Publications

This dissertation consists of the present summary and of the following publications:


Author’s contribution

The author contributed to Publications [I]–[V] as follows.

Publication [I]: The paper was initiated by Siikonen and primarily written by Sorsa. Siikonen provided data for the simulation experiment and Ehtamo helped in the mathematical formulations. Sorsa developed the prototype group control and simulation model for the double-deck conventional control and conducted the numerical experiments.

Publication [II]: The paper was initiated by Siikonen and primarily written by Sorsa. Sorsa developed the prototype group control and simulation model for the double-deck destination control and conducted the numerical experiments.

Publication [III]: The paper was initiated and primarily written by Sorsa. Sorsa developed the software for analyzing the double-deck elevator dispatching problem off-line and provided solution candidates for the problem instances to which Ruokokoski evaluated the pareto-optimal solutions.

Publication [IV]: The paper was initiated by Sorsa and primarily written by Ruokokoski. The elevator dispatching problem was jointly formulated by Ruokokoski and Sorsa. Ruokokoski conducted the numerical experiments.

Publication [V]: The paper was initiated by Ehtamo and primarily written by Sorsa. The elevator dispatching problem was jointly formulated by Sorsa and Ehtamo. Sorsa created the risk scenarios and conducted the numerical experiments. Kuusinen and Ruokokoski helped in developing the models to estimate stochastic demands and customers.
Errata (Publication IV)

The algorithm to solve the assignment formulation of the elevator dispatching problem is incorrectly stated as Branch-and-Bound. The algorithm actually was Branch-and-Cut.
Contents

1 Introduction .......................................................... 1
   1.1 Background .................................................. 1
   1.2 Objectives .................................................. 7
   1.3 Research methods .......................................... 8
   1.4 Dissertation structure ..................................... 11

2 Single-deck elevator group control ......................... 11
   2.1 Theoretical foundations .................................. 11
   2.2 Results .................................................... 16

3 Double-deck elevator group control ....................... 19
   3.1 Theoretical foundations .................................. 19
   3.2 Results .................................................... 20

4 Discussion .......................................................... 23
   4.1 Theoretical and practical implications ................. 23
   4.2 Future research directions ............................... 26

References ............................................................ 28
1 Introduction

1.1 Background

Elevator systems for tall buildings

The invention of elevator safety equipment in 1852 by Elisha Otis marked the start of using elevators for passenger transportation, which allowed architects to design higher and higher multi-storey buildings (Otis 1861). Already in the early years of modern elevators, elevator group handling capacity was recognized as a bottleneck for buildings with high traffic demand but limited space for elevators. To meet this challenge, the elevator industry has developed several technologies. One natural solution is to install another elevator car in the same elevator shaft, which resulted in the inventions of a double-deck elevator and a dual elevator system (Gumpel 1916, Sprague 1930). Later, the development of computing hardware enabled the destination control system (DCS) in which passengers register their destination floors already in the lobby using a passenger terminal with a numeric keypad (Schröder 1990).

In tall buildings, elevators are typically gathered into groups which share call-giving devices and operate under the coordination of an elevator group control system (EGCS). The main task of the EGCS is to dispatch the elevators to passenger calls or transportation requests. In the conventional control system (CCS), passengers use up and down call buttons to give a call. Hence, a transportation request consists of a passenger’s origin floor and desired traveling direction. However, the destination floor remains uncertain for the EGCS until the passenger registers it in the elevator. The serving elevator is signalled by hall lanterns and gongs, i.e., the elevators provide bulk service with shared signalization. The moment the EGCS signals the serving elevator defines the assignment policy of the EGCS and how it is allowed to react to new transportation requests. Typical to western manufacturers, the signalling of an approaching elevator is delayed until it starts to decelerate to the call floor which allows the reassignment of the requests until the last moment. However, Japanese manufacturers started to provide immediate signalization as early as in the 1970s (Hirasawa et al. 1978). Such a system does not allow reassignments.

Based on the idea of Port (1961) and promising results of Closs (1970), the destination control system (DCS) was introduced in the 1990s. The *de facto* standard DCS operates with the immediate assignment policy, as it shows the serving elevator on the terminal display immediately after a passenger has registered the call. Hence, the DCS provides
personalized guidance and signalization. Since all passengers are assumed to register their
destination floors, the DCS has accurate information about the passengers. As a result,
the DCS can group passengers with common destination floors in the same elevator, which,
under incoming traffic conditions, reduces the number of stops and elevator roundtrip time,
and increases the handling capacity (Schröder 1990, Koehler and Ottiger 2002, Smith and
Peters 2002, Sorsa et al. 2006). However, the DCS has its limitations during lunch traffic
in office buildings since the traffic mix presents only few chances of taking advantage of
the destination floor information. The situation is further complicated by the immediate
assignment policy.

A double-deck elevator consists of two elevator cars which are stacked one above the
other and fixed in the same car frame. A double-deck elevator can serve two adjacent
landings simultaneously if their distance equals the distance between the decks. Double-deck
elevator groups are usually arranged with a dual ground-floor lobby. These lobby levels are
interconnected by escalators to ease the journey of the passengers. Typically, incoming
passengers destined for odd floors are guided towards the lower lobby level to board the
lower deck and, respectively, those destined for even floors are guided to the upper lobby
level to board the upper deck (Fortune 1996). As a result, a double-deck elevator stops only
on every other floor. In addition, the rate of coincident stops during which both decks allow
passengers to either board or alight is relatively high. Hence, double-deck elevators increase
up-peak handling capacity not only because of the double car capacity, but also due to the
reduced number of stops per roundtrip and simultaneous passenger transfers to and from
both decks (Kavounas 1989, Siikonen 2000). However, during lunch traffic, the inter-floor
passengers between the upper floors easily break the even-odd split of the decks, which is
otherwise an efficient strategy for handling both incoming and outgoing traffic. This leads
to many non-coincidental stops and reduces the efficiency of double-deck elevators.

In the dual elevator system, two elevator cars run in a single shaft independently using
their own hoist mechanisms. Additionally, the dual elevator system benefits from the dual
ground-floor lobby if the served floors are split into two sub-zones. In such an arrangement,
the lower car serves the lower half of the floors and, respectively, the upper car the upper
half. However, the inter-floor traffic between the sub-zones may cause destination floor
combinations that make one car cross the path of another. For this reason, one shaft of the
elevator group can be dedicated to an ordinary single-deck elevator (Thumm 2004).

The first double-deck elevators were installed in the Eiffel Tower as early as in 1889 (Vo-
gel 1961). However, it was not until the beginning of the 1970s when double-deck elevators
reached sufficient maturity for commercial high-rise buildings. During the first wave, approx-
imately 10 buildings were constructed with double-deck elevator groups. On the other hand, the first dual elevator system was installed and put in regular service in 1930 but was later dismantled (Fleming 1931). Neither of these systems succeeded in becoming common technology partly because of the inefficient early CCSs. The DCS turned out to be the necessary technological breakthrough in starting a new boom of double-deck elevators and enabling the dual elevator system in practice (Thumm 2004, Fortune 2005). These technologies do not yet perform optimally during lunch traffic, which consequently gives the motivation to develop optimization models and algorithms for an EGCS further.

