Computers & Industrial Engineering 113 (2017) 589–601

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

Real-time recovering strategies on personnel scheduling in the retail industry

Michael Mac-Vicar[a], Juan Carlos Ferrer[a], Juan Carlos Muñoz[b], César Augusto Henao[c,*]

[a] Industrial and Systems Engineering Department, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Macul, Postal code: 7820436, Santiago, Chile
[b] Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Macul, Postal code: 7820436, Santiago, Chile
[c] Department of Industrial Engineering, Universidad del Norte, km. 5 Vía Puerto Colombia, Barranquilla, Colombia

ARTICLE INFO

Keywords:
Labor contingency
Multiskilling
Personnel scheduling
Real-time adjustment
Retail services
Workforce flexibility

ABSTRACT

Retailers must frequently deal with alterations in planned customer service levels due to unexpected demand variations or unscheduled employee absences. Although personnel scheduling techniques have been extensively studied and successfully applied, previous treatments of scheduling adjustments in response to demand and employee contingencies have not systematically considered all of the relevant issues. After presenting a mathematical specification of the problem, this study develops various algorithms that search for the best adjustments among all available contingency recovery resources, including transfers of multiskilled employees between different store areas. The proposed formulations also permit interaction between the user/decision maker and the affected employees. The underlying objective is to maximize profits, favoring solutions with fewer schedule modifications in order to minimize worker dissatisfaction. Due to the complexity of the basic model, the problem is divided and simplified using two greedy heuristics. Both algorithms can be implemented with real-world size problems and reach good solutions within minutes. Multiskilled employees prove to be an important reserve capacity for recovery of service levels in the face of unexpected variations. Empirical results using real data from a Chilean chain retailer show that in the worst scenario, the proposed model's schedule adjustments reduced lost profits due to unexpected variations by 18%.

1. Introduction

The retail trade is highly labor-intensive as well as extremely competitive (Nissen & Günther, 2010). Businesses in the rapidly growing industry face both predictable phenomena such as demand seasonality and unpredictable ones like demand uncertainty and unscheduled personnel absences. Unlike other sectors of the economy where demand is regular and predictable, retailers must deal with demand volumes that may vary dramatically over the course of a single day or the days of the week (Cuevas, Ferrer, Klapp, & Muñoz, 2016). In this scenario, and given the often complex labor law restrictions on working hours in the industry, planning shift schedules to efficiently meet the requirements of customer demand is no easy task. The goal of retail firms is to minimize labor costs while maintaining the best possible customer service levels, but even those with sophisticated workforce planning systems may find themselves confronted at different moments of the week with overstaffing or understaffing problems (Henao, Muñoz, & Ferrer, 2015). Overstaffing in this context refers to periods when there is an excess of employees on duty (and often idle) for a given desired level of customer service while understaffing denotes periods when employee requirements exceeds on-duty personnel. In latter case, the staffing imbalance can be propagated across successive periods. Thus, poor management of these issues may result in significant sales revenue losses, deterioration in customer service and a negative impact on a firm's business reputation (Kabak, Ülengin, Aktas, Önsel, & Topcu, 2008).

Personnel scheduling is traditionally a static process in that the scheduling decisions are all made at the start of the planning horizon, the assumption being that they cannot be adapted dynamically to address the stochasticity inherent in supply and demand. The process of planning a given day, week or month usually begins several weeks beforehand on the basis of demand forecasts. In some cases, it is executed by sophisticated shift-assignment systems that optimally allocate available personnel according to demand requirements. However, in the days and weeks following definition of the assignments, decision-makers will inevitably be confronted with a series of new and unpredictable events that force them to make adjustments.

In Bard and Wan (2005), the authors suggest a hierarchical three-phase approach to schedule adjustment. In the first phase, employees are assigned to daily shift and days-off patterns over a short-term

* Corresponding author.
E-mail addresses: mmacvicar@gmail.com (M. Mac-Vicar), jferrer@ing.puc.cl (J.C. Ferrer), jcm@ing.puc.cl (J.C. Muñoz), cahenao@uninorte.edu.co (C.A. Henao).

http://dx.doi.org/10.1016/j.cie.2017.09.045
Received 15 February 2017; Received in revised form 26 August 2017; Accepted 27 September 2017
Available online 28 September 2017
0360-8352/

planning horizon, typically a single week. In the second phase, new information is added to the weekly scheduling process on predictable events such as scheduled personnel absences (e.g., medical leave, holidays, resignations) and expected average demand increases (e.g., sales promotions, introduction of new products). The third phase, and the one that will concern us here, concerns the real-time adjustments (i.e., day to day) in personnel scheduling in response to the occurrence of events more difficult to predict.

Contingencies arising at the last minute that affect the quality of employee workforce planning may originate in demand and/or staffing level factors. In the case of demand, variations may occur due to internal events such as marketing campaigns, unanticipated public events or other eventualities such as the weather. As regards staffing, levels may be altered if an employee for any reason cannot work the assigned hours or because of new recovery resources that were not included in the plan. An example of a recovery resource is overtime. In either case, to ensure a day's shift scheduling requirements are met in the face of an unexpected occurrence, recovery resources must be applied once demand and actual personnel availability are known so that the contingency can be mitigated at minimum cost. The recovery resources most widely used by service industry companies are the following:

