Dynamic normative and descriptive optimal strategies for intelligent transit systems: a unified Markov decision problem approach

Nuzzolo A., Comi A.,
{comi; nuzzolo}@ing.uniroma2.it

Department of Enterprise Engineering

XXII SIDT National Scientific Seminar
TOWARDS SUSTAINABLE CITIES: INNOVATIVE SOLUTIONS FOR MOBILITY AND LOGISTICS
POLITECNICO DI BARI – Aula Magna “Attilio Alto” 14 – 15 SETTEMBRE 2017

Session G2-PM: Methods and models for greener transportation systems I
(Aula Magna A. Alto), Thursday 14th of Sept. 16.15-18.00
Summary

✓ Dynamic strategy-based path choice in unreliable transit networks
✓ Optimal strategy as optimal policy of a Markov Decision Problem
  ➢ Transition probabilities and expected rewards

✓ Optimal strategy search in Intelligent Transit Trip Planners
  ➢ Data-driven transition probabilities and expected rewards
  ➢ Normative optimal strategy in advanced trip planners.

✓ Subjective optimal strategies of traveler’s in real-time simulation models
  ✓ (Traveller-driven transition probabilities and expected rewards)

✓ Concluding remarks and possible developments
Stochastic (Unreliable) Transit Networks
Given an O-D pair of a transit network, let $SG$ the sub-graph including the feasible paths connecting O-D.

If some path attributes $X$ (e.g. waiting time, on-board time, on-board occupation degree,.....) are random variables, the network $N$ is classified as stochastic (unreliable) service network $SN$.

In the following, only ordinary randomness are considered; e.g. effects due to large service disruptions are not considered.
Path choice in stochastic networks

In unreliable or stochastic transit networks, in order to take into account, at diversion nodes, the stochastic occurrences (e.g. arrival time of buses at a stop, transit vehicle crowding, failure to board) and the updated attribute forecasts (if available), a real-time dynamic travel strategy should be used.
A travel strategy (Spiess, 1983) is a set of coherent decision rules which, for each diversion node, establish which diversion links are to be considered and how to choice among them (diversion rule), with the objective to optimise an objective function.

For example, the objective can be, according the Decision Theory in a uncertainty context, to minimise the expected travel cost or maximise the expected travel utility.

An optimal strategy is a strategy which allows the traveler to obtain a (long-run) maximum expected travel utility on the O-D pair.
Hyperpath-hs

Nguyen and Pallottino (1988) highlighted the underlying graph structure of Spiess’ basic strategy concept, introducing a **graph-theoretic framework** and the concept of **hyperpath**: an acyclic subnetwork, connecting the origin to the destination and including a sub-set of diversion nodes (and a sub-set of diversion links).

At each diversion node, the choice of the diversion link depends on the **occurrences** of transit services and on the **decision rules**, therefore there are **probabilities** for choosing a link among the alternative diversion links (Khani et al., 2015).
Strategies and diversion rules

Given a line service graph, a travel strategy $S$ is defined through:

- a line hyperpath $HP$,
- an objective function $Of$
- a diversion rule $dr_i$, for each diversion node $i$, which determines the choice of the diversion link, at that node.

A rule could be to board the arriving line if its expected travel time is less than or equal to the sum of the expected waiting time and the expected travel time of the other line.

The strategy $S^*[HP^*, Of, dr]$ which optimizes the objective function is the optimal strategy conditional upon the given diversion rule $dr$ and the given objective function $Of$. 
Optimal travel strategies as solutions of Markov Decision Problems

✓ Path choice in an unreliable service network entails decision making without comprehensive knowledge of all relevant factors and their possible future evolution. Hence the outcomes of any decision depend partly on randomness and partly on the agent’s decisions.

✓ Therefore, a general theoretical framework for optimal strategy-search can be found in stochastic decision theory.

✓ If path choice is considered as decision-making in a Markov decision process (MDPs), the Markov decision problem (MDPm) approach can be used.
Markov decision problems – MDPm 1

A Markov decision process MDPs can be defined by the quintuple

\[(T; SS^t; A_{st}; p^t [j/s, a]; r^t [s, a])\]

✓ \(T\) is a set of stages \(t\) at which the decision maker observes the state of the system and may make decisions; *(the set of times when traveler is at a diversion node and a diversion link has to be chosen)*;

✓ \(SS\) is the state space, where \(SS^t\) refers to the possible system states for a specific time \(t\); *(the set of diversion nodes among which traveler moves)*;

✓ \(A_{s,t}\) is the set of possible actions that can be taken after observing state \(s\) at time \(t\); *(the different diversion link choice sets with a diversion rule)*;
Markov decision problems – MDPm 2

\[(T; SS^t; A_st; p^t [j/s, a]; r^t [s, a])\]

✓ \(p^t [j/s,a]\) are the \textit{transition probabilities}, determining how the system will move to the next state. In particular, \(p^t [j/s,a]\) defines the transition to state \(j\) belonging to \(SS^{t+1}\) at time \(t + 1\); \textit{(the probabilities of going from a diversion node to each of the following diversion nodes, if action “a” is applied)};

✓ \(r^t [s, a]\) is the \textit{reward function}, which determines the consequence for the decision maker’s choice of action \(a\) while in state \(s\). The value of the reward depends on the next state of the system, becoming an \textit{“expected reward”}; \textit{(the expected utility from the diversion node to destination, if action “a” is applied)};

✓ Transition probabilities and expected utility only depend on state \(s\) and chosen action \(a\) at time \(t\), so \textit{Markov property is satisfied}.
Markov decision problems – MDPm 3

✓ An MDPs with a specified optimality criterion (hence forming a sextuple) is called *Markov decision problem MDPm*.