**Optimal control of an elevator group**

Traditionally, it has been assumed that passenger arrivals at an elevator lobby follow a Poisson process (Alexandris 1977). Accordingly, passenger arrivals are independent and their inter-arrival times follow an exponential distribution with rate parameter $\lambda$ (Johnson et al. 2005). However, passengers often travel in socially connected batches towards the same destination floors, which can be modelled as a compound Poisson process (Kuusinen et al. 2012). In the compound Poisson process, the batches arrive randomly according to a Poisson process with rate $\lambda/\beta$; the batch sizes are independent and identically distributed random variables with mean $\beta$ (Johnson et al. 2005). Hence, $\lambda$ represents the expected number of passenger arrivals during an observation period, i.e., passenger demand, while $\lambda/\beta$ is the expected number of batch arrivals. If the batch sizes are distributed according to a geometric distribution with parameter $\theta = 1/\beta$, the resulting process is a geometric Poisson process, which is also known as a Pólya-Aeppli process (Galliher et al. 1959).

Modern elevators detect the counts of boarding and alighting passengers by using a load weighing device and door safety equipment (Siikonen 1997). The EGCS can accumulate these passenger counts for predefined time periods on each floor both in the up and down direction, which define the rate parameters $\lambda_{ti}$ for period $t$ and arrival process $i$. A typical period length is as long as 15 minutes to guarantee sufficient arrivals for detecting the Poisson process, although during peak times five minutes could be sufficient. The batch sizes, however, cannot be directly detected from the counts but can be estimated by solving an elevator trip origin-destination (OD)-matrix for each one-directional elevator trip (Kuusinen et al. 2015). The accumulated OD matrices of a particular time period describe the arrival process parameters between all OD floor pairs $(i,j)$, i.e., $\lambda_{tij}$, $\beta_{tij}$, as well as the discrete distributions $Y_{tij}$ of batch sizes, $\Pr(Y_{tij} = k)$. Generally, rate parameters, batch size distributions as well as the traffic mix vary from period to period throughout the day (Strakosh and Caporale 2010,
Siikonen 1997, Peters et al. 2011, Kuusinen et al. 2012, Siikonen et al. 2014). Thus, the arrival processes are non-stationary.

For elevator groups with the CCS, morning up-peak in office buildings is the most difficult traffic condition since the EGCS is typically able to dispatch only one elevator to the only transportation request. Therefore, the EGCS needs to detect the prevailing up-peak condition and automatically dispatch idle elevators back to the entrance floor. Traffic pattern recognition has been solved by using many artificial intelligence methods (e.g. So et al. 1995, Siikonen 1997, Kim et al. 1998, Luo et al. 2005, Cortés et al. 2012). On the other hand, dispatching of idle elevators to the ground floor and letting them depart as soon as possible is not necessarily an optimal strategy. It has been shown that passenger waiting times are minimized by a threshold policy, according to which an elevator is sent off from the ground floor after its load exceeds either a constant or adaptive threshold value (Pepyne and Cassandras 1997, 1998).

In a general traffic situation, the EGCS faces a dynamic on-line control problem with uncertainties. The uncertainties arise from the available information at the time of making dispatching decisions and from the near-future passenger arrivals. Formally, the EGCS controls a set $E$ of elevators with finite capacity, which can be modelled as an asynchronous event-driven stochastic optimal control problem. A typical EGCS operates as a certainty equivalent controller as follows (Bertsekas 2005):

1. at time step $k = 0$, the EGCS measures the system state, replaces the uncertain quantities by their typical values and solves the snapshot elevator dispatching problem (EDP) in an on-line fashion;

2. the solution to the snapshot problem assigns an elevator to each transportation request and defines the optimal control sequences $\bar{\pi}_e = \{\bar{u}_{e,0}, \ldots, \bar{u}_{e,N_e-1}\}$ for each elevator $e \in E$ for stages $k = \{0, \ldots, N_e - 1\}$;

3. the EGCS applies the first control inputs $\bar{u}_{e,0}$ of the sequences and dispatches the elevators to serve the corresponding transportation requests;

4. at time step $k = 1$, the EGCS measures the system state, solves the snapshot EDP anew resulting in the new control sequences $\bar{\pi}_e$, and applies $\bar{u}_{e,0}$.

This process continues until all the elevators have reached their final stages $N_e$ and have become idle. If the system state at time step $k = 1$ has not changed significantly since time
step $k = 0$, the applied control input $\hat{u}_{e,0}$ should equal the corresponding original control input $\bar{u}_{e,1}$. However, if the system state has changed significantly, the solution to the first snapshot EDP may have become suboptimal or even infeasible and, consequently, $\hat{u}_{e,0}$ does not correspond to $\bar{u}_{e,1}$. In such a case, an EGCS operating under the delayed assignment policy can cope with the new situation by re-optimizing all the assignments and applying the new sequences $\hat{\pi}_e$. On the other hand, under the immediate assignment policy, the new sequences may divert some of the elevators from their current assignments. This may significantly delay their service since the original elevator assignments cannot be changed. Therefore, the uncertainties should be considered in the EDP especially in the case of the immediate assignment policy.

The elevator dispatching problem (EDP) is a snapshot optimization problem that models the dispatching decisions and elevator routes for a finite planning horizon. The EDP is conceptually similar to the pickup and delivery problem (PDP) of multiple vehicles in which a fleet of vehicles transports customers from their origins to their destinations (Ruland and Rodin 1997, Savelsbergh and Sol 1995). This dissertation considers several variants of the EDP that may differ with respect to the handling of various constraints. The variants, however, share the same objective functions that allow their comparison numerically. Common to all the variants under consideration, the main decision variables define a binary choice of whether elevator $e \in E$ is assigned to transportation request $i \in V$, or, in the case of double-deck elevators, whether elevator $e \in E$ and deck $d \in \{1,2\}$ are assigned to transportation request $i \in V$. In particular, the bilevel formulation of the EDP follows the distributed structure of a typical real-world elevator system: it separates the problem to an upper-level problem of assigning elevators to transportation requests optimally according to a suitable objective function, and a set of lower-level problems for finding minimum-time routes for the elevators. Furthermore, the bilevel approach is similar to the well-known decomposition of a closely related optimization problem, the vehicle routing problem (Fisher and Jaikumar 1981, Bertsekas 2005).

The bilevel EDP can be formulated as a problem of minimizing a passenger service quality criterion. Here, the details of the lower level elevator routing problems (ERPs) are omitted. Let $V$ denote the set of transportation requests which consists of origin-destination (OD) pairs $(i^+,i^-)$, $i^+ \in V^+$, $i^- \in V^-$. The EDP assigns an elevator to each transportation request by binary assignment variables $x_{ei}$ for all $e \in E$ and $i \in V$. Furthermore, each transportation request is associated with demand $D_i$ (number of passengers) and time since
registering the request $\gamma_i$. With this notation, the upper-level problem becomes,

$$\min_{x_e} \sum_{e \in E} f_e (x_e, \bar{\pi}_e (x_e)) = \sum_{e \in E} \sum_{i \in V'} D_i (\gamma_i + s_{ei})$$

subject to

$$\sum_{e \in E} x_{ei} = 1, \forall i \in V,$$

$$x_{ei} = 1, \forall i \in \bar{V}, e \in E,$$

$$x_{ei} \in \{0, 1\}, \forall i \in V, e \in E,$$

$$\bar{\pi}_e (x_e) \in \arg \min_{\pi_e (x_e)} g_e (x_e, \pi_e (x_e)),$$

where $f_e (\cdot)$ denotes the upper-level and $g_e (\cdot)$ the lower-level objective function for elevator $e$, respectively. The control sequence $\bar{\pi}_e (x_e)$ is the solution to the lower-level problem of elevator $e$ with the given assignment vector $x_e = (x_{ei})$, which also defines the elevator arrival times $s_{ei}$ to transportation request $i \in V$ and the number of passengers $q_{ei}$ inside the elevator after serving it. The arrival times $s_{ei}$ are accumulated by non-linear elevator flight times $t_{eji}$ between requests $j, i \in V$, which are calculated accurately using the nominal speed, acceleration and jerk of the elevator (Motz 1976, Roschier and Kaakinen 1979). In addition to the precedence and capacity constraints, a feasible elevator route $\pi_e (x_e)$ must satisfy the basic rules of elevator operation, first stated by Closs (1970):

1. A car may not stop at a floor where no passenger enters or leaves it.
2. A car may not pass a floor at which a passenger wishes to leave it.
3. A passenger may not enter a car carrying passengers and traveling in the reverse direction to his required direction of travel.
4. A car may not reverse its direction of travel while carrying passengers.