(1) *Overtime:* Involves extending an employee's shift for the day, either by moving up a shift's start time or moving back its end time. This is the most common recovery resource given that it is simple to implement, always available and can be assigned as soon as a staffing deficit appears. On the other hand, overtime is associated with lower productivity, increased stress and other potentially negative consequences of extended shifts.
(2) *Employee call-in:* Refers to employees contracted for a certain minimum guaranteed number of weekly hours but who can be called in for additional shifts. As well as the hourly remuneration, they receive standby pay for being on call during specified hours. The additional shifts must be of a certain minimum length and are subject to a predefined advance notice requirement.
(3) *Temporary employees:* These are employees supplied by an external firm, thus ensuring a certain level of availability. The pros and cons of such an arrangement depend in large measure on the terms of the contract, especially those regarding productivity and availability guarantees (Milner & Pinker, 2001). The main disadvantages are the training, remuneration and administrative costs incurred by the retailer.
(4) *Reallocation of multiskilled employees:* Consists in transferring employees between different store areas or departments, implying that the employees must be multiskilled for each area or department they may be allocated to. This strategy allows shift hours to be dynamically redistributed by making such reallocations for the precise period the contingency lasts without incurring significant labor cost increases. The main cost factors in this resource are the training program and related workforce planning activity. For the service industries and the retail trade in particular, Henao et al. (2015) and Henao, Ferrer, Muñoz, and Vera (2016) showed that total multiskilling is unnecessary. The best cost-effective performance was in fact obtained from a combination of specialized (i.e., employees trained in a single skill) and multiskilled employees, the majority of the latter trained in one additional skill.
(5) *Shift modification and new shift assignment:* A shift modification is defined as a change made to an existing shift while a new shift assignment entails creating a shift for an employee on a scheduled day off. In the latter case it likely also involves the cancellation of some other previously assigned shift. Generally, shift modification is possible if employee acceptance is obtained. The disadvantage of this resource is its lack of flexibility given that it does not allow for small local adjustments, and like employee call-in it is subject to a predefined advance notice requirement.

Various researchers have studied personnel scheduling adjustment processes using recovery resources to mitigate the negative effects of uncertainty (e.g. Bard & Wan, 2005; Pinker & Larson, 2003; Menezes, Kim, & Huang, 2006; Hur, Mabert, & Bretthauer, 2004a; Hur, Mabert, & Bretthauer, 2004b; Orsoni, 2004; Easton & Goodale, 2005). However, they propose make the adjustments automatically and in a single step, with no interaction involving the decision maker and the affected employees. This implicitly assumes that all of the changes suggested will be accepted by both the decision maker and the affected employees. In practice, however, the limited availability of each recovery resource leads inevitably to a sequential decision process. Since the willingness of any given employee to accept a proposed adjustment cannot be known in advance, it is required a sequential decision process that allows interaction between the decision maker and the affected employees.

The present article proposes a contingency control system with interaction based on employee schedule adjustment. Retailers generally define a number of control points for weekly workforce planning depending on their particular needs (e.g., one week beforehand, start of the week, start of the working day). The proposed system is designed to guide the decision maker at each such point in defining the best feasible action plan for recovering as much as possible of the benefits that would otherwise be lost due to a contingency. Additionally, the proposed system also permits an interaction between the user/decision maker and the affected employees. In Section 3, we provide a detailed description of our approach to develop a real-time system for adjusting personnel schedules.

The contribution made by the present study to the personnel scheduling problem consists in the development of a system that provides real-time support for decision makers tasked with determining scheduling adjustments to recover from a contingency once the system have the information about the current demand, staff availability, and the recovery resource alternatives. The proposed solution methodology considers two types of contingency: demand uncertainties and unscheduled personnel absences. Two different methods for addressing an identified contingency are designed. The first one is based on an iterative algorithm that solves a relatively big mixed integer programming model using column generation. The second one incorporates a much more practical approach, evaluating the scheduling adjustment alternatives by performing a local search routine and then greedily applying the most beneficial adjustments. Finally, we present the results of the implementation of both proposed methods for a study case of a Chilean retailer. Experiments are conducted using real-world instances in order to evaluate and compare both methods under various absenteeism scenarios and levels of priority for schedule planning stability. Note that, although contingencies can be attributed to demand variations and/or unplanned personnel absences, in this study case the experiments are conducted only for absenteeism.

2. Review of the literature

There exists an extensive literature on personnel scheduling that addresses various types of problems and solutions with applications to different industries. Thorough surveys of the field may be found in Ernst, Jiang, Krishnamoorthy, Owens, and Sier (2004)a), Ernst, Jiang, Krishnamoorthy, and Sier (2004)b), Alfares (2004) and Van den Bergh, Belien, De Bruecker, Demeulemeester, and De Boeck (2013). It has been shown that solution approaches to the general scheduling problem can be either deterministic or stochastic, in the latter case by incorporating

Table 1
Characteristics of previous studies on contingent resources for schedule adjustment.

Reference	RTSA	A	DU	Adjustment options	Objective function	Method	Application
Easton and Goodale (2002)	No	Yes	Yes	Overstaffing	Max. Expected revenues - Expected labor expenses for turnover	Markov process, Simulation	Service delivery system
Pinker and Larson (2003)	No	Yes	Yes	Overtime, Temporary workers, On-call workers	Min. Cost of regular workers + Cost of contingent workers	Dynamic programming	Services
Hur et al. (2004a) and Hur et al. (2004b)	Yes	Yes	Yes	Shift changes, Reallocation of multiskilled employees	Min. Labor shortage + Direct labor cost - Overall productivity + Shift changes + Task rotations	Goal programming, Heuristics, Simulation	Quick service restaurants
Orsoni (2004)	Yes	Yes	No	Overtime, Reallocation of multiskilled employees	Max. Quality service - Overtime costs + Job satisfaction	Genetic algorithms, Simulation	Retail
Bard and Wan (2005)	No	Yes	Yes	Overtime, Shift changes, Increasing part-time hours, Calling temporary workers	Min. Direct labor cost + Overtime cost + Cost of temporary workers + Cost of additional part-time hours	Mixed integer linear program, Heuristic	United State postal service
Easton and Goodale (2005)	Yes	Yes	Yes	Overtime, Temporary workers, Call-in workers	Max. Expected revenues - Expected labor expenses for turnover	Markov process, Simulation	Services
Menezes et al. (2006)	No	Yes	Yes	Reallocation of multiskilled employees among stores or among sections in a store	Max. Overall expected profit rate of the chain	Markov process, Greedy-type algorithm	Retail
Zhang et al. (2009)	Yes	Yes	Yes	Overtime, Call-in casual employees	Min. Total overtime + Casual costs	Integer program, LP-based decomposition algorithm	United State postal service
Easton (2011)	Yes	Yes	Yes	Reallocation of multiskilled employees	Max. Sum of fringe benefits - Direct labor costs	Mixed integer linear program, Simulation	Services
Campbell (2011)	Yes	No	Yes	Reallocation of multiskilled employees	Max. Total expected utility - Direct labor costs	Mathematical programming, Heuristic	Services
Cuevas et al. (2016)	No	No	No	Reallocation of multiskilled employees	Max. Total expected utility - Multiskilled employee transfer cost	Mixed integer linear program	Retail
Henao et al. (2015)	No	No	No	Reallocation of multiskilled employees	Min. Direct labor cost + Training cost + Understaffing cost	Mixed integer linear program	Retail
Henao et al. (2016)	No	No	Yes	Reallocation of multiskilled employees	Min. Training cost + Understaffing cost + Overstaffing cost	Mixed integer linear program, Heuristic, Robust optimization	Retail

