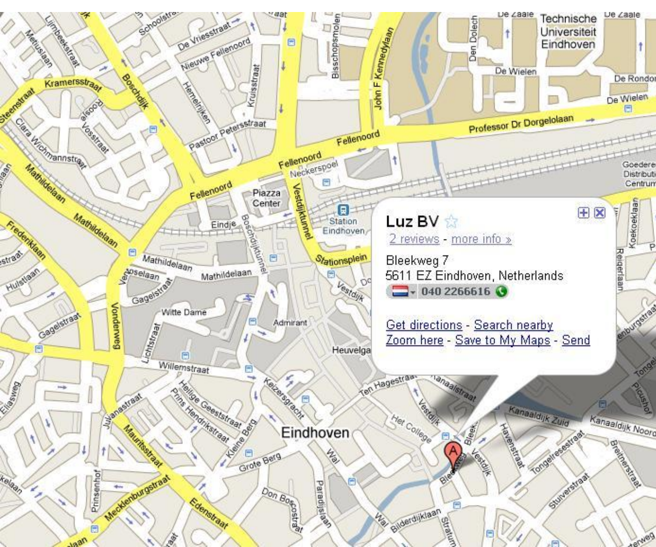
Objectives

- **To develop an agent-based framework to simulate**
  - **Social activities**
  - **Social selection and network dynamics**
  - **Social influence**

- **Activity scheduling**
  - **Current activity-based models: within-household interactions at best**
  - **Objective is to extend person interactions to whole social networks**
  - **A new process model is needed**