In the objective function (1), the set $V'$ represents either pickup or delivery requests depending on whether the objective is to minimize passenger waiting or journey times. In the CCS, the demand $D_i$ is estimated according to the expectation of the Poisson process (Siikonen 1997, Tyni and Ylinen 2001),

$$D_i = 1 + \lambda_i (\gamma_i + s_{ei}).$$
In the DCS, the demand equals one since each passenger is assumed to register the transportation request. Constraint (2) ensures that each transportation request is assigned to exactly one elevator. The system constraint (3) reflects the assignment policy through the subset of transportation requests having an unchangeable elevator assignment, $\bar{V} \subset V$.

Owing to integer variables, uncertainties and non-linear vehicle dynamics, the EDP is a rather challenging optimization problem to formulate in the standard form. The advantage of the standard form lies in the fact that it could be solved with well-known exact optimization algorithms such as branch-and-bound (Bertsimas and Weismantel 2005). However, with exact algorithms, the computation times become excessive in large-scale instances since the EDP is an NP-hard problem. Due to these reasons, the EDP has not been studied much as a formal optimization problem, although many kinds of mathematical methods have been applied to optimize elevator group performance (Fernández and Cortés 2015). On the other hand, metaheuristics such as genetic algorithm, tabu search, particle swarm optimization and ant colony optimization are heuristic optimization algorithms which explore the solution space on the basis of heuristic rules and are well suited for solving combinatorial optimization problems (Goldberg 1989, Deb 2001, Glover 1989, Kennedy and Eberhart 1995, Dorigo et al. 1996). These algorithms provide high-quality solutions fast, but they have no guarantee of finding the global optimum. The development of computing hardware has enabled the implementation of ever more complex methods in the EGCS. Optimization did not become a feasible alternative until the late 1990s along with the computational power of industrial PCs and the development of system software. These factors enabled the implementation of a tailored real-time genetic algorithm for a real-world elevator product that optimizes dispatching decisions on-line (Tyni and Ylinen 2001).

### 1.2 Objectives

In general, the main objectives of this dissertation are, how to model elevator dispatching as an optimization problem, how to obtain optimal or high-quality solutions to problem instances, and how the optimization affects elevator group performance and passenger service quality. New mathematical optimization models and algorithms are presented for different elevator products as shown in Table 1.

The double-deck elevator dispatching problem (DD-EDP) is formulated for the CCS and DCS in Publications [I] and [II], respectively. A prototype EGCS solves the DD-EDP by a real-time genetic algorithm. Elevator traffic is simulated for different traffic conditions.
Table 1: Elevator products and control systems studied in the publications of this dissertation.

<table>
<thead>
<tr>
<th>Elevators</th>
<th>Conventional control</th>
<th>Destination control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-deck</td>
<td>–</td>
<td>[IV], [V]</td>
</tr>
<tr>
<td>Double-deck</td>
<td>[I]</td>
<td>[II], [III], [VI]</td>
</tr>
</tbody>
</table>

to study passenger service quality and elevator group handling capacity, i.e., the maximum number of passengers that an elevator group can transport in five minutes.

In Publications [III], [IV] and [VI], the DCS with the delayed assignment policy is studied by analyzing separate EDP instances. The objectives of these publications are to estimate the improvement potential of the delayed assignment policy in passenger service quality and/or to develop efficient solution algorithms for the EGCS. In particular, the DD-EDP is considered as a multi-objective optimization problem in Publication [III] in which passenger waiting and transit times are conflicting objectives. In Publication [IV] the EDP is formulated as a mixed integer linear programming problem and solved by an exact algorithm. The DD-EDP is formulated anew as a bilevel mixed integer linear programming problem in Publication [VI] for which new solution algorithms are developed to be applied to the DCS with delayed assignments.

In Publication [V], necessary modelling techniques are developed to estimate passenger-related uncertainties for the robust EDP with the DCS. To achieve robustness, the EDP is formulated as a stochastic bilevel optimal control problem in which the uncertainties are modelled explicitly in multiple risk scenarios. The objective is to compare the forecasting accuracy of different estimation methods and recommend the most suitable approach for the on-line optimization of the EGCS.

1.3 Research methods

Elevator traffic simulation plays a key role in the development of elevator dispatching since it enables the rapid comparison of control systems under a wide variety of traffic conditions (Siikonen 1993, Peters 1998, Siikonen et al. 2001, Cortés et al. 2006). The results of this dissertation are based on case study simulations for which the key parameters are summarized in Table 2. The simulator generates virtual passenger agents with random arrival times according to either the Poisson or the geometric Poisson process (Sorsa et al. 2013). Furthermore, each agent is associated with the origin and destination floor randomly according to the traffic mix. The simulator models the journeys of the agents in the building and the use of the transports. In the simulation, the EGCS responds to the transportation requests of
the agents by solving the EDP on-line. As a result of these processes, the simulation events define a multitude of performance measures related either to the elevators or the passengers.

Table 2: Traffic patterns (UP = up-peak, DP = down-peak, L = lunch), elevator types (SD = single-deck, DD = double-deck), control systems (CCS = conventional control, DCS = destination control) and key elevator group parameters used in the simulations of this dissertation.

<table>
<thead>
<tr>
<th>Publication</th>
<th>Traffic pattern</th>
<th>Elevator type</th>
<th>Control system</th>
<th>Total population (persons)</th>
<th>Population distribution</th>
<th>Number of populated floors</th>
<th>Number of elevators</th>
<th>Rated speed (m/s)</th>
<th>Passenger capacity (persons)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>L, DP</td>
<td>DD</td>
<td>CCS</td>
<td>3540</td>
<td>Uneven</td>
<td>29</td>
<td>8</td>
<td>5.0</td>
<td>21</td>
</tr>
<tr>
<td>II</td>
<td>UP</td>
<td>DD</td>
<td>DCS</td>
<td>1800</td>
<td>Even</td>
<td>18</td>
<td>6</td>
<td>3.5</td>
<td>21</td>
</tr>
<tr>
<td>III</td>
<td>L</td>
<td>DD</td>
<td>DCS</td>
<td>1585</td>
<td>Uneven</td>
<td>17</td>
<td>5</td>
<td>4.0</td>
<td>21</td>
</tr>
<tr>
<td>IV</td>
<td>UP, DP, L</td>
<td>SD</td>
<td>DCS</td>
<td>450...1100</td>
<td>Even</td>
<td>9...22</td>
<td>1..6</td>
<td>3.0</td>
<td>21</td>
</tr>
<tr>
<td>V</td>
<td>DP</td>
<td>SD</td>
<td>DCS</td>
<td>N/A</td>
<td>Even</td>
<td>10</td>
<td>1</td>
<td>2.5</td>
<td>4.0</td>
</tr>
<tr>
<td>VI</td>
<td>L</td>
<td>DD</td>
<td>DCS</td>
<td>1800</td>
<td>Even</td>
<td>18</td>
<td>5</td>
<td>4.0</td>
<td>21</td>
</tr>
</tbody>
</table>

Usually, the differences between control systems are studied with serial simulations for stereotypical traffic patterns (Hakonen and Siikonen 2009). Such traffic patterns include up-peak traffic consisting of 100% of incoming traffic, down-peak traffic (100% of outgoing traffic) and lunch traffic (40% incoming, 40% outgoing, and 20% inter-floor traffic) (Siikonen 1993). The serial simulation itself contains several independent simulations, each of which keeps passenger demand constant for the duration of one simulation. Passenger demand is increased by a specified step from one simulation to another. Usually, the performance measures are reported as averages for each simulation but excluding the first five minutes of each simulation step so that the results represent steady-state simulation. Simulation time is critical to the validity of the analysis and should be at least 30 minutes for each arrival rate to guarantee sufficient statistical accuracy.

