Agent-based modelling for transport planning

KW Axhausen

IVT
ETH
Zürich

September 2019
In lieu of acknowledgements: “Local” contributors

Prof. Kay Axhausen
Dr. Milos Balac
Dr. Michael Balmer
Henrick Becker
Dr. Joschka Bischoff
Dr. David Charypar
Billy Charlton
Dr. Nurhan Cetin
Dr. Artem Chakirov
Yu Chen
Dr. Francesco Ciari
Dr. Christoph Dobler
Dr. Alexander Erath
Dr. Matthias Feil
Dr. Gunnar Flötteröd
Dr. Pieter Fourie
Dr. Christian Gloor
Dr. Dominik Grether
Dr. Jeremy K. Hackney
Dr. Andreas Horni
Sebastian Hörl
Anugrah Ilahi
Ihab Kaddoura
Grace Kagho
Nicolas Lefebvre
Clarissa Livingston
Dr. Johannes Illenberger
Dr. Gregor Lämmel
Dr. Michal Maciejewski
Patrick Manser
Dr. Konrad Meister
Dr. Ming Lu
Joe Molloy
Manuel Moyo
Dr. Krill Müller
Prof. Kai Nagel
Dr. Andreas Neumann
Dr. Thomas Nicolai
Dr. Benjamin Kickhöfer

Dr. Sergio Ordonez
Stefano Penazzi
Dr. Bryan Raney
Dr. Marcel Rieser
Dr. Nadine Schüssler
Dr. Lijun Sun
Dr. David Strippgen
Christopher Tchervenkov
Therasa Thunig
Dr Michael van Eggermond
Dr. Rashid Waraich
Dominik Ziemke
Dr. Michael Zilske
Basic assumptions for transport planning
Basic definitions

Social generalised costs is the sum of

individual generalised costs, i.e. decision relevant generalised costs & overlooked individual costs

And the

externalities caused
Basic assumption 1

Accessibility ~
Opportunities, Speeds
Basic assumption 2

Traffic is a system of moving, self-organising

Queues
Basic assumption 3

The crucial short-term interaction between capacity, i.e. the

number of *slots*

for the desired speed and the

**current demand**
Basic assumption 4

Societies chose their number of *slots* by the design/operation of the road/rail/bike network for the desired speeds.
Basic assumption 5

Travel demand (pkm or tkm) is a

normal good

i.e. it grows with

decreasing individual “generalised costs”
Basic assumption 6

Decision relevant generalised costs are the

sum of the risk and comfort weighted monetary expenditure and the time spent
The travellers chose their average decision relevant generalised costs with their package of locations (residence, work) and mobility tools.
A person’s travel demand is the result of its out-of-home activity participation constrained by the currently available time and money resources and their chosen average generalised costs.
A person’s travel experience is the result of the queues (joined or avoided) and can be addressed by mostly costly changes.
How do we label our models?
### A terminological problem?

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution</td>
<td>Agents, flows</td>
</tr>
<tr>
<td>Scheduling model</td>
<td>Trip, tour, daily chain (with breaks)</td>
</tr>
<tr>
<td>Choice model</td>
<td>DCM, rules&amp;heuristics</td>
</tr>
<tr>
<td>Route choice</td>
<td>Integrated, external (with consistent valuations?)</td>
</tr>
<tr>
<td>Choice set construction</td>
<td>Explicit, implicit</td>
</tr>
<tr>
<td>Solution method</td>
<td>Whole population (&amp; MSA or similar)</td>
</tr>
<tr>
<td></td>
<td>Sample enumeration (&amp; MSA or similar), co-evolutionary search</td>
</tr>
<tr>
<td>Schedule equilibrium</td>
<td>Yes, no</td>
</tr>
</tbody>
</table>
## The typical four-stage model

<table>
<thead>
<tr>
<th>Resolution</th>
<th>Agents, flows</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheduling model</td>
<td>Trip, tour, daily chain (with breaks)</td>
</tr>
<tr>
<td>Choice model</td>
<td>DCM, rules&amp;heuristics</td>
</tr>
<tr>
<td>Route choice</td>
<td>Integrated, <strong>external without consistent valuations</strong></td>
</tr>
<tr>
<td>Choice set construction</td>
<td><strong>Explicit</strong>, implicit</td>
</tr>
<tr>
<td>Solution method</td>
<td><strong>Whole population</strong> (&amp; MSA or similar)</td>
</tr>
<tr>
<td></td>
<td>Sample enumeration (&amp; MSA or similar), co-evolutionary search</td>
</tr>
<tr>
<td>Schedule equilibrium</td>
<td>(Yes), no</td>
</tr>
</tbody>
</table>
## The typical activity-based model (ABM)