Model design



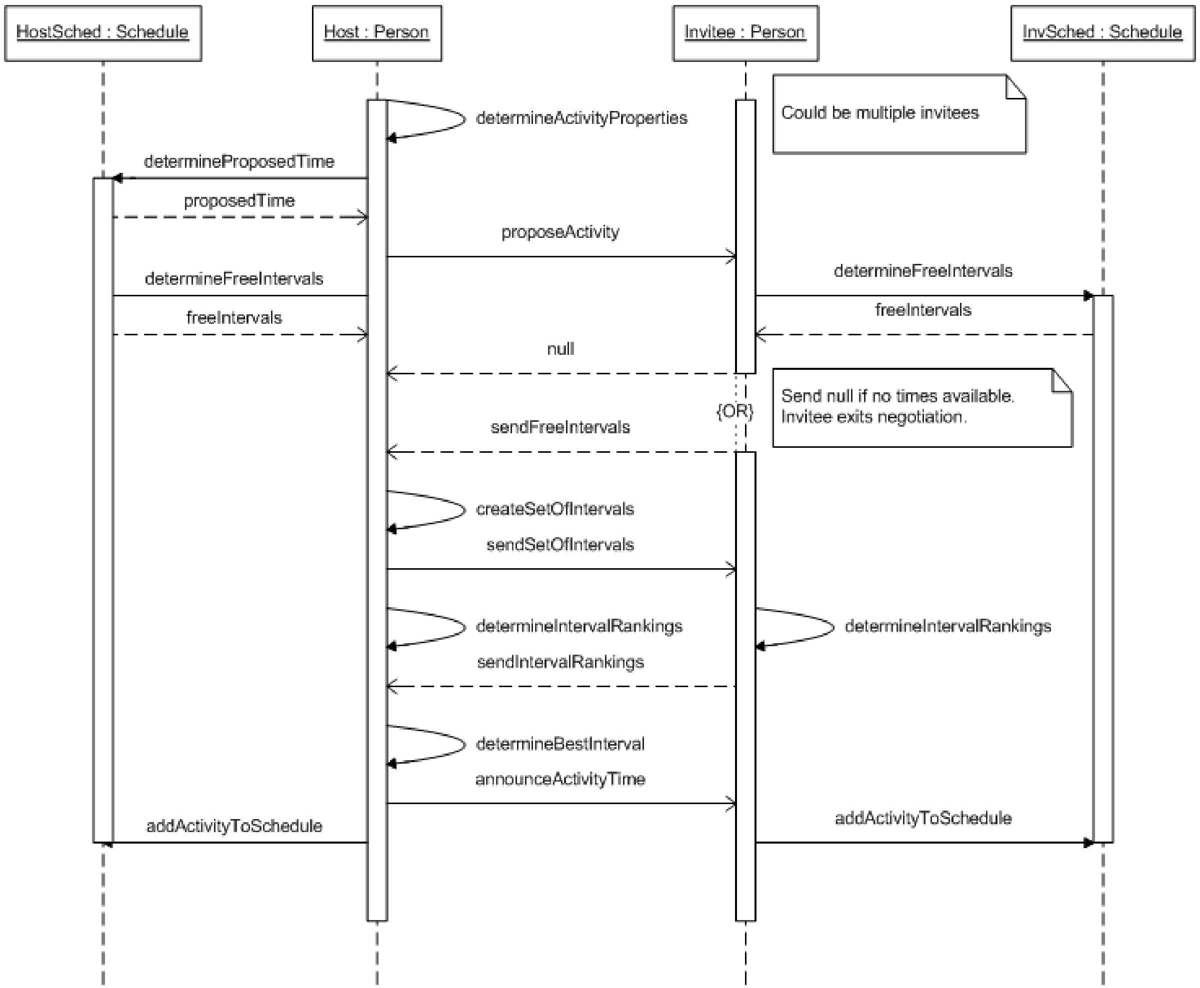


ms

Actions agents can perform

- Determining who to un/form a connection with
- Keeping a schedule
- Determining whether to initiate a discussion about an activity or send an invite
- Evaluating invite/activity proposals
- Undertaking activities
- Evaluating activities, in terms of satisfaction with location, people and time
- Updating own properties
- Sharing information about the environment with others
- Participating in a club, i.e., deciding to join/leave

Multi-issue negotiation protocol for joint activity scheduling



Utility function

$$U_{idy}(a,d,y,J,l,r) = f(r)(V_i^{apg} + V_{ij}^{aJ} + V_i^{apl} + \varepsilon_{ij})$$

- **Duration (*r*), day (*d*), time of day (*y*), type (*a*) and purpose (*p*) of the activity, location (*l*), and the people involved (*J*)**