Simulation, however, is not the best tool for comparing optimization models and algorithms, since the performance measures of the simulation are an accumulated account of solving a large number of EDP instances. The performance of the solution algorithms is related to a particular instance of the EDP. Therefore, custom-built software is used to analyze the EDP instances off-line for which instance data has been produced by traffic simulation. The off-line analysis enables the comparison of different optimization models and algorithms with respect to, e.g., objective values and computation times. Simulation performance measures and optimization objectives use the same terminology interchangeably, although the definitions of the parameters differ slightly. Simulation results are always averages of the realizations during the simulations and are based on the exact timing of simulation events. On
the other hand, the values of the optimization objectives represent predictions of a snapshot instance.

In Publication [I], the Advanced Lift Traffic Simulator (ALTS) is used to simulate a group of double-deck elevators (Siikonen 1993). In the ALTS, a double-deck elevator group is modelled with two entrance floors between which the entrance bias is even, i.e., both entrance floors attract 50% of incoming traffic. The next-generation simulator, the Building Traffic Simulator (BTS) is utilized in Publications [II]–[VI] (Siikonen et al. 2001). The BTS puts a building and its transports in 3D-geometry as well as simulates multiple transports at the same time. As shown in Figure 1, the BTS simulates double-deck elevator groups as is typical in real buildings. All incoming passengers arrive at the only entrance floor. Then, they either walk to the lower entrance level or take an escalator to the upper entrance level depending on their destination floor. Furthermore, the advanced passenger models of the BTS allow the modelling of the destination control as in real installations where the walking times from the call-giving devices to each elevator are configured in the EGCS. The results in Publications [I] and [II] are based on serial simulations of different traffic patterns. The reported results include average passenger waiting time, journey time and time to destination as well as average car load factor (Barney 2005, Hakonen and Siikonen 2009). The average car load factor is also used to determine the simulated handling capacity for a particular traffic pattern and control system.

![Figure 1: A screenshot showing the simulation of two double-deck elevator groups with the Building Traffic Simulator, where passengers destined for even floors use the escalator to reach the upper lobby level.](image)

In Publications [III]–[VI], individual problem instances are solved off-line. The branch-and-cut algorithm of IBM ILOG CPLEX Optimization Studio 12.5.1 is used in Publication
[IV] to study both the computation times and objective values of the instances. In Publications [III] and [VI], custom-built software reads instance data from log-files stored during the simulations and then solves the instances using either the genetic algorithm or a branch-and-bound algorithm. In Publication [V], EDP instances for one elevator are created to compare the elevator routes with different estimation methods of the uncertainties.

1.4 Dissertation structure

The rest of this dissertation is organized as follows. The key results and main contributions of this dissertation are presented in Chapters 2 and 3 for single- and double-deck elevators, respectively. In Chapter 4, implications of the key results are discussed and directions for future research are proposed.

2 Single-deck elevator group control

2.1 Theoretical foundations

Optimization models and numerical algorithms for EGCSs have been researched rather widely since the early approaches using dynamic programming (Closs 1970, Levy et al. 1977). The researched models and methods have been bound to available computing technologies: first relays, integrated circuits in the 1970s, microprocessors in the 1980s, and industrial PCs in the late 1990s. Traditionally, due to the limited computing power, elevator dispatching has dealt with simple questions such as to which transportation requests the elevators are dispatched next. This dissertation, however, departs from this tradition by formulating the EDP in such a way that it builds complete elevator routes through all the transportation requests in a general traffic situation.

Many EDP studies define elevator routes for the given assignments using the collective control principle according to which an elevator serves the nearest pickup or delivery request in front of it (Tyni and Ylinen 2001, 2006, Koehler and Ottiger 2002, Luh et al. 2008). The advantage of the collective control lies in the fact that it undoubtedly satisfies the basic rules of elevator operation and, hence, provides feasible elevator routes by only considering simple rules. Further, the resulting routes have been found nearly optimal, within 5% of the optimal policy, with respect to route duration and passenger journey time (Closs 1970). Even in the
case of one elevator operating under the DCS, passenger journey times can be improved compared to the collective control and especially so if an elevator is allowed to reverse its travelling direction while carrying passengers (Tanaka et al. 2005a). Thus, the optimization of an elevator route may become a rather challenging task. For a practical application, the feasibility of a route with respect to the basic rules and a fast solution algorithm are more important than optimality.

The EDP can be approached methodologically in many ways:

1. as a bilevel problem in which the upper-level or leader’s problem contains lower-level or followers’ problems as constraints (Publications [V] and [VI]; Luh et al. (2008); for a general textbook on the topic, see Shimizu et al. (1997));

2. as a dynamic problem in which new transportation requests may emerge while the elevators are executing their current routes (Publications [I], [II] and [V]; Closs (1970); Levy et al. (1977); see Psaraftis (1988) on the dynamic vehicle routing problem and Bertsekas (2005) on the optimal control theory);

3. as a stochastic problem in which some problem variables are uncertain (Publication [V]; Nikovski and Brand (2003); Brand and Nikovski (2004); Gendreau et al. (1996) give an excellent review of stochastic vehicle routing problems);

4. as a multi-objective problem with two or more possibly conflicting objective functions (Publication [III]; Tyni and Ylinen (2006); for a general textbook on multi-objective optimization problems and their solution by evolutionary algorithms, see Deb (2001));

5. as a mixed integer programming problem in which some variables can obtain only integer values (Publications [IV] and [VI]; Friese and Rambau (2006); Hiller et al. (2014); Ruokokoski et al. (2015); for a general text on the topic, see Bertsimas and Weismantel (2005)).

Bilevel programming concerns problems that have a decomposable structure. Bilevel problems are usually modelled either as an optimal value function or a Lagrange multiplier coordination (Shimizu et al. 1997). The decomposition of the vehicle routing problem into the generalized assignment problem and a set of travelling salesman problems for each vehicle has been known for a long time but its bilevel form has not gained much attention (Fisher and Jaikumar 1981, Bertsekas 2005, Marinakis et al. 2007). Stochastic bilevel models were first introduced for pricing and knapsack problems but have recently also been considered
for the vehicle routing problem (Luh et al. 1987, Patriksson and Wynter 1999, Audestad et al. 2006, Alizadeh et al. 2013, Özaltın et al. 2010, Kosuch et al. 2012, Ma and Xu 2014). The bilevel structure has also been adopted in many modelling approaches of the EDP starting from Hirasawa et al. (1978). Luh et al. (2008) formulated the EDP as a bilevel problem by using the Lagrange multiplier coordination approach. Differing from the classical bilevel forms, the EDP can also coordinate elevator routes by an auction process (Koehler and Ottiger 2002). The decomposition of the EDP has also been used in other studies although not formally presented as a bilevel problem (Tyni and Ylinen 2001, Friese and Rambau 2006, Hiller et al. 2014).