uncertainties into the schedule adjustment process. However, the great majorities of studies stick to a deterministic view of the problem and simply ignore contingencies involving demand variations or unscheduled personnel absences (Van den Bergh et al., 2013). A few deal with the scheduled absence problem by assigning in advance more hours than are needed to cover demand, but this technique cannot solve the problem of adjusting personnel schedules when a demand or personnel supply contingency arises. As is noted in Easton and Goodale (2005), anticipating absences via structural overstaffing not only increases labor costs just as the contingencies it is designed to mitigate do, but also reduces the future availability and efficiency of contingency resources—hence the need for a real-time system for adjusting personnel schedules.

Previous works that have tackled the use of recovery resources for schedule adjustment in the face of demand variations and/or unplanned personnel absences are listed in Table 1 along with brief indications for each study on the following key characteristics:

(1) *Real Time Schedule Adjustment* (RTSA): Indicates whether the study addresses contingency management by making schedule adjustments in real time (i.e., day to day or during the day) to cope with unexpected demand changes and/or disruptions of labor supply.
(2) *Absenteeism* (A): Indicates whether the model of the problem addresses unscheduled personnel absences.
(3) *Demand Uncertainty* (DU): Indicates whether the model of the problem addresses uncertainty of demand.
(4) *Adjustment Options:* Indicates what actions are taken and resources used by the proposed model as contingency management measures to recover service levels.
(5) *Objective Function:* Indicates the components of the objective function.
(6) *Method:* Indicates what type of solution methodology is used.
(7) *Application:* Indicates in which industry or economic sector the study was applied.

Brief reviews of each of the studies summarized in Table 1 are presented below. We begin with those that do not present a solution approach in real time and then consider those that do

2.1. Non-real time contingency control models

In Easton and Goodale (2002) the authors develop a short-term model that optimally determines workforce size and work hours to compensate for anticipated employee turnover and absenteeism. Their premise is that both demand and employee availability are random variables. They also assume a dynamic workforce that varies in experience and size due to resignations and hirings. In Pinker and Larson (2003) a dynamic programming model is developed that determines the optimal number of regular employees and the optimal size of a contingent employee pool. Under this approach the decision maker makes two sequential decisions in each period: how many contingent employees to hire and how many extra hours to schedule. However, the affected employees are not involved in the decision process. Since it is primarily concerned with staffing size decisions, the study focuses mainly on the early planning stages but it draws some interesting conclusions on the impacts of labor flexibility and information availability on contingent resources. In Bard and Wan (2005) the authors create a model of weekly scheduling adjustments for the mail processing and distribution centers of the U.S. Postal Service. The model's main purpose is to generate shift adjustments in the stages preceding their execution. A set of heuristics are developed to solve the most difficult instances, which are then evaluated in terms of computational cost and solution quality.

In Menezes et al. (2006), the authors attempt to determine the optimal number of employees at a retail chain and how to distribute them among the various stores. The problem is formulated as a multi-server queueing system in which customers are served only when there is an employee available. A greedy algorithm defines the optimal allocation policy for each possible number of employees present based on demand characteristics and the parameters of each store. Experiments are performed to compare a scenario in which the stores coordinate their actions to share employees under the optimal policy with one in which they act independently. In Cuevas et al. (2016) and in Henao et al. (2015), the authors develop optimization models to simultaneously assign days off and working shifts for each employee where each working shift could have a different activity, shift start period and length of workday. Particularly, Henao et al. (2015) explain that in the presence of understaffing, the multiskilled employees always generate greater savings in total costs than with strictly specialized personnel. Finally, in Henao, Ferrer, Muñoz, and Vera (2016), the authors propose a mixed integer linear programming model for the problem of weekly assigning employees to the departments of a retail store under uncertain demand. The solution method seeks to determine the optimal staff levels of multiskilled and specialized employees for each department store that minimizes the supply/demand mismatch.

2.2. Real time contingency control models

In Hur et al. (2004a) the authors study schedule adjustment policies in the presence of absenteeism and demand fluctuations. Their objective is to evaluate the impacts on adjustment quality of decision maker experience, the percentage of part-time employees and demand forecast uncertainty. They conduct an experiment in which a large number of decision makers make adjustments under hypothetical scenarios with variations in workforce mix, demand forecast error and employee willingness to accept schedule changes. Two worker adjustment acceptance levels were assumed. Worker acceptance captures the extent that employees are willing to accept requested schedule changes. The responses are evaluated by simulation and the results are compared to the benefits of a benchmark solution generated by an automated shift optimization system. In Hur et al. (2004b) the same authors focus on the final phase of the adjustment period, developing a real-time adjustment model for a fast-food chain that uses contingent resources. The first of two proposed solution policies maximizes productivity and minimizes cost while the second stresses worker convenience in terms of the number of adjustments and task rotations per employee. Heuristics are developed to solve the problem rapidly. In this second paper, the proposed formulation assumes that the decision maker has identified all employees who are willing to accept schedule changes.