- **Credit is a component of the utility function so that agents can make compromises**

Illustration

- **Prototype in Python with limited parameters**
- **100 individuals with contacts and locations**
- **Three activity profile preferences:**
  - **Prefer cultural/going out activities**
  - **Prefer home/greenspace activities**
  - **Prefer social activities**
- **Random environment with 100 locations**
- **Random starting social network**

Illustration - activity classification

Activity purpose

Location type

	Experience	Info	Social	Support
Cultural				
Green				
Home		**Specific interactions**		
Horeca				
…				

Illustration - utility function

$$U_i(a,p,l,d,J) = V_i^{ap} + V_i^{l} + V_i^{J} + \varepsilon$$

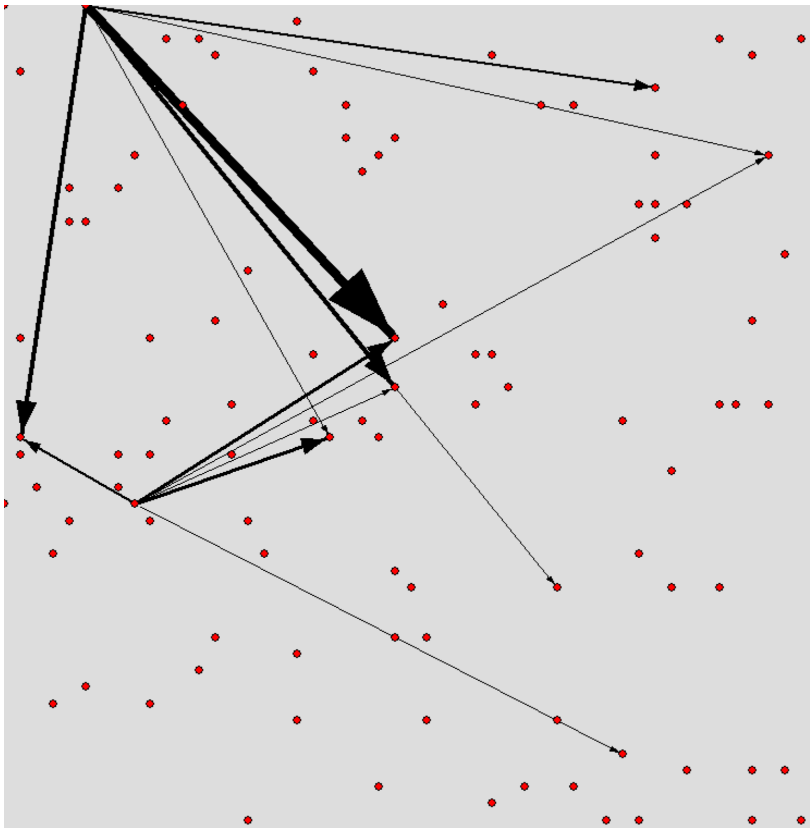
$$V_i^{ap}(d) = f_t\ (t_{ap}, \beta_i^{ap}) + \alpha_i^{ap}$$