The assignment policy defines how an EGCS is allowed to react to new transportation requests. The delayed assignment policy, which has been a popular choice for the CCS, allows the reassignment of a transportation request to another elevator if something unexpected happens for the currently assigned elevator (e.g. Siikonen 1997). Recently, it has also been proposed for the DCS, although such a system has not yet been realized in practice (Hiller et al. 2014). On the other hand, the immediate assignment policy does not allow the reassignment, which increases the risk of the planned routes becoming suboptimal or infeasible. Hence, the EGCS needs to predict the uncertain future somehow. For that purpose, methods based on artificial intelligence were developed already for the early CCS with the immediate assignment policy (e.g. Hirasawa et al. 1978, Markon et al. 1995, Imasaki et al. 1995, Sasaki et al. 1996). Also the current DCS products operate with the immediate assignment policy and are vulnerable to passenger-related uncertainties although the more accurate information about passengers is sufficient for most traffic conditions (Schröder 1990, Koehler and Ottiger 2002, Smith and Peters 2002, Sorsa et al. 2006). Hence, advance information about or estimation of future passengers improves passenger service quality with the immediate assignment policy (Hirasawa et al. 1978, Nikovski and Brand 2003, Brand and Nikovski 2004, Luh et al. 2008).

Stochastic programming aims at modelling the uncertainties of an optimization problem proactively in the first-stage decision, which has to be taken without knowing the uncertain quantities exactly (Kall 1982). Then, the second (recourse) stage makes the required corrections if the revealed uncertainties violate some constraint. In stochastic routing problems, the demands or the entire set of transportation requests may be uncertain, which are called stochastic demands and customers, respectively (Gendreau et al. 1996). The first-stage decision can be modelled either as a chance constrained problem in which the probability of violating a constraint during the second stage is small, or as a stochastic program with recourse in which the expected cost of the recourse stage is included in the objective function.
Another approach to model stochastic routing problems is called a priori optimization (Bertsimas et al. 1990, Bertsimas 1992). The first stage designs a priori routes for the vehicles to visit all the transportation requests by minimizing the expected length of the route. During the second stage, the vehicles visit either all the possible requests or only those present.

Stochastic programming, however, considers only the expected objective value but not its variability due to uncertainties and, hence, it does not measure the risk involved in the decision. Robust optimization considers explicitly those scenarios in which the uncertain quantities take realizations from bounded sets around their nominal or expected values. The resulting solutions are robust in the sense that they should not be affected much by the realizations of the uncertainties (Ben-Tal et al. 2009). Many of the robust optimization approaches aim at minimizing the worst case objective value or satisfying the constraints within the uncertainty set (e.g. Soyster 1973, Bertsimas and Sim 2004). In complex optimization problems, uncertainties arise from multiple sources. As the number of sources of uncertainty becomes high, it is unlikely that each of them attains its worst-case value (Bertsimas and Thiele 2006). Therefore, worst-case approaches may lead to robust but costly solutions. Instead of the mean or min-max objective function, the mean-variance model or concave utility functions should be used since they also capture the risk levels of the decisions (Markowitz 1952, Von Neumann and Morgenstern 1953, Mulvey et al. 1995).

Many real-world optimization problems are inherently multi-objective. Solutions to such problems should perform well with respect to two or more conflicting objectives. Generally, multi-objective optimization problems do not have a unique optimal solution but a set of Pareto-optimal solutions. The globally Pareto-optimal set represents all non-dominated solutions among the set of all feasible solutions (Deb 2001). Consequently, the Pareto-optimal solutions define the trade-offs between the objectives, since it is not possible to reduce one objective without increasing another. If the Pareto-optimal frontier is convex, then the multi-objective problem can be transformed into a single-objective linear scalarization of the objectives that has a unique optimum.

In the EDP, passenger service quality and elevator energy consumption are conflicting objectives. They can be combined into the objective function as a weighted sum (Tyni and Ylinen 2006, Kobori et al. 2010). The weights, however, should be adapted to the prevailing traffic conditions throughout the day. This challenge is further complicated by the fact that elevator energy consumption depends on elevator group handling capacity, daily traffic patterns and elevator products (Siikonen et al. 2010, Segercrantz 2010, Siikonen 2012). A PI-controller can adjust the weights by comparing measured call times as an indicator of the
service quality to a preset target value (Tyni and Ylinen 2006). If the measured values are greater than the target, the weight of passenger service quality will be increased, and, in the reverse case, the weight of the energy consumption will be increased. As a result, the EGCS is capable of maintaining passenger service quality at the target level and reducing elevator energy consumption when passenger demand is below the handling capacity (So et al. 2005, Ylinen 2008). As the demand approaches or even exceeds the handling capacity, the weight of passenger service quality is increased and elevator routes start to resemble the ones of a single-objective EDP that optimizes only the service quality.

Integer optimization problems such as the vehicle routing problem (VRP), the travelling salesman problem (TSP) and the generalized assignment problem (GAP) belong to the class of NP-hard problems for which a polynomial-time exact solution algorithm does not exist. In other words, the number of iterations and the computation time of the algorithms increase exponentially with respect to the number of modelled entities. The branch-and-bound (B&B) was the first tree-search algorithm to solve the TSP (Little et al. 1963). The algorithm enumerates the solution space gradually by a branching procedure. For each branch, it keeps track of the lower bound value for the objective function, which directs the search to promising sub-problems. Furthermore, once the algorithm has found the first feasible solution that defines the upper bound of the objective value, unprocessed sub-problems are discarded if their lower bounds are inferior to the upper bound. Then, the remaining unprocessed sub-problems are processed to find an improving feasible solution. Similar to the B&B, the branch-and-cut (B&C) forms polyhedral cutting planes to gradually construct facets of a polytope containing all the feasible solutions, while branch-and-price (B&P) solves a separate pricing problem to choose the branching variable (Padberg and Rinaldi 1991, Savelsbergh 1997). Recently, exact methods have also been applied to the EDP. The EDPs arising from pallet transportation in a distribution centre as well as from passenger transportation were formulated as set partitioning models and solved by the B&P (Friese and Rambau 2006, Hiller et al. 2014). Ruokokoski et al. (2015) formulated the EDP as a complete mixed integer programming problem and solved it by the B&C. Tanaka et al. (2005b) used B&B to solve the route of an elevator.

Since large-scale instances of integer programming problems take an exponentially growing solution time, metaheuristics provide an alternative solution strategy with the promise of terminating fast but without any guarantee of global optimality. Usually, they mimic natural phenomena which exhibit optimization-like processes. Further, metaheuristics include randomized mechanisms to explore the solution space and to avoid convergence to local optima. Genetic algorithms, and more generally evolutionary algorithms, simulate evolution
computationally in which the fittest individuals survive. A genetic algorithm creates new generations of individuals by crossover and mutation and terminates with the final population whose best individual determines the solution of the problem instance (e.g. Goldberg 1989). Genetic algorithms have been applied successfully to the VRP, the GAP and the EDP (Chu and Beasley 1997, Baker and Ayechew 2003, Tyni and Ylinen 2001, Cortés et al. 2004).

Even though a genetic algorithm in itself is fast compared to an exact algorithm, real-time optimization of an elevator group control system sets strict requirements for the computation time: 500 milliseconds for the delayed assignment and 50 milliseconds for the immediate assignment policy. The genetic algorithm implemented for a real-world elevator product was accompanied with elitism, decision recycling and the GeneBank, which speed up the convergence, stabilize decisions of subsequent problem instances and store the objective values in a cache memory for fast retrieval if the same candidate solution is encountered again during the algorithm execution (Tyni and Ylinen 1999, 2001). These techniques are also used in the solution algorithms of this dissertation where applicable. Other metaheuristics, such as the ant colony algorithm, tabu search, and particle swarm optimization, have also been applied to the EDP (Liu and Liu 2007, Bolat and Cortés 2011, Bolat et al. 2013). The performance of these algorithms could also be studied for the models presented in this dissertation but the real-time performance would be laborious to achieve.