In Orsoni (2004) the author proposes a decision support system for the assignment of resources to workgroups in a retail setting. The system uses genetic algorithms and simulations to iteratively construct assignments. One algorithm assigns the resources available for each area twice a week and a second one operates when the assigned resources are unavailable. Unlike the first algorithm the second one works online, rescheduling resources every hour. It searches for available resources in other areas based on customer demand in real time and the current degree of utilization of each resource. In fact, at beginning of each time unit the pool of available resources is updated according to a stochastic available rate. The author shows that the amount of overtime can be significantly reduced while still maintaining service and employee satisfaction levels through careful resource reallocation. In Easton and Goodale (2005) the authors develop a model that determines the best shift schedules and workforce sizes for a service sector firm and a second model that uses contingent resources to handle demand fluctuations and unplanned employee absences. In this paper, all the adjustments are made assuming that demand and actual personnel availability are already known. The model can use planned overtime and anticipated absenteeism during the process of creating schedules. The study compares the performance of various contingent resource use policies including the effects of factors exogenous to the model such as worker productivity, the pattern and amplitude of demand and per-capita labor costs.

In Zhang, Chakravarthy, and Gu (2009), the authors study how to adjust production and workforce scheduling for a U.S. Postal Service mail processing and distribution centre. The problem formulation considers demand fluctuations and absenteeism, and the schedule adjustments are made day-to-day using overtime and employee call-ins. In Campbell (2011), a two-stage stochastic program is developed for scheduling and allocating multiskilled employees in a multi-department service environment with random demand. The first stage schedules days off over a time horizon such as a week or a month while the second stage allocates the multiskilled employees at the start of each day to match staffing with realized demand. Finally, in Easton (2011) the author proposes a two-stage stochastic model that assumes demand is variable and integrates decisions on multiskilling, staffing level and employee scheduling and allocation. The first stage solves a deterministic model that optimizes these decisions. The results are then inputted to the stochastic second stage, which uses a simulation model to reallocate available multiskilled employees in response to random demand variations and unscheduled personnel absences. In these three papers, all the adjustments are made assuming that demand and actual personnel availability are already known by the decision maker.

The literature on the use of recovery resources for schedule adjustment in real time demonstrates a considerable effort to improving its understanding. All of the works reviewed above use methodologies of one sort or another to determine the use of contingent resources in the face of demand and/or supply variability. However, they employ a one-stage decision process and do not permit interaction between the decision maker and the affected employees. In all cited papers, the affected employees are not involved in the decision process, since that all the adjustments are made assuming that a pool of employees are willing to accept the requested schedule changes. By contrast, we propose a sequential decision process that allows interaction between the decision maker and the affected employees. Thus, in our study a proposal for a change can be or cannot be accepted/rejected by the decision maker and the affected employees.

3. Description of the problem

The problem under study consists in providing real-time adjustments in personnel scheduling to mitigate the negative impacts of the inherent uncertainties involved in this process. Specifically, the problem we propose consists in a two-phase approach to schedule adjustment and to improve customer service levels in a retail store. In the first phase, employees are assigned to daily shift and days-off patterns over a weekly planning horizon. The second phase, and the one that is the core of this paper, concerns the real-time adjustments (i.e., day to day) using shift scheduling recovery resources in response to the occurrence of unexpected events (with interaction involving the decision maker and the affected employees).

To develop a real-time system for adjusting personnel schedules we propose the following approach. At the beginning of each working day, the proposed system must answer three questions: (i) what recovery resources should be used, (ii) where the resources can be obtained, and (iii) how they should be applied. To answer these questions, the system must have information on the updated expected demand for the day, the staff availability that is expected to show up, and the recovery resource alternatives. With this data we conduct an exhaustive search of the best actions for adjusting staff levels to the updated demand forecast involving the remaining days of the planned week. But since the ultimate feasibility of the available options cannot be determined without the consent of the affected employees, the system suggests several actions for one employee, each of them involving one resource - overtime, for example, or a shift modification or a new hiring. The decision maker then submits these options to the employee in question, who may reject some or all of them. Based on the employee response, new actions for a second employee are generated by the system. The process is iterated with other employees until the system has no more alternatives to propose or the decision maker determines that the contingency has been resolved.

3.1. Cost of recovery resources

Each of the system's recovery action options is represented by a new weekly schedule for an employee and has associated costs that depend on the nature of the action. These costs include the expense of a person-hour, an extra hour (i.e., overtime), a penalty for changing an existing schedule, replacing a shift with a day off, replacing a day off with a new shift, and an multiskilled employee transfer from one store area to another in a given period of the day.

3.2. Objective

The objective of the recovery system's search for the best action is defined in terms of two factors: productivity and convenience.

In the case of productivity, the objective of any retail store is to maximize store profits by reducing losses due to the mismatch between demand and staffing level. For this reason, we agree with several studies listed in Table 1, which chose to disregard the cost minimization objective and replace it with a utility function depending upon the person-hours covering labor requirements for each area at each period of time. Similar to Cuevas et al. (2016), we modelled the profits by a marginal benefit function for each store area and period. This function describes the marginal benefits of an additional person-hour, that is, the benefits conferred by the presence of the n-th employee relative to demand, and is assumed to be convex and decreasing in n. Cuevas et al. (2016) explain that considering a decreasing marginal utility function allows decision makers to assign the next employee according to the areas and periods where the greatest marginal impact can be achieved. Accordingly, the function is linearized and discretized into a certain number of marginal benefit percentage ranges for each area and period of the day. Each such benefit range represents a percentage of coverage of demand of the labor requirements. It is assumed that the coverage of one person-hour at a particular segment generates a fixed marginal utility. Later, in Section 5.1, we will introduce in detail our approach to calculate these marginal benefits.

As for the convenience factor, the idea is to reduce administrative costs and favour a more stable personnel planning by minimizing the cost of the staffing actions taken, the number and magnitude of the modifications made to existing shift schedules. For this purpose, costs are assigned to each available schedule depending on its characteristics and the recovery resources used.