$$V_i^{l}(d,a,p) = \beta_i^{l} f_t^{\,l}(t_l\ ) + \alpha^{apl'} - d_i^{l}$$

$$V_i^{J} = \sum_{j \in J} s_{ij} f_h\ (t_h)$$

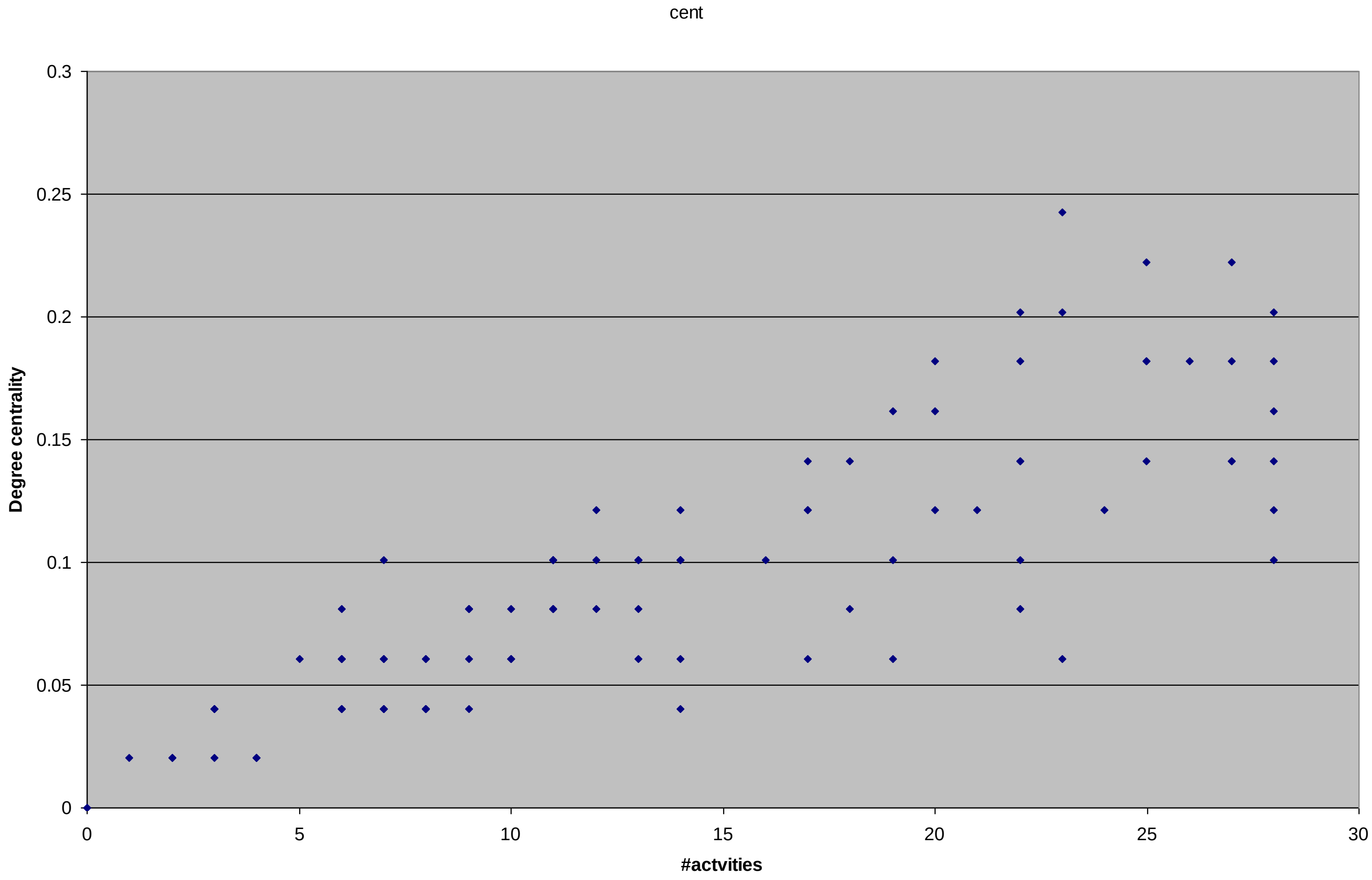
$$s_{ij} = (1 - \left(\frac{\mathrm{x}_a^i - \mathrm{x}_a^j}{\mathrm{r}_a}\right)^{\lambda_a})(1 - (\varepsilon_{gl(i)l(j)})^{\lambda_g})$$