A bilevel programming problem with linear objective functions and constraints becomes a nonconvex problem which may have multiple local optima (Shimizu et al. 1997). When modelling bilevel problems with value functions, they are typically solved by gradient or subgradient methods but, when modelling them as the Lagrange multiplier coordination, various secant (Broyden) methods are used to obtain the solution. On the other hand, the structure of the genetic algorithm applies well to combinatorial bilevel problems where the genetic algorithm guides the search process of the upper level problem but leaves flexibility of solving the lower level problem by any known method, i.e., by an exact or heuristic algorithm. In a co-evolutionary approach, both the upper and the lower level problem are solved by a genetic algorithm but with separately evolving populations that may exchange information between the generations (Oduguwa and Roy 2002, Deb and Sinha 2010).

2.2 Results

In Publication [IV], the EDP is formulated as a deterministic mixed integer linear programming problem, which can be solved by common optimization libraries. Such a mathematical
formulation has been known for the closely related vehicle routing, pickup and delivery as well as dial-a-ride problem for a long time but not for the EDP (Dantzig and Ramser 1959, Dumas et al. 1991, Psaraftis 1980). One challenge in formulating the EDP is its constraints arising from a generally accepted elevator behaviour "because these constraints are applied to the way in which states may be reached rather than to specific states" (Closs 1970). However, the main challenge lies in the dynamic and stochastic nature of the EDP as evidenced in Publication [V]: the objective function and/or constraints of the problem may be non-linear or they cannot be expressed by closed form equations. Such problems cannot be solved by the well-known exact algorithms but require tailored heuristic algorithms.

The assignment formulation of the EDP introduced in Publication [IV] is an alternative to the routing formulation given in Ruokokoski et al. (2015). The assignment formulation assigns transportation requests to the elevators through main decision variables \(x_{ei}\) but assumes that elevator routes follow the collective control principle and are restricted to only one roundtrip. The routing formulation allows more complex routes that may span multiple roundtrips. These mathematical EDP formulations allow the use of exact solution algorithms in common optimization libraries. The assignment formulation is solved by the CPLEX default branch-and-cut algorithm but the routing formulation by adding some valid inequalities to the problem during the solution process (Ruokokoski et al. 2015).

To compare these EDP formulations, some benchmark instances are first defined in Publication [IV] for up-peak, down-peak and lunch traffic arising from traffic simulations with the DCS. These instances contain only one non-assigned transportation request, which is typical for the DCS with the immediate assignment policy, while all the other requests already have a fixed assignment. The system constraints related to the fixed requests are then relaxed gradually to find out the limiting number of non-assigned requests that can be solved in a reasonable time. These instances correspond to the DCS with the delayed assignment policy. When minimizing the average waiting time, the optimal solutions of the assignment formulation equal the ones of the routing formulation. On the other hand, when minimizing the average passenger journey time, the routing formulation provides better objective value than the assignment formulation in down-peak and lunch traffic instances. Hence, the optimal solution of the assignment formulation is not always optimal for the routing formulation. The results also show that the B&C is able to find the solution for the assignment formulation of the EDP with up to 10 non-assigned transportation requests, which is not sufficiently large to allow the use of exact algorithms in the EGCS with the delayed assignment policy. The solutions provided by B&C could, in principle, be compared with the solutions of the genetic algorithm to gain insight into the solution quality. However, practical problems of
an EGCS with the delayed assignment policy may have much more non-assigned requests than the B&C is capable of solving within a reasonable time. Therefore, the quality of the solutions of the genetic algorithm cannot be easily studied.

The EDP of Publication [IV] is a deterministic optimization problem which uses only the information of the current transportation requests. This approach is valid most of the time since all passengers are assumed to register their transportation requests in the DCS. However, the immediate assignment policy makes the DCS vulnerable to near-future passenger arrivals and unexpected passenger behaviour. For example, passengers traveling in a batch may register only one transportation request or one passenger may try to abuse the system by registering multiple requests. Figure 2 shows the arising challenge with an arrival process example where two batches consisting of one and four passengers have registered two transportation requests within a short period. It is easy to spot the bias of the deterministic EDP in this case: the two transportation requests are interpreted as two passengers while the actual number of waiting passengers is five. Thus, the deterministic EDP underestimates the number of waiting passengers and, as a result, also the number of passengers in the elevator after serving a request, $q_{el}$. This, on the other hand, may lead to such a realized elevator route that violates the capacity constraint although the originally planned route is feasible. Therefore, in Publication [V], the EDP is formulated as a stochastic bilevel optimal control problem and methods to estimate stochastic demand and customers in the EDP are developed.

![Graph](figure2.png)

Figure 2: Realized batch arrivals in a traffic simulation and expected demand $E[D_i]$ as estimated by the EDP by assuming the Poisson process $X(t)$ and the geometric Poisson process $Z(t)$.

The results of Publication [V] show that the EDP variables are greatly underestimated with respect to the realized values if the uncertainties are not estimated at all, i.e., passenger
demands are assumed deterministic and static. Under heavy traffic conditions, an EGCS using such an EDP has a risk of prematurely overcrowding an elevator or substantially delaying the service of an existing transportation request. To minimize these adverse effects, the EDP should explicitly model the uncertainties. However, if the uncertainties are only modelled by their expected values, the EDP variables exceed the realizations with a large margin and define an overly pessimistic solution. On the other hand, robust optimization uses uncertainty sets within which the uncertainties are allowed to vary in different scenarios. The uncertainty sets are usually defined symmetrically around mean values. In contrast, in Publication [V], risk scenarios are defined using specified cumulative probabilities, which allow an exact parametrization of the unsymmetrical Poisson distributions. An analysis of three EDP instances shows that the estimation error among the scenarios is the lowest when modelling the uncertain passenger arrivals by a compound Poisson process. This result is of particular importance: if the passengers arrive at elevator lobbies in batches, the scenarios based on the individual arrivals, i.e., an ordinary Poisson process, exhibit a large estimation error. Hence, the robust EDP should model the uncertain passenger arrivals by using the compound Poisson process, whose parameters, i.e., arrival rate and batch size distribution, the EGCS needs to predict from historical data. When the scenarios are explicitly defined for the EDP, the EGCS becomes robust with respect to the near-future passenger arrivals.

3 Double-deck elevator group control

3.1 Theoretical foundations

The literature covering the dispatching of double-deck elevators is rare. Patents are the best source in finding information about early approaches for the CCS. For example, so called trailing deck principle prefers the deck lagging behind with respect to elevator traveling direction (Nowak and Luce 1986). This principle tends to fill the decks unevenly, i.e., the trailing deck may become full even though the leading deck still has unused capacity. In such a case, the elevator must bypass remaining pickup requests and empty the full deck first before picking up new passengers. The BestDeck method improves the use of the capacity by assigning the requests for the deck with smaller load and by also considering the estimated passenger journey times (Siikonen 1998, 2001). More recently, Hirasawa et al. (2008) proposed a rule-based double-deck DCS in which the EGCS optimizes the rules by genetic network programming and by evaluating multiple objectives such as elevator arrival
time to a request, difference of the number of passengers between the decks and coincidences of the requests.