3.3. Modeling assumptions

Our formal definition of the problem incorporates a series of assumptions: (1) The proposed solution is a system designed to take into account the current week staff plan and demand forecast, all events affecting staffing levels (e.g. absenteeism or training that was not originally planned), demand variations with respect to the plan expected for the day, and recovery resources availability. (2) The recovery resources considered are: overtime, employees call-in; the transfer of multiskilled employees between store areas, and the modification or elimination of existing shifts and the creation of new ones. (3) The retail store has different areas, and the personnel requirements for each area, time period, and day, is known and it is allowed to satisfy it only partially, which involves a penalty cost. This cost varies by store area. (4) The daily planning horizon is divided into small time periods (e.g., 15 or 30 min) in which the demand forecast is assumed constant. (5) Each employee works under specific type of contract (i.e., full-time or part-time). (6) For the sake of simplicity in the model, employees are homogeneous in their productivity regardless of how many departments each of them is trained to work in, or in the number of hours that the employee have already worked in the day. Henao et al. (2015) and Henao et al. (2016) justify this assumption because the complexity levels of tasks in retail are both low and rather uniform (compared with other industries such as manufacturing, health, and call centers, for which heterogeneous productivity is often assumed). (7) Each employee can work in a given set of feasible areas. Thus, our formulation integrates multiskilling as an input. (8) Each employee has a base area in the store and any "transfers" to other (non-base) areas have a cost proportional to the hours assigned. (9) All shift adjustments to be offered to employees still satisfy all legal constraints (i.e. minimum or maximum number of hours to be worked in a day).

4. Methodology

As explained earlier, two different methods are proposed for solving the problem. In the first one, we formulate a mixed integer linear programming (MIP) model for solving iteratively the shift scheduling problem. At each iteration the method searches for the best schedule adjustments for each store area and permits transfers of multiskilled employees from one area to another. In the second method, different transfer alternatives and schedule modifications are evaluated by a local search routine. The best alternative is then applied and the search continues. The two methods are developed in the following subsections: (4.1) Method 1: Iterative MIP system; and (4.2) Method 2: Greedy local search system.

4.1. Method 1: Iterative MIP system

Barring a few minor differences, the iterative MIP model used in the first method closely resembles a number of others found in the literature. The model optimizes a set of areas but only one of them can receive the multiskilled employees from the others. There are two reasons for this restriction. One is that it greatly reduces the complexity of the general problem in which transfers are made to and from any area, thus ensuring the problem can be solved in a reasonable amount of time. The second reason is that organizational structure and hierarchical issues may prevent centralized decision-making on shift modifications and planning changes for individual areas. In effect, decisions must often be made separately for each case, a characteristic we exploit here to simplify the model. By breaking down the problem thus, each partial solution can be evaluated and the definitive solution can then be constructed progressively, integrating accepted changes one by one. Furthermore, the area receiving transfers can be checked after each iteration.

The method for solving the iterative MIP approach is explained in the following subsections: Notation, Pre-processing, MIP model, Column generation, and Execution of the system.

4.1.1. Notation

In this subsection we introduce the sets, parameters, and variables that are used in the formulation of the model. We define a schedule as the complete set of shifts an employee is planned to work over the weekly planning horizon. For the purpose of this model, a schedule j is defined by the elements α_{jt}, where $t \in T$, being equal to 1 if the employee is assigned to work in period t during the day being optimized, and 0 otherwise.

Model sets:

- T Subdivisions or periods over the daily planning horizon, indexed by t. Periods of 15 or 30 min are typical examples
- R Benefits ranges, indexed by r
- A Store areas or departments, indexed by a
- E Employees of the store, indexed by i
- E_a Employees belonging to store area a , $\forall\ a \in A$
- Y_a Employees qualified to work in area a but belonging to a different area, $\forall\ a \in A$
- S_i' Selected weekly schedules for employee i , $\forall\ i \in E$
- S_{it}^+ Schedules generated for employee i in which a transfer can start in period t , $\forall\ i \in E$, $t \in T$, $S_{it}^+ \subset S_i'$
- S_{it}^- Schedules generated for employee i in which a transfer can end in period $t+1$, $\forall\ i \in E$, $t \in T$, $S_{it}^- \subset S_i'$.

Model parameters:

- a^* Area to which transfers can be made, $a^* \in A$. We define $A' = A \backslash \{a^*\}$, such that the set A' does not include the area a^*.
- α_{jt} Equal to 1 if an employee assigned to schedule j works in t , $\forall\ j \in S_i', t \in T$
- γ_{jt}^+ Equal to 1 if schedule j allows a transfer to start in period $t+1$, $\forall\ j \in S_i', t \in T$
- γ_{jt}^- Equal to 1 if schedule j allows a transfer to end in period t , $\forall\ j \in S_i', t \in T$
- m_{at} Minimum staffing level in person-hours for area a in period t , $\forall\ a \in A, t \in T$
- c_{ij} Total cost of all recovery resources (excluding relocation of a multiskilled employee to a different area) needed for employee i to work schedule j , $\forall\ i \in E$, $j \in S_i'$
- δ_a Cost associated with a transfer from area a to area a^*, $\forall\ a \in A'$
- h_a Employee deficit penalty for each employee below the required minimum in area a in any period, $\forall\ a \in A$
- r_{atr} Marginal benefit of a person-hour in area a, in period t, with benefit range r , $\forall\ a \in A$, $t \in T$, $r \in R$
- ω_{atr} Proportion of desired person-hours assigned to area a, in period t, with benefit range r , $\forall\ a \in A$, $t \in T$, $r \in R$
- ϕ_{at} Desired number of person-hours for area a in period t , $\forall\ a \in A$, $t \in T$.

Model decision variables:

- U_{at} Number of employees present in area a in period t , $\forall\ a \in A, t \in T$
- X_{ij} Equal to 1 if employee i is assigned to schedule j, otherwise 0, $\forall\ i \in E$, $j \in S_i'$
- H_{at} Employee deficit in relation to required minimum number in area a and period t , $\forall\ a \in A$, $t \in T$
- R_{atr} Number of employees covering area a, in period t, with benefit range r , $\forall\ a \in A$, $t \in T$, $r \in R$
- V_{it}^+ Equal to 1 if employee i is transferred to area a^* starting with period $t+1$, otherwise 0, $\forall\ i \in Y_{a^*}$, $t \in T$
- V_{it}^- Equal to 1 if employee i stops working in area a^* starting with period $t+1$, otherwise 0, $\forall\ i \in Y_{a^*}$, $t \in T$
- P_{it} Equal to 1 if employee i is transferred to area a^* in period t, otherwise 0, $\forall\ i \in Y_{a^*}$, $t \in T$
- P_{it}' Equal to 1 if employee i is present in area a^* in period t, otherwise 0, $\forall\ i \in Y_{a^*}$, $t \in T$. Note that P_{it}' is equal to P_{it} for any period t in which employee i is not on lunch break.