✓ The objective of MDPm is to provide the decision maker with an *optimal policy* $p^*$ that associates to states $S$ actions $A$ optimizing a predefined objective function (in our case an *optimal strategy* $S^*$ which associate to each diversion node a set of diversion links among which to choice with the given diversion rule, to maximise expected utility).
MDPm solution methods

✓ When the transition probabilities $P$ and the reward function $R$ are known, exact methods can be applied, for example Linear Programming (LP) or Dynamic Programming (DP).

✓ The most common DP algorithm to solve MDP is the value iteration (Bellman, 1957).
Optimal strategy search methods 1

Analytical probabilities

In some cases a direct computation of transition probabilities is possible through analytical methods. For example, if it is assumed:

✓ a completely random pattern of arrival at stops for vehicles and a uniform pattern for travellers
✓ a non-congested transit system
✓ known average link travel utilities (times)
✓ travellers boarding the first arriving line of an attractive line set

the probability $p(ln)$ that line $ln$ is the first arriving at the stop, and hence the probability to use line $ln$, can be expressed as function of line frequencies (Spiess and Florian, 1989).

Often the hypotheses required are not congruent with the study cases.
Optimal strategy search methods 2

✓ When the optimal strategy is sought in an attempt in real time to take into account the actual transit service functioning and the available info systems, such as individual predictive info, the probabilities $P$ and rewards $R$ are very difficult to determine and approximate time-consuming MDPm methods have to be applied, such as those based on the reinforcement learning (RL) approach.

✓ For example, Wahba and Shalaby, within an agent-based transit simulation model, apply a Reinforcement Learning-RL method and a Q-learning algorithm, with explorations of available actions and exploitations of experience, to converge to the optimal policy.
OPTIMAL STRATEGY
SEARCH METHODS

INTELLIGENT TRANSIT SYSTEMS

DATA-DRIVEN APPROACH
Optimal strategy search methods 3

*Intelligent transit systems – data-driven approach*

✓ **Intelligent transit networks** allow an efficient forecasting of path attributes, and thus an **exact DP method** can be applied (Nuzzolo and Comi, EWGT2017), using a **data-driven** approach.

✓ If a **time series approach** is used in path attribute forecasting, then:

  ➢ the **expected reward** is the **expected path utility** obtained through the **(data-driven)** expected forecasted path attributes

  ➢ the **transition probabilities** can be computed starting from the distribution probabilities of the **(data-driven)** bus travel time forecasting errors.
The (objective) optimal strategies are also defined as **normative strategies**, because they indicate how travellers should or ought behave to obtain a given objective.
SUBJECTIVE
OPTIMAL STRATEGIES
Subjective Optimal Strategies

Due to cognitive limitations, travellers apply strategies here called *subjective (or descriptive) strategies*, different from the *objective optimal strategies* (Hickman, 2017; Kurauchi et al., 2012; Schmocker et al., 2013; Fonzone et al., 2013).
SUBJECTIVE OPTIMAL TRAVEL STRATEGIES in within-day and day-to-day agent-based transit network simulation
Subjective optimal travel strategies in transit network simulation 1

✓ Due to cognitive limitations, travellers apply subjective optimal strategies, which can differ from the objective optimal ones.

✓ Therefore, in transit network simulation, subjective optimal travel strategies have to be considered, but a consolidated paradigm of subjective travel strategy modelling for different traveller categories, does not exist yet.

✓ Thus, in transit network simulation a unique strategy is assumed valid for all the users and an optimal strategy conditional to a given diversion rule is applied.

✓ Usually, the Spiess and Florian optimal strategy approach is applied, but often the hypotheses required are not congruent with the study cases.
Subjective optimal travel strategies in transit network simulation 2

In run-oriented agent-based meso-simulation of unreliable transit service networks, in order to search subjective optimal strategies, taking into account the specific study case conditions,

- some authors (Wahba) use Reinforcement Learning-RL methods, but this approach requires too long computation time (e.g. several minutes at each simulation step).

- other authors assume diversion rules too complex in relation to travellers’ cognitive capacity (Nuzzolo et al., in DYBUSRT).

- other authors (for example Cats in BUSMEZZO) consider the service stochasticity implicitly in a path choice random utility model.
Subjective optimal travel strategies

*Proposed search method*

*(Traveller-driven approach)*

A new path choice behavioural framework is proposed in Nuzzolo and Comi (submitted to TRB) for agent-based transit simulation, which allows to obtain the *subjective optimal travel strategy* as a solution of a simplified MDPm, with a simplified state-action tree and proxies for transition probabilities and for expected rewards, obtained through a “*traveller-driven*” approach.
Conclusions and further developments

✓ The search of optimal travel strategies on unreliable transit network was analysed as Markov decision problem and two applications for normative strategy search and subjective optimal strategy search were considered.

✓ Several research issues still need to be resolved. These include:
  ➢ developments within theories other than expected utility,
  ➢ in relation to subjective optimal strategy:
    • master line hyperpath choice modeling
    • new models of travel strategy generation for different categories of users
    • introduction of stochastic path choice models which take into account user’s heterogenities and analyst modelling