<table>
<thead>
<tr>
<th>Resolution</th>
<th>Agents, flows</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheduling model</td>
<td>Trip, tour, <strong>daily chain</strong> (with breaks)</td>
</tr>
<tr>
<td>Choice model</td>
<td><strong>DCM</strong>, rules&amp;heuristics</td>
</tr>
<tr>
<td>Route choice</td>
<td>Integrated, <strong>external without consistent valuations</strong></td>
</tr>
<tr>
<td>Choice set construction</td>
<td><strong>Explicit</strong>, implicit</td>
</tr>
<tr>
<td>Solution method</td>
<td>Whole population (&amp; MSA or similar) <strong>Sample enumeration</strong> (&amp; MSA or similar), co-evolutionary search</td>
</tr>
<tr>
<td>Schedule equilibrium</td>
<td>Yes, <strong>none reported it yet</strong></td>
</tr>
<tr>
<td>MATSim</td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td><strong>Resolution</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Agents, flows</td>
</tr>
<tr>
<td><strong>Scheduling model</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trip, tour, <strong>daily chain</strong> without breaks</td>
</tr>
<tr>
<td><strong>Choice model</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>DCM</strong> and/or <strong>rules</strong> &amp; heuristics</td>
</tr>
<tr>
<td><strong>Route choice</strong></td>
<td></td>
</tr>
<tr>
<td><strong>external</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Integrated with consistent valuations</strong>,</td>
</tr>
<tr>
<td><strong>Choice set construction</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Explicit, <strong>implicit</strong></td>
</tr>
<tr>
<td><strong>Solution method</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Whole population (&amp; MSA or similar)</td>
</tr>
<tr>
<td></td>
<td>Sample enumeration (&amp; MSA or similar),</td>
</tr>
<tr>
<td></td>
<td><strong>co-evolutionary search</strong></td>
</tr>
<tr>
<td><strong>Schedule equilibrium</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Yes</strong>, no</td>
</tr>
</tbody>
</table>
What is the task?
# Time horizon of transport planning

<table>
<thead>
<tr>
<th></th>
<th>System</th>
<th>Person</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Long term</strong></td>
<td><em>slots</em></td>
<td>Home and work locations</td>
</tr>
<tr>
<td></td>
<td>Regulation</td>
<td>Mobility tool ownership</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Social networks</td>
</tr>
<tr>
<td><strong>Medium term</strong></td>
<td>Services</td>
<td>Season tickets</td>
</tr>
<tr>
<td></td>
<td>Prices</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Awareness</td>
<td></td>
</tr>
<tr>
<td><strong>Short term</strong></td>
<td>Operations</td>
<td><strong>Scheduling</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Generic model structure

```
“Scenario”

Competition for slot in facilities and the network

k(t,r,j)_{i,n}

Mental map

Scheduling

Population \( \beta_{i,t, r,j,k} \)

\[ q_i \equiv (t,r,j)_{i,n} \]
```
Market model

For all goods i of the market:

\[ k'_{i, togz} = f(q'_{i, togz} (k'_{i, toqz}, B_{ogz}), A_{i, togz}) \]

- \( k' \): Estimated generalised costs [SFr/good]
- \( q' \): Estimated demand [Elements/Unit time]
- \( A \): Supply of the goods
- \( B \): Population (natural and legal)
- \( t \): Time t
- \( o \): Place o
- \( g \): Group g
- \( z \): Year z
Key points of the critique of equilibrium approaches

- Travel is derived demand, with some exceptions
- The travellers are constrained by their commitments and mobility tool ownership
- Travellers aren’t in equilibrium
- Travellers don’t know all alternatives
- Travellers don’t plan their whole day (week) in advance
MATSim – A GNU open source project
MATSim: A GNU public licence software project

Main partners:
• TU Berlin (Prof. Nagel)
• ETH Zürich & FCL Singapore
• Senozon, Zürich (Dr. Balmer)
• Simunto, Zürich (Dr. Rieser)