Joint activities: example

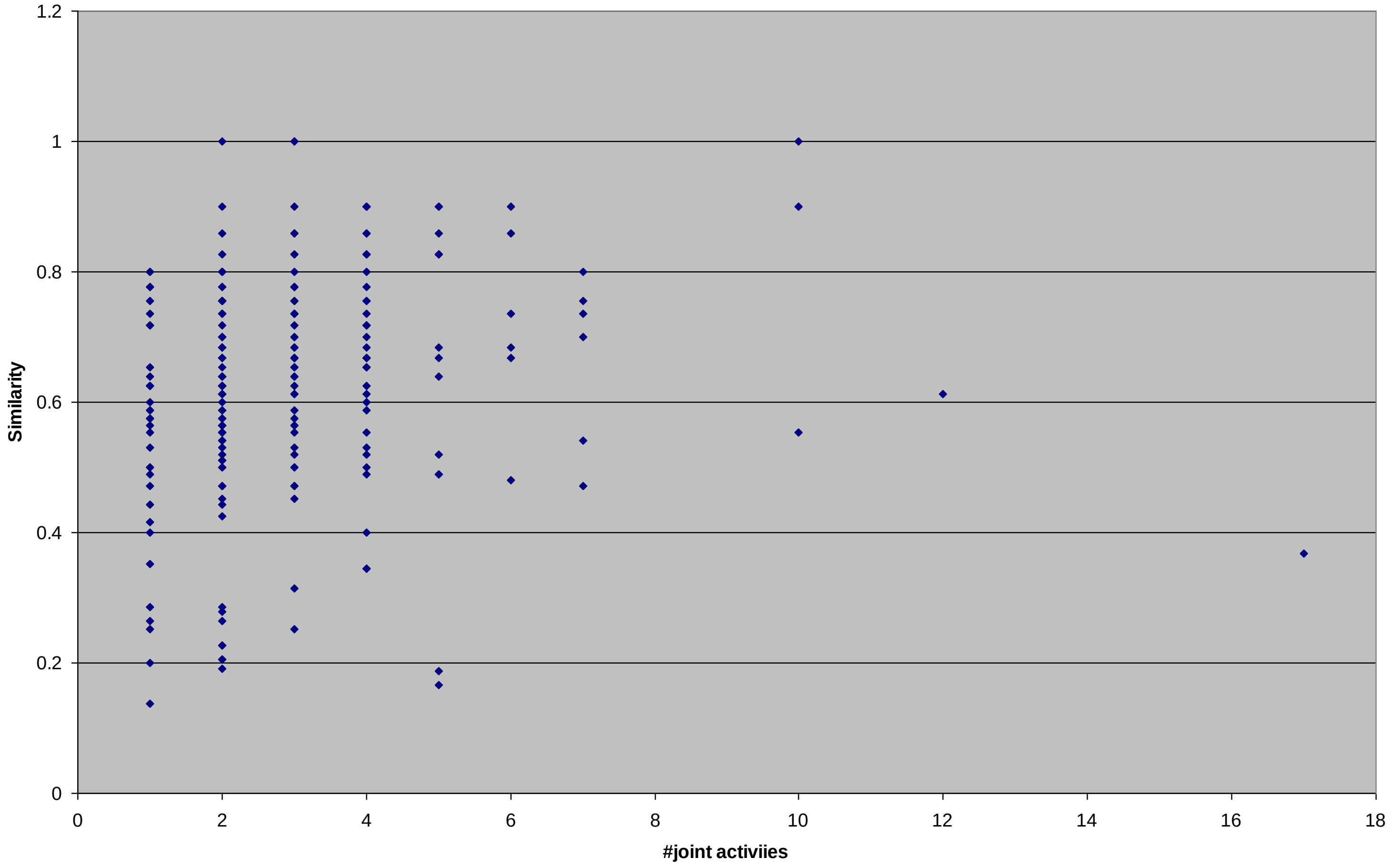


Day	Person 1	Person 2
29	museum	museum
30	horeca	
31	horeca	
32	horeca	
33	cinema	cinema
34	museum	park
35		
36		
37	horeca	
38	horeca	horeca
39		
40	horeca	home
41		park
42	horeca	

Some statistics



Some statistics



Conclusions

- **First step in developing an agent-based framework has been made**
- **A utility function and protocol for joint activity scheduling has been developed and tested**
  - **Deals with a full range of activity choice facets**
  - **Allows for limited information and credit built-up**
  - **Shows face validity**
- **Remaining steps**
  - **Social selection, influence and dynamics**
  - **Estimating parameters**

Conclusions overall studies

- **We identified three topic areas for social networks and travel behavior**
  - **Activity generation – how often, what**
  - **Social selection – where, with whom**
  - **Social influence – choice sets, choice behavior**

- **Studies showed approaches and results**
  - **Within household interaction**
  - **Data collection and data analysis**
  - **Synthesis of social networks**
  - **Agent-based simulation framework**

TU/e Technische Universiteit Eindhoven University of Technology

Discussion

- **Can we extend activity-based models to incorporate social networks?**
  - **Synthesis of networks √**
  - **Activity scheduling protocol √**
  - **Dynamics**
  - **Influence**

- **If we have, would it make a difference?**
  - **Extended task allocation**
  - **Extended coupling constraints**
  - **Non-linear effects and clustering of behavior**

Thank you for your attention!

Activity classification - Albatross

No	Activity Types	Grouped Activity	Personal (P) or Household (HH) Level	Activities
1	Work	Mandatory	P	Full-time and part-time
2	Business		P	Work-related
3	Other		P	Other mandatory activities (school, etc)
4	Bring/get person	Household Task	HH	Drop-off/pick-up children or spouse to a certain location
5	Shop-1-store		HH	Daily shopping
6	Shop-n-store		HH	Non-daily shopping
7	Service-related		HH	Renting movie,getting (fast) food, institutional purposes (bank, post office, etc)
8	Social-alone	Non Household Task	P	Meeting friends, religions, etc
	Social-joint		HH	Meeting relatives, social activies, etc
9	Leisure-alone		P	Sports,café/bar,eating out,movie,museum,library
	Leisure-joint		HH	Recreational activities with children, café/bar,eating out, movie, museum, library
10	Touring-alone		P	Making a tour by car,bike,or foot (eg., letting out the dog, etc)
	Touring-joint		HH	