4.1.2. Pre-processing

The goal of the preprocessing stage is to determine sets of feasible shifts assignments for each day and employee as a function of their contract terms, personal restrictions and recovery methods. In a retail store with real-world instances, the number of flexible shifts that could be generated for a single employee is extremely high to be included in a MIP model. To avoid having to enumerate all possible schedules for each employee i, a set of possible schedules S_i is expressed as a set J_i containing all possible series of workday lengths of the form $h_1,h_2,h_3,...,h_7$ that can be worked on days in the weekly planning horizon. These series take into account restrictions associated with the employee, the number of hours stipulated in his or her weekly contract and the amount of overtime he or she can work. They do not, however, indicate the workday start period.

A set of possible shifts S_{id} for each day d in the horizon is also generated. In addition to considering employee availability and restrictions, it specifies both the starting period of each workday and the lunch breaks. As will be explained in the subsection on column generation, the use of these sets of workday length series J_i and possible shifts S_{id} for each day will allow the model to search for the best candidates to be added to S_i' without having to enumerate all the elements of S_i, such that $S_i' \subset S_i$.

4.1.3. MIP model

In this subsection we formulate the MIP model for solving iteratively the shift scheduling problem as follows:

$$Max \quad \underbrace{\sum_{a\in A}\sum_{t\in T}\sum_{r\in R} r_{atr}R_{atr}}_{(a)} - \underbrace{\sum_{i\in E}\sum_{j\in S_i'} c_{ij}X_{ij}}_{(b)}$$
$$-\underbrace{\sum_{a\in A'}\sum_{i\in E_a}\sum_{t\in T} \delta_a V_{it}^{+}}_{(c)} - \underbrace{\sum_{a\in A}\sum_{t\in T} h_a H_{at}}_{(d)} \tag{1}$$

s.t.

$$\sum_{j\in S_i'} X_{ij} \leqslant 1 \quad \forall\ i\in E \tag{2}$$

$$U_{a^*t} - \sum_{i\in Y_{a^*}} P_{it}' - \sum_{i\in E_{a^*}}\sum_{j\in S_i'} \alpha_{jt}X_{ij} \leqslant 0 \quad \forall\ t\in T \tag{3}$$

$$U_{at} + \sum_{i\in Y_{a^*}\cap E_a} P_{it}' - \sum_{i\in E_a}\sum_{j\in S_i'} \alpha_{jt}X_{ij} \leqslant 0 \quad \forall\ t\in T, a\in A' \tag{4}$$

$$R_{atr} - \omega_{atr}\phi_{at} \leqslant 0 \quad \forall\ a\in A, t\in T, r\in R \tag{5}$$

$$\sum_{r\in R} R_{atr} - U_{at} \leqslant 0 \quad \forall\ a\in A, t\in T \tag{6}$$

$$m_{at} - H_{at} - U_{at} \leqslant 0 \quad \forall\ a\in A, t\in T \tag{7}$$

$$P_{it} - \sum_{t_1<t} (V_{it_1}^{-} - V_{it_1}^{+}) = 0 \quad \forall\ i\in Y_{a^*}, t\in T \tag{8}$$

$$V_{it}^{+} - \sum_{j\in S_{it}^{+}} X_{ij} \leqslant 0 \quad \forall\ i\in Y_{a^*}, t\in T \tag{9}$$

$$V_{it}^{-} - \sum_{j\in S_{it}^{-}} X_{ij} \leqslant 0 \quad \forall\ i\in Y_{a^*}, t\in T \tag{10}$$

$$P_{it}' - \sum_{j\in S_i'} \alpha_{jt}X_{ij} \leqslant 0 \quad \forall\ i\in Y_{a^*}, t\in T \tag{11}$$

$$P_{it} - P_{it}' + \sum_{j\in S_i'} \alpha_{jt}X_{ij} \leqslant 1 \quad \forall\ i\in Y_{a^*}, t\in T \tag{12}$$

$$P_{it}' - P_{it} \leqslant 0 \quad \forall\ i\in Y_{a^*}, t\in T \tag{13}$$

$$X_{ij} \in \{0,1\} \quad \forall\ i\in E\ j\in S_i' \tag{14}$$

$$V_{it}^{+} \in \{0,1\} \quad \forall\ i\in Y_{a^*}, t\in T \tag{15}$$

$$V_{it}^{-} \in \{0,1\} \quad \forall\ i\in Y_{a^*}, t\in T \tag{16}$$

$$P_{it}' \in \{0,1\} \quad \forall\ i\in Y_{a^*}, t\in T \tag{17}$$

$$P_{it} \geqslant 0 \quad \forall\ i\in Y_{a^*}, t\in T \tag{18}$$

$$U_{at} \geqslant 0 \quad \forall\ a\in A, t\in T \tag{19}$$

$$H_{at} \geqslant 0 \quad \forall\ a\in A, t\in T \tag{20}$$

$$R_{atr} \geqslant 0 \quad \forall\ a\in A, t\in T, r\in R \tag{21}$$

The objective function (1) is given below, its four component terms presented separately for ease of identification. The first term (a) maximizes the expected benefits; the second and third terms promote the stability of the employees' assignments by penalizing the costs of modifying schedules and applying the new version (b) and of transferring employees between different areas (c); and the last term (d) penalizes assignments that generate employee deficits relative to the minimum requirements for certain critical periods such as store opening and closing times. Recall that, the model optimizes a set of areas but only the area a^* can receive the multiskilled employees from the others.