Contributors, users, e.g.:
• TU Poznan
• University of Pretoria
• SBB, Bern

• Systems Group, DINF, ETH Zürich
2018 status
# Current status

<table>
<thead>
<tr>
<th>Known implementations:</th>
<th>About 45</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research groups:</td>
<td>About 35 (including some beyond transport)</td>
</tr>
<tr>
<td>Uses:</td>
<td>Research</td>
</tr>
<tr>
<td></td>
<td>Some initial commercial uses</td>
</tr>
<tr>
<td></td>
<td>Some policy consulting</td>
</tr>
<tr>
<td>Software:</td>
<td>Last reimplementation in 2012/13</td>
</tr>
<tr>
<td></td>
<td>Stable API</td>
</tr>
<tr>
<td></td>
<td>Daily tests</td>
</tr>
<tr>
<td></td>
<td>JAVA</td>
</tr>
</tbody>
</table>
Current progress: Singapore
MATSim: Base approach
Equilibrium search in MATSim

Initial schedules

Simulation of flows on networks and to facilities

Score (utility) calculation

Optimal Replanning (inc. connection) & plan choice

$q_i \equiv (t,r,j)_{i,n}$

$k(t,r,j)_{i,n}$

$U_i(t,r,j)_{i,n}$
Following the agents
MATSim: Logic of the co-evolution – Step 0

Agent 1
  Plan 1.1  H-W-H; 8:00, 17:00; C,C;

Agent 2
  Plan 2.1  H-W-H; 8:00, 17:00; C,C;

Agent 3
  Plan 3.1  H-W-H; 8:00, 17:00; C,C;
## Co-evolution – Step 1.1 – Simulation/scoring

<table>
<thead>
<tr>
<th>Agent</th>
<th>Plan 1.1</th>
<th>Time: 8:00, 17:00; C,C;</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agent 1</td>
<td>Plan 1.1</td>
<td>H-W-H; 8:00, 17:00; C,C;</td>
<td>35</td>
</tr>
<tr>
<td>Agent 2</td>
<td>Plan 2.1</td>
<td>H-W-H; 8:00, 17:00; C,C;</td>
<td>35</td>
</tr>
<tr>
<td>Agent 3</td>
<td>Plan 3.1</td>
<td>H-W-H; 8:00, 17:00; C,C;</td>
<td>35</td>
</tr>
</tbody>
</table>
Co-evolution – Step 1.2 – After replanning (1/3)

Agent 1
  Plan 1.1  H-W-H; 8:00, 17:00; C,C; 35

Agent 2
  Plan 2.1  H-W-H; 8:00, 17:00; C,C; 35

Agent 3
  Plan 3.1  H-W-H; 8:00, 17:00; C,C; 35
  Plan 3.2  H-W-H; 8:15, 17:30; C,C
Co-evolution – Step 1.3 – After plan selection (best/MNL)

Agent 1
Plan 1.1  H-W-H; 8:00, 17:00; C,C;  100%

Agent 2
Plan 2.1  H-W-H; 8:00, 17:00; C,C;  100%

Agent 3
Plan 3.1  H-W-H; 8:00, 17:00; C,C;  35
Plan 3.2  H-W-H; 8:15, 17:30; C,C;  New
Co-evolution – Step 2.1 – Simulation/scoring

Agent 1
Plan 1.1 H-W-H; 8:00, 17:00; C,C; 45

Agent 2
Plan 2.1 H-W-H; 8:00, 17:00; C,C; 45

Agent 3
Plan 3.1 H-W-H; 8:00, 17:00; C,C; 35
Plan 3.2 H-W-H; 8:15, 17:30; C,C; 60
|
|---|
|**Co-evolution – Step 2.2 – After replanning (1/3)**|
|**Agent 1**| |
|Plan 1.1 | H-W-H; 8:00, 17:00; C,C; | 45 |
|Plan 1.2 | H-W-H; 8:00, 17:00; B,B; | |
|**Agent 2**| |
|Plan 2.1 | H-W-H; 8:00, 17:00; C,C; | 45 |
|**Agent 3**| |
|Plan 3.1 | H-W-H; 8:00, 17:00; C,C; | 35 |
|Plan 3.2 | H-W-H; 8:15, 17:30; C,C; | 60 |
### Co-evolution – Step 2.3 – After plan selection (best/MNL)

<table>
<thead>
<tr>
<th>Agent 1</th>
<th>Plan 1.1</th>
<th>H-W-H; 8:00, 17:00; C,C;</th>
<th>45</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Plan 1.2</td>
<td>H-W-H; 8:00, 17:00; B,B;</td>
<td>New</td>
</tr>
</tbody>
</table>

| Agent 2          | Plan 2.1 | H-W-H; 8:00, 17:00; C,C; | 100% |

| Agent 3          | Plan 3.1 | H-W-H; 8:00, 17:00; C,C; | 38%  |
|                  | Plan 3.2 | H-W-H; 8:15, 17:30; C,C; | 62%  |
Co-evolution – Step 3.1 – Simulation/scoring