3.2 Results

The dispatching of double-deck elevators needs to consider additional decision variables that assign a serving deck to each transportation request. In addition, the model must correctly take into account coincident stops during which an elevator simultaneously serves transportation requests on two adjacent floors. The model can be built in two different ways by using the bilevel approach. In Publications [I]–[III], based on the original ideas of Tyni and Ylinen (2001), the upper-level decision variables $x_{edi}$ combine the elevator and deck assignment of transportation request $i \in V$. This approach is called the deck assignment model (DAM) since the main decision variables assign the serving deck directly to each transportation request. For the given elevator and deck assignments, a lower-level problem for each elevator defines the serving order of the requests.

Although straightforward to implement, the DAM has some drawbacks. First, the assignment variables combine the elevator and deck assignments, which grows the solution space in a combinatorial manner compared to single-deck elevators. This is likely to slow down the convergence of the solution algorithm and increase the computational effort. Second, the efficiency of the elevator routes depends on the upper-level objective function, i.e., there are no control mechanisms to ensure efficient routes. Based on these observations, the elevator assignment model (EAM) is introduced in Publication [VI]. The EAM assigns only an elevator to each transportation request by the upper-level variables $x_{ei}$. A lower-level problem defines the deck assignments and serving order of the requests that are assigned to a particular elevator. The lower-level problem is formulated as a two-vehicle pickup and delivery problem to minimize the route duration, which, consequently, makes the elevator route independent of the upper-level objective function.

The modelling approach of the double-deck elevator dispatching problem affects, naturally, the structure and the encoding of the genetic algorithm, which is used to solve the problem in Publications [I]–[III] and [VI]. However, the genetic algorithm for solving both double-deck models need not be modified from the algorithm used for single-deck elevators. In the genetic algorithm, a candidate solution is defined by a chromosome that contains a gene for each transportation request. In the case of the DAM, the value of a particular gene defines a car index, which is a unique mapping to elevator and deck indexes. Thus, the gene
value uniquely determines the decision variables \( x_{ed_i} \) for some transportation request \( i \in V \). Then, the serving order of the requests is solved for each elevator using the collective control principle with the fixed deck assignments. In the EAM, on the other hand, the gene values correspond to the serving elevators, i.e., the decision variables \( x_{ei} \). The EAM allows further decomposition of the lower-level problem into deck assignments and ordering of the transportation requests. The deck assignments can be solved either by minimizing the number of stops or by heuristic rules after which the service order is determined by the collective control principle.

As shown in Publication [VI], the EAM along with the new structure of the genetic algorithm decreases the computation times and enables the implementation of the delayed assignment policies for the DCS in practice. The EAM also shortens elevator routes by one or two stops compared to the DAM in some of the benchmark instances, which is due to the explicit minimization of route lengths. Furthermore, the heuristic rules can be used to solve the deck assignments since they provide nearly optimal routes. This is of particular importance for keeping the computation times short for real-time optimization of an EGCS.

In Publication [I], the double-deck CCS is studied. It is based on the delayed assignment policy to cope with the uncertainties. The stochastic demand is estimated by equation (6). The optimization task is to minimize either the total call time or estimated passenger journey time. Simulation results for lunch and down-peak traffic show that the optimization objectives affect passenger service quality as expected. The average passenger waiting time is usually shorter when minimizing call times than when minimizing journey times. The opposite is true for the average passenger journey time. In addition, the minimization of journey times increases the handling capacity of the elevator group. The optimization-based methods outperform the rule-based BestDeck. In Publication [II], all the four elevator products depicted in Table 1 are analyzed under incoming traffic conditions by applying traffic simulation and up-peak equations (Roschier and Kaakinen 1979, Sorsa et al. 2006). The case study with six elevator shafts shows that the group of double-deck elevators with the DCS has more than three-fold handling capacity in up-peak compared to a group of single-deck elevators with the CCS (see Figure 3).

In Publications [III] and [VI], two new assignment policies are introduced for the double-deck DCS. These are important new techniques to improve passenger service quality during lunch traffic. The delayed deck assignment (DDA) policy allows the reassignment of the serving deck, while the delayed elevator assignment policy (DEA) allows the reassignment of the serving elevator until the last moment. In practice, the DDA policy can be implemented with the current personal signalization and guidance, but the DEA policy requires new
Figure 3: Performance comparison of single-deck and double-deck elevator groups with both the CCS and the DCS. The numbers beside the curves indicate the average car load factor. Handling capacity can be interpreted as the arrival rate in which the average car load factor reaches 80%.

shared devices. According to the solutions to the benchmark instances in Publication [VI], the DDA policy could decrease passenger waiting and journey times by more than 10% compared to the immediate assignment (IA) policy of Publication [II]. The DEA policy could reduce waiting times by about 35% and journey times by about 15%. On the other hand, the results of Publication [III] show that the Pareto-optimal solutions for the DEA policy have about 10-15% shorter waiting and transit times than the IA policy. The rather large deviation in the estimates can be explained by the different objective functions. In Publication [III], the times include the elapsed time a passenger has already been waiting or in transit. Hence, the improvement percentages are affected by the IA policy applied during the simulations. In the case of Publication [VI], the times represent the remaining waiting and journey times, which better describe the improvements caused by repeated optimization.

The computation times of the genetic algorithm are studied in Publication [VI]. The results show that the EAM outperforms the DAM under the IA and DDA policy. Of particular importance is the fact that the computation times of the EAM are more than an order of magnitude shorter compared to the DAM when considering the DDA policy. This means that the DDA policy cannot probably be implemented with the DAM for a real-world EGCS. Moreover, in the case of the EAM, the DDA policy increases the computation time only marginally from the IA policy. Therefore, the EAM can readily be used to implement the DDA policy in practice. Regarding the DEA policy, the EAM solves the benchmark
instances faster than the DAM if the deck assignments in the lower level problem are solved by heuristic rules. Thus, the EAM is probably fast enough to apply the DEA policy in practice.

The single-objective approach may be adequate for the CCS due to incomplete information about passengers. On the other hand, experience has shown that waiting and journey time minimization in the DCS lead to quite different results. Building on this observation, the DD-EDP with the DCS is formulated as a multi-objective optimization problem in Publication [III]. Passenger waiting and transit times are considered as objective functions. The real-time genetic algorithm of Publication [II] is modified to minimize the weighted sum of two objectives and to explore the solution space. The modified algorithm is used to solve DD-EDP instances occurring in lunch traffic simulations. The Pareto-optimal solutions of these instances exhibit a convex Pareto-frontier (see Figure 4).

![DD-EDP Pareto-optimal solutions for 30 passengers](image)

Figure 4: Pareto-optimal solutions for the large DD-EDP instance presented in Publication [III].

4 Discussion

4.1 Theoretical and practical implications

In the light of Publications [IV] and [VI], the collective control principle often defines the optimal or nearly optimal route of an elevator. Hence, its frequent use seems justified.
In Publication [IV], some counter-examples show situations in which the collective control is suboptimal with respect to a passenger service quality criterion. Although such situations seem to arise rarely in practice, more complex elevator routes consisting of multiple roundtrips could offer a step towards more efficient routes (Tanaka et al. 2005a, Ruokokoski et al. 2015). However, care must be taken to satisfy the basic rules of elevator operation as described in Closs (1970). Interestingly, the decomposition approach of solving the double-deck elevator routing problem by look-ahead rules in Publication [VI] can be considered as the collective control principle for one double-deck elevator with the destination control system.

In the bilevel formulations of Publications [V] and [VI], the upper-level problem optimizes passenger service quality and the lower-level problems solve the shortest elevator routes. As shown in Publication [VI], also other objectives are important in the EDP in addition to passenger service quality, and especially so in the case of double-deck elevators. The benefit of the bilevel formulation is that the optimality of the lower-level problem is independent of the upper-level objective function. In other words, the problem of finding minimum-time routes remains always the same for the given upper-level assignments.