The solution of the model must satisfy a series of constraint sets. The first set, given by (2) limits the assignment of schedules to a maximum of one per employee, the objective function itself guaranteeing that no employee receives none. Constraints (3) count the presence of employees in area a^*, including transfers from other areas. Constraints (4) count the presence of employees in the remaining areas; deducting those who have been transferred to area a^*. Constraints (5) restrict the size of each benefits range while constraints (6) assign the employees present to each range. Note that the properties of the benefits function ensure the employees are assigned to ranges in strict order from highest to lowest, each range being filled before passing to the next lowest one. Constraints in (7) count the employee deficit relative to the minimum personnel requirement. Constraints (8) regulate the transfers, forcing variables P_{it} to take a value of 1 only for the periods from the start to the end of a transfer inclusive. The start and end periods themselves are indicated by variables V_{it}^{+} and V_{it}^{-}, respectively. Constraints (9) and (10) restrict the respective moments at which a transfer can start and end, depending on the schedule assigned to the transferred employee. Lastly, constraints (11)–(13) impose that a transferred employee complies with the assigned schedule for the entire duration of the transfer. Finally, constraints (14)–(21) define the domain of each variable.

4.1.4. Column generation

As was noted in Section 4.1.2., workdays J_i and schedules S_{id} for each employee are generated during pre-processing. In this subsection, a selection of schedules is accomplished by a column generation routine set forth below as Algorithm 1. That is, to avoid that the model enumerates all the elements of S_i, it is possible to use only a subset of these elements (S_i') using a column generation routine so that the problem can be solved at minimal cost to the quality of the solution (Easton & Rossin, 1991). Phase 1 of the algorithm evaluates, for each day, the contribution of each possible shift S_{id} to the reduced cost of a new column according to the formula given in (22). To simplify the presentation, we have used vector notation. Thus, λ_n, λ_{an} and λ_{in} are dual cost vectors of dimension $\#T$ associated with the constraints n, for employee i and area a. The α_j term indicates the availability of schedule j and is also a $\#T$-dimension vector, defined by the parameters α_{jt}. The γ^{+} and γ^{-} terms indicate the possibility of transferring multiskilled employees at beginning and ending of any period, respectively. Both are defined by the parameters γ_{jt}^{+} and γ_{jt}^{-}. All of the multiplications in the equation are scalar products.

$$c_{ij} = \begin{cases} \theta_{ij} - \alpha_j\lambda_3 & if \quad i\in E_{a^*}, \\ \theta_{ij} - \alpha_j\lambda_{a4} - \gamma^{+}\lambda_{i9} - \gamma^{-}\lambda_{i11} - \alpha_j\lambda_{i11} + \alpha_j\lambda_{i12} & otherwise. \end{cases} \tag{22}$$

Thus, the contribution of a possible shift on a particular day to the term θ_{ij} is estimated from the costs associated with the shift such as the cost of modifying it or adding extra hours.

Algorithm 1: Column generation routine

Phase 1: Evaluation of daily shifts
Step 1: For each day d, calculate the contribution of each shift in S_{id} to the dual cost of any schedule.
Step 2: For each day, order the possible shifts by the value computed in *Step 1* and choose the best ones $\widehat{S}_{id}$.
Phase 2: Generation of weekly schedules
Step 3: For each series of workdays in J_i, enumerate all constructible schedules from the best shifts for each day, $\widehat{S}_{id}, d \in D$.
Step 4: For each schedule not yet in $\widehat{S}_i$, calculate the reduced cost c_j and add those with a positive cost to $\widehat{S}_i$.

In Phase 2, the best daily shifts are chosen according to some criterion and are combined using the series of workdays J_i calculated in Pre-processing (4.1.2.). The criterion we use is to take a selection of schedules with the greatest reduced cost for each possible workday length. Since the space we want to explore is defined by the number of daily schedules used to generate the shift combinations, it is generally advisable to avoid starting with many schedules that will never be used and gradually increase the number as the solution improves. The costs of the schedules themselves must also be added to θ_{ij}. Among these are the cost of eliminating an existing shift and, in the case of an call-in employee, the cost of any extra hours above and beyond the minimum guaranteed number.

4.1.5. Execution of the system

The result of Algorithm 1 is a weekly plan for each employee. These plans may include changes from the original schedule and/or transfers to area a^*. Algorithm 2, which we now present, iteratively executes the contingency control system to provide real-time support for decision makers. Note that, the Algorithm 3 (subroutine of the Algorithm 2) executes the MIP model to allow transfers of multiskilled employees to a different area at each iteration. These routines completely describe the optimization process.

Algorithm 2: Iterative recovery process

Phase 1: Preprocess
Step 1: Generate all possible workday series J_i and daily shift options S_{id} for each employee.
Step 2: Order the areas according to their priority in a list denoted *AreasOptimize*.
Step 3: Initially there are no modifications: *Modifications* $\leftarrow \varnothing$.
Step 4: Initially there are no transfers: *Transfers* $\leftarrow \varnothing$.
Phase 2: Optimization
Step 5: Remove the first area in *AreasOptimize* and assign it to a^*.
Step 6: Solve the MIP model using the *Algorithm 3* routine.
Step 7: Evaluate the changes in the solution. If they are all accepted by the decision maker, go to Phase 3; otherwise, restrict the possibilities to exclude the infeasible ones.
Step 8: Return to step 6.
Phase 3: Entry of modifications
Step 9: Enter each new or modified shift in the solution in *Modifications*.
Step 10: Enter each transfer in the solution in *Transfers*.
Step 11: If *AreasOptimize* $\neq \varnothing$, go to Phase 2; otherwise, process terminates.

Algorithm 2 begins with the preprocess and then orders the areas in accordance with some priority criterion. This ordering is important because the first areas that are optimized will have priority access to personnel excesses in other areas. One possible heuristic criterion is to order the areas according to the benefits lost due to contingencies compared to the original plan. Once it has ordered the areas, the algorithm then solves the problem iteratively for each one. After each iteration, the modifications made must be validated by the decision maker. If they are rejected by an employee, the iteration must be repeated with added restrictions prohibiting the rejected possibilities for that employee as infeasible. If the modifications are accepted, they become permanent and any other possibility for the employee on the days in question is prohibited. Since that the willingness of any given employee to accept a proposed adjustment cannot be known in advance, this step 7 is very important in our Algorithm 2. It generates a sequential decision process that allows interaction between the decision maker and the affected employees.