Agent 1
- Plan 1.1  H-W-H; 8:00, 17:00; C,C;  45
- Plan 1.2  H-W-H; 8:00, 17:00; B,B;  70

Agent 2
- Plan 2.1  H-W-H; 8:00, 17:00; C,C;  45

Agent 3
- Plan 3.1  H-W-H; 8:00, 17:00; C,C;  45
- Plan 3.2  H-W-H; 8:15, 17:30; C,C;  60
## Co-evolution – Step 3.2 – After replanning (1/3)

### Agent 1
- **Plan 1.1**
  - H-W-H; 8:00, 17:00; C,C;
  - 45
- **Plan 1.2**
  - H-W-H; 8:00, 17:00; B,B;
  - 70

### Agent 2
- **Plan 2.1**
  - H-W-H; 8:00, 17:00; C,C;
  - 45

### Agent 3
- **Plan 3.1**
  - H-W-H; 8:00, 17:00; C,C;
  - 45
- **Plan 3.2**
  - H-W-H; 8:15, 17:30; C,C;
  - 60
- **Plan 3.3**
  - H-W-H; 7:30, 17:15; B,B
Co-evolution – Step 3.3 – After plan selection (best/MNL)

Agent 1
- Plan 1.1  H-W-H; 8:00, 17:00; C,C; 36%
- Plan 1.2  H-W-H; 8:00, 17:00; B,B; 64%

Agent 2
- Plan 2.1  H-W-H; 8:00, 17:00; C,C; 100%

Agent 3
- Plan 3.1  H-W-H; 8:00, 17:00; C,C; 45
- Plan 3.2  H-W-H; 8:15, 17:30; C,C; 60
- Plan 3.3  H-W-H; 7:30, 17:15; B,B  New

(The (worst) plan, more then memory allows, is deleted)
## Co-evolution – Summary of best scores

<table>
<thead>
<tr>
<th></th>
<th>Iteration 1</th>
<th>Iteration 2</th>
<th>Iteration 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agent 1</td>
<td>35</td>
<td>45</td>
<td>80</td>
</tr>
<tr>
<td>Agent 2</td>
<td>35</td>
<td>45</td>
<td>45</td>
</tr>
<tr>
<td>Agent 3</td>
<td>35</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td><strong>35</strong></td>
<td><strong>50</strong></td>
<td><strong>62</strong></td>
</tr>
</tbody>
</table>
SUE search example

\[ \beta_{\text{score}} = 2.0, \alpha = 1.0 \]

- \( \text{pct}_{\text{SUE, min}} = 97.5\% \)
- \( \text{pct}_{\text{SUE, select}} \)
- \( \text{pct}_{\text{SUE, router}} \)
- \( \text{pct}_{\text{SUE, tam}} \)
- \( \text{pct}_{\text{SUE, all}} \)
- \( \text{avg}(V_{\text{plan, worst}}) \)
- \( \text{avg}(V_{\text{plan, best}}) \)
- \( \text{avg}(V_{\text{plan, executed}}) \)

iteration

LVMT 19
Co-evolution – Issues

• Size of search space ~ Behavioural alternatives

• Rate of replanning (~ MSA)

• Size of the choice set ~ RAM

• Similarity of the daily schedules
• Integration into a log-sum term
Activity schedule dimensions
Activity scheduling dimensions

Number and type of activities
Sequence of activities

• Start and duration of activity
• Composition of the group undertaking the activity
• Expenditure division
• Location of the activity

• Movement between sequential locations
  • Location of access and egress from the mean of transport
    • Parking type
  • Vehicle/means of transport
  • Route/service
  • Group travelling together
  • Expenditure division
Current Vickrey-type utility function

\[ U_{plan} = \sum_{i=1}^{n} U_{act,i} + \sum_{i=2}^{n} U_{trav,i-1,i} \]

\[ U_{act,i} = U_{dur,i} + U_{late\_ar,i} \]
## Future whole day utility function?