Validiation: sensitivity testing

INDICATORS	Old model		New model	
	m_1-m_0 (%)	sign	m_1-m_0 (%)	sign
Total travel time	-3.89		-3.38	
Travel time car driver	4.56	**	5.08	**
Travel time public	-36.74		-36.33	
Travel time slow	0.15		0.94	
Travel time car passenger	0.45		0.2	
Number of tours	1.11		1.32	**
Number of trips	1.16		1.63	**
Ratio trips-tours	0.05		0.31	**
Total travel distance	4.09	**	4	**
Distance car driver	4.73	**	4.93	**
Distance car passenger	1.01		0.36	
Distance slow	1.11	*	2.03	**
Distance public	5.79	**	5.2	**

The models predict different effects regarding the number of trips and ratio of trips-tours

Validation: distributions (example)

Activity Type	Observed Data (%)	Predicted Data (%)	
		Old Version	New Version
Work	20.47	18.88	18.39
Business	5.8	6.43	5.79
Bring-Get	7.96	8.3	8.65
Shop-1 store	20.92	22.53	21.61
Shop-n store	4.07	4.46	3.97
Service	5.28	5.65	5.08
Social	13.17	11.67	13.86
Leisure	12.9	12.51	12.92
Tour	8.04	8.16	8.27
Other	1.39	1.41	1.47
Total	82584	76842	78812

Predictions of activity choice on an aggregate level at least as accurate as the existing model

Validation: distributions (example)

INDICATORS	Observed		Old model		New model	
	Mean	Stdev	Mean	Stdev	Mean	Stdev
Total travel time	63.9	71	45.7	60.8	46.5	57.4
Travel time car driver	31.3	54.8	23.3	40.7	22.6	38.8
Travel time public	5.3	30.8	4.5	28.1	4.6	27.4
Travel time slow	18.6	39.5	12.9	38.6	13.6	35.1
Travel time car passenger	8.7	31	4.6	17.4	5.5	19.1
Number of tours	1.4	1	1.3	1	1.4	1.1
Number of trips	3.1	2.4	2.9	2.4	3.1	2.5
Ratio trips-tours	2.30		2.27		2.21	
Total travel distance	38.7	82.4	35	64.6	34.8	61.8
Distance car driver	23.8	62.7	25.8	57	24.4	53.6
Distance car passenger	7.1	35.4	4.9	23.3	6	26.1
Distance slow	4.8	34.3	1.7	6.9	1.7	6.5
Distance public	3.1	24.2	2.7	24.1	2.7	22.5

Predictions of mobility indicators on an aggregate level at least as accurate as the existing model

Validiation: sensitivity testing

- **Scenario**
  - **Increase of 40% of labor participation of women**

- **Analysis method**
  - **Create a synthetic population for the Netherlands for baseline and scenario**
  - **Run the model to predict activity schedules for the baseline and scenario: old and new model**
  - **Compare predicted effects between old and new model**

Validiation: sensitivity testing

Activity Type	Old model		new model	
	m1-m0 (%)	sign	m_1-m_0 (%)	sign
Work	13.15	**	12.79	**
Business	9.89	**	10.78	**
Bring-Get	-4.98	**	2.6	**
Shop-1-store	-5.58	**	-3.15	**
Shop-n-store	-6.54	**	-3.81	**
Service	-1.33		2.2	**
Social	-0.66		-1.21	**
Leisure	1.76	**	-0.69	
Tour	0.33		-3.53	**
Other	-6.84	**	-3.15	*
Total	1.2		1.89	**

Household task activities

The old model predicts a decrease

The new model predicts a slight increase in bring-get and service activities and smaller decreases of the other household activities

Validiation: sensitivity testing

Activity Type	Old model		new model	
	m1-m0 (%)	sign	m_1-m_0 (%)	sign
Work	13.15	**	12.79	**
Business	9.89	**	10.78	**
Bring-Get	-4.98	**	2.6	**
Shop-1-store	-5.58	**	-3.15	**
Shop-n-store	-6.54	**	-3.81	**
Service	-1.33		2.2	**
Social	-0.66		-1.21	**
Leisure	1.76	**	-0.69	
Tour	0.33		-3.53	**
Other	-6.84	**	-3.15	*
Total	1.2		1.89	**

Social and leisure activities

The old model predicts no change or an increase

The new model predicts a decrease at least for social and touring activity

TU/e Technische Universiteit Eindhoven University of Technology