In Publication [V], an important result on the predictability of stochastic demands and customers is provided. When they are estimated for a wide range of scenarios, at least one of the scenarios closely conforms to the actual events. The key to the robust EDP is not to find the best-matching scenario but many realistic ones. Then, the robust solution to the EDP should perform well with respect to the objective function as well as deviate only slightly due to the uncertainties across all the scenarios. The robust EDP cannot be formulated as a linear mixed integer programming problem since the estimation of the uncertainties inevitably results in non-linearities. The stochastic bilevel optimal control formulation in Publication [V] seems a promising approach, since it hides the estimation of the uncertainties from the upper-level problem and, thus, does not make restricting assumptions about the estimation methods. To solve the robust EDP, the algorithm also needs to consider the scenarios. For example, the genetic algorithm could be developed further to also evolve another population which defines the scenarios (Herrmann 1999).

On the part of a double-deck elevator group with the conventional control system, the simulation results of Publication [I] confirm an earlier observation that up-peak boosting technologies such as double-deck elevators and the destination control do not yet perform optimally during lunch traffic (Siikonen 2000, Sorsa et al. 2006). This means that the increased handling capacity for up-peak cannot be fully exploited when building designers attempt to squeeze the elevator core to the minimum. This dissertation studies two solu-
tions to the lunch traffic challenge: the DCS with the delayed (elevator) assignment policy and the robust EDP accompanied with the prediction of near-future passengers by the compound Poisson process. The DCS with the delayed elevator assignment policy requires new shared signalization and guidance devices. The first such system will be realized in 2018 in a 118-storey building in Kuala Lumpur. On the other hand, the prediction of the near-future passengers requires reliable supporting processes to model building traffic, namely, passenger counting, batch size estimation as well as forecasting of the arrival process parameters (Siikonen 1997, Backlund 2015, Kuusinen et al. 2015). When designing tall buildings with double-deck elevators according to the current industry practice, elevator shafts are reduced by about 30% compared to single-deck elevators with the CCS (Fortune 1996). By applying the studied methodologies, tall buildings could possibly be designed with 50% fewer elevator shafts. To reduce the elevator core further, however, requires multi-deck elevators and other multi-car solutions.

In Publications [I], [II] and [VI], optimization models and solution algorithms are developed for a double-deck EGCS. Further, in Publications [I] and [II], they are implemented in a prototype EGCS and integrated with elevator traffic simulation. The principles presented in Publication [II] were also used in the first realized double-deck DCS with the immediate assignment policy. Based on the results of Publication [VI], the new bilevel approach and its solution by the genetic algorithm should be fast enough to implement the double-deck DCS with the delayed deck and elevator assignment policy in practice, which may not be the case for the original approach of Publication [II]. In addition, the approach of Publication [VI] should also outperform that of Publication [I] for the double-deck CCS with the delayed assignment policy. The optimization models and algorithms of Publications [I], [II] and [VI] could be generalized for multi-deck elevators as well as for elevator groups consisting of elevators with an unequal number of decks. In addition, the bilevel model of Publication [VI] forms a good basis for modelling two or more independently moving elevator cars in a single shaft.

The simulation results of Publications [I] and [II] show that the minimization of passenger journey times in the double-deck elevator dispatching problem increases elevator group handling capacity. This also applies to single-deck elevators with the DCS (Sorsa et al. 2006). Publication [III], on the other hand, reveals that passenger waiting and transit time are conflicting objectives producing a convex Pareto-frontier. In practice, the EGCS needs to balance between the optimization of handling capacity and passenger service quality, and the transition between these should be smooth. The multi-objective approach of Publication [III] should provide desirable characteristics for handling varying passenger demands.
In practice, the weights of the objectives should adaptively change according to the underlying traffic condition. Such a mechanism would be further complicated if elevator energy consumption was considered as the third objective (Tyni and Ylinen 2006).

It seems that exact solution algorithms cannot be used in the EGCS since real-world EDP instances may be of arbitrary complexity. The EGCS must solve instances of the elevator dispatching problem in a fraction of a second. Nevertheless, exact algorithms provide an important alternative for benchmarking metaheuristics like the genetic algorithm. Even the use of metaheuristics in the EGCS requires additional techniques and high-performance software, as shown in Tyni and Ylinen (2001) and Publication [VI]. The new modelling approaches introduced in Publications [III], [V] and [VI] show that the EDP falls naturally into the fields of bilevel, multi-objective and robust optimization. All these methodologies increase the complexity of solving the EDP and require an ever more growing need for computational resources. Luckily, multi-core processors are becoming common also for industrial computing and should be up to the challenge.

The increasing complexity of the models also affects the requirements for simulation programs that are used to demonstrate the advantages of new control principles. Since the robust optimization proposed in Publication [V] aims at taking into account the uncertainties arising from the passengers’ tendency to travel in batches, its performance should be demonstrated by a traffic simulation that models both the batch arrivals and passengers’ behaviour in batches (Sorsa et al. 2013). Therefore, as experience and knowledge of the batch behaviour accumulate, batch arrivals should also be taken as the standard in the simulation studies conducted during the building design phase. This, on the other hand, indicates that design practices as well as the used simulation programs need to be modified accordingly within the industry.

### 4.2 Future research directions

The optimization models and algorithms presented in Publication [VI] should be integrated with a traffic simulation program as well as tested with real EGCS hardware. The bilevel approach can be further developed by studying metaheuristics to solve the lower-level problem since the exact solution algorithm was found too slow for on-line optimization. Then, the space savings of the double-deck DCS with the delayed elevator assignment policy could be assessed using the latest elevator group design practices and standards. The optimization model can be extended for the dual elevator system. Then, the traffic simulation provides
insight into the ongoing debate within the industry about the most efficient space saving technology.

In Publication [V], robust optimization based on discrete risk scenarios is presented as the next step in improving passenger service quality of the DCS with the immediate assignment policy. Risk, in this context, means the deviation from the planned elevator routes which are determined by a solution to a snapshot elevator dispatching problem. Many technical questions of the robust approach remain unanswered, for example, how to solve the robust elevator dispatching problem, how to balance between taking and avoiding risks and how inherent estimation errors in distribution parameters affect the optimization results. Once these questions have been adequately answered, the robust optimization model should be integrated with a traffic simulation program. The simulations also help finding good default values for control parameters that inevitably are needed to fine-tune the performance of the EGCS in real installations.

Optimization objectives of the elevator dispatching problem could also be studied further. This dissertation considered only linear objective functions with respect to passenger service quality criteria such as waiting, transit and journey time. Also quadratic and piecewise linear objective functions could be considered. Linear objectives, however, simplify the analytic use of the formulations of Publications [IV]–[VI]. For example, it should be possible to obtain analytic optimality conditions for the collective control principle, which could serve as useful heuristics in the EGCS.

In Publication [III], the multi-objective approach for the EDP is studied only briefly leaving many open questions. First of all, one should decide whether to apply multiple objectives in both the upper- and lower-level problems or only in the upper-level problem. Especially in the case of the double-deck elevators, also other objectives may arise for the lower-level problem in addition to those considered in Publication [VI]. Since eco-efficiency has become an important part of elevator operation, energy and power consumption could be studied further in the multi-objective approach. Naturally, the multi-objective and robust optimization approaches should be combined in the end. This requires further research on the genetic algorithm, i.e., how to efficiently solve the multi-objective problem across different risk scenarios. Also other metaheuristics could be tried on solving these kinds of problems in a real-time EGCS even though the genetic algorithm achieves the real-time performance.
References


Segercrantz, N. 2010. Impact of traffic on annual elevator energy consumption in high-rise buildings.