<table>
<thead>
<tr>
<th>Time elements</th>
<th>linear</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Travel time</td>
<td>By mode and type of service;</td>
</tr>
<tr>
<td></td>
<td>by crowding level</td>
</tr>
<tr>
<td></td>
<td>by comfort level (parking search, stop&amp;go)</td>
</tr>
<tr>
<td>• Transfer penalty</td>
<td></td>
</tr>
<tr>
<td>• Late penalty</td>
<td>by activity type</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Activity time</th>
<th>log (Vickrey) or S-shape (Joh) (all, individual)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Minimum duration</td>
<td>by activity type</td>
</tr>
<tr>
<td>• Preferred duration</td>
<td>by activity type</td>
</tr>
<tr>
<td>• Duration</td>
<td>by time of day (might go away if participation is included)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Destination</th>
<th>Attractiveness, Value for money (on-line, off-line)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expenditure</td>
<td>by activity</td>
</tr>
<tr>
<td></td>
<td>by mode/type of service</td>
</tr>
</tbody>
</table>

LVMT 19
Schedule detail possibilities (in current stable MATSim)

Number and type of activities
(Stéphane Balac)
Sequence of activities
(William Ordonez)

- Start and duration of activity
- Composition of the group undertaking the activity (Francois Dubernet, Warren Fourie)
- Expenditure division
- Location of the activity (Friedrich Hörl, Jan Vitins)

- Movement between sequential locations

- Location of access and egress from the mean of transport
  - Parking search and type (Rajiv Waraich)
  - Vehicle/means of transport (Hans Bösch, Friedrich Hörl)
  - Route/service
  - Group travelling together (Francois Dubernet, Warren Fourie)

- Expenditure division
Finding short cuts
Turning Big Data into Smart Data

Data → Models (Dwell times, Speed, Behavior) → Simulation (MATSim) → Insight
Dwell time model

Boarding and alighting process

Results of statistical model

Critical occupancy at 63% of total capacity.

Low floor allows short dwell processes.

Double decker alighting time per pax 0.285 seconds longer.

With higher occupancy and number of boarding and alighting passenger -> shorter activity time

Heteroscedasticity of dwell times

Accounting for travel time variability

- Road network
- Travel time distributions by time of day
- Stop to stop link
- Dwell link
Modelling stop to stop travel times

Derive from Smart Card Data records travel times between stops

Each observed travel time between two subsequent stops constitutes one observation

Independent variables to be either derived from smart card data or GIS data, but do not require any other data source (e.g. traffic flow)

Static variables

- Availability of bus lane
- Number of intersections
- Number of left/right turns
- Curviness
- Deviation from crowfly distance
- Number of traffic lights
- Intersection density

Time-dependent variables

- Boarding/alighting activities in 500m radius
Validation

**Bus speed**

- MATSim
- EZ Link

**Transfer times**

- MATSim
- EZ Link

**Trip duration (Bus)**

- MATSim
- EZ Link

**Journey duration all modes**

- MATSim
- EZ Link

LVMT 19

N = 880230  Bandwidth = 60

Access, egress times removed from matsim bus times
Case study: network reconfiguration

Evaluation of new services and routes:

- How can new network designs improve reliability and tackle overcrowding?
- How many passengers will be attracted by a new service?

Simulation and analysis:

- A full day simulated in just about 40 minutes.
- Leverage on off-the-shelf business analytic software for interactive analysis.
The reliability of a long bus line
The effect of splitting the line

Time-space graph after split

Stop number

Occupancy

Split

Time of day
Reliability before and after line split

**Before line split**

**After line split**

Stop number

Time of day

Occupancy

0

128

LVMT 19
Challenges
Challenges for MATSim

- Econometric estimation of the whole day scoring function
- Increase the size and variance of the implicit choice set
- Link to a log-sum formulation (Chakirov)
- Accelerating the iterative equilibrium search
- Gridlock modeling (& stability of equilibrium)
- Modelling “irrational/uninformed” behaviours
- Generation of artificial social networks in the agent-population
- Co-generation of joint activities
- Multiple agent-type equilibria, e.g. stores, PUDOs, agents
Questions?

www.matsim.org
www.ivt.ethz.ch
www.futurecities.ethz.ch
www.senozon.com
www.simunto.com
Conclusions for modelling
Conclusions for modelling

We have to account for

**self-selection everywhere**

And we have to account for

**spatial-temporal correlations** and

**joint choices producing the queues**
Conclusions for modelling

We have to better understand the

**system capacities (e.g. mMFD)**

And the willingness to

**costly change (individual/joint) behaviour** and

**joint decision-making (group; collectives)**