Algorithm 3: Solution of the MIP model

Phase 1: Initialization
Step 1: Generate an initial solution using the current schedules and the detected events (contingencies).
Step 2: For each modification in *Modifications* or shift in *Transfers*, limit the possibilities of the employee to the modified shift.
Step 3: For each transfer in *Transfers*, if it is made to an area other than a^*, prohibit transfers during the period in question and modify the model to include the presence of the transferred employee in the receiving area.
Phase 2: Solution
Step 4: Solve the dual of the relaxed problem.
Step 5: Obtain the set $\widehat{S}_i$ using the *Algorithm 1* routine for each employee $\forall\ i \in E$. If there is an i such that $\widehat{S}_i \neq \varnothing$, add the best n schedules to S_i' and return to *Step 4*.
Phase 3: Elimination
Step 6: Eliminate all inactive schedules ($X_{ij} = 0$).
Phase 4: Termination
Step 7: Solve the model using the integrality constraints and the set of remaining schedules.

To consider new transfers in the following iterations, in which they may be received by a different area, the presence of each transfer is simulated prohibiting new transfers during the affected periods and modifying the areas' presence constraints (3) and (4). That is, the assigned transfers are marked in order to insure that the same employee is not again transferred when looking at another area.

4.2. Method 2: Greedy local search system

Unlike the MIP system, the second proposed method employs a greedy algorithm that searches for recovery resources by applying heuristics similar to those a decision maker working manually would use. The areas and periods affected by a contingency are identified previously so that the actions taken can be directed exclusively at mitigating the unexpected event. Although this algorithm explores a very much reduced space in the set of existing recovery possibilities, the strategy here is that since the number of such possibilities remains high, a good level of recovery can still be achieved using only the most obvious ones. One of the advantages of this algorithm is that each action is applied individually without depending on the feasibility of any other action. Another advantage is that since the possibility space is so much smaller, the problem can be solved much more quickly.

4.2.1. Options used by local search

The options considered will depend on the employee but there are only three basic movements:

1. *Shift modification:* consists in modifying a shift the same day the contingency arises.
2. *New shift:* if the employee is not scheduled to work on the day of the contingency, a new shift is assigned for that day and if necessary, a

shift for a different day is eliminated.

3. *Reallocation of multiskilled employees:* for employees who belong to an area other than the affected one, there are three types of transfers: for the whole shift, for the period before lunch break and for the period after lunch break.

4.2.2. Search routine

The search for options is performed iteratively for each area and day affected. For each employee the movements are evaluated in combination with all of his or her shift possibilities for each day. In evaluating each option, the change in benefits is divided into two parts. The first part is the changes in the benefits and costs associated with the areas involved, computed in terms of the marginal benefits and the penalties for not meeting the minimum requirements in each period. The second part is the change in costs associated with the employee's weekly schedule and the type of recovery resource. All changes that have a positive net benefit are stored for later evaluation by the decision maker. Each step in the optimization routine is set out in Algorithm 4.

Algorithm 4: Local search optimization process

Phase 1: Initialization
Step 1: *Modifications* ← ∅, *Transfers* ← ∅, *Actions* ← ∅.
Step 2: For each employee, generate all possible shifts for each day.
Phase 2: Movement search
Step 3: For each day and area affected by a contingency, search for all the movements with a positive benefit and add them to *Actions*.
Step 4: If *Actions* = ∅ the algorithm terminates.
Step 5: Order *Actions* by the benefit increase obtained with each one.
Step 6: Present the options to the decision maker.
Step 7: Add the option chosen by the decision maker to the modifications and transfers lists and then return to *Step 3*. If no modifications or transfers are accepted, the algorithm terminates.

5. Experiments, results and discussion

In this section we describe the experiments conducted to measure the effectiveness of the solutions delivered by the two recovery system methods developed in the preceding section. Then we present the results of the implementation of both methods for a study case of a Chilean retailer. Measurements were made first for the iterative MIP system in different scenarios and were then compared to the greedy algorithm. The main indicators were the lost benefits recovered by the systems and the number of modifications made to achieve the recovery. Recall that, although contingencies can be attributed to demand variations and/or unplanned personnel absences, in this case study the experiments are conducted only for absenteeism.

5.1. Benefits function

The calculation of marginal benefits proceeds from two simple data sets in the possession of the retailer. The first is the approximate person-hour cost C_H, which includes employee wage and all other employee-related expenses. The second data set is the established customer service standards for each area, from which sales and transactions forecasts can be converted into the optimal number of person-hours h^* for each period of the day. This is done either by quantitative methods or on the basis of company experience. As for the retailer's utility function $B(h)$, it must satisfy the following basic conditions: (a) $B'(h) < 0$, $\forall\, h \leqslant h^*$, and (b) $B''(h) < 0$, $\forall\, h \geqslant 0$. The first condition states that it is always beneficial to have an additional employee as long as the optimum has not yet been reached while the second condition establishes that marginal benefits are decreasing.

In Cornejo (2007) the author attempted to determine which one of a group of simple functions would best fit the data for a certain company where experiments were to be performed. The planning tool developed in this study is now used by the company to schedule its work shifts. Based on these results, we chose the following function where $a > C_H$:

$$B(h) = \begin{cases} a\log(h) + b - C_H h & \text{if } h \geqslant 1, \\ (a-C_H)h + (b-a) & \text{otherwise.} \end{cases} \tag{23}$$

If, given the parameters of the company, h^* is an optimum, marginal benefits at this point are equal to marginal costs. Parameter a as a function of h^* is then:

$$\frac{dB}{dh}(h^*) = 0 \Rightarrow a = h^* C_H \tag{24}$$

Once the parameters are calculated we can calibrate the benefit ranges. Assuming that range r extends from $\omega_1 h^*$ to $\omega_2 h^*$, its utility for each person-hour is given by:

$$r_{atr} = \frac{B_{at}(\omega_2 h^*) - B_{at}(\omega_1 h^*)}{(\omega_2-\omega_1)h^*} \tag{25}$$

5.2. Experimental design

All of the experiments used instances based on real data from two stores belonging to a Chilean retailer. The processing times indicated here are thus the actual figures for any one of its stores. The demand forecasts, required for obtaining the benefits functions, were the same ones originally used to draw up the retailer's work schedules with a shift scheduling computer tool. These schedules are the best shift plan known to the company and therefore will serve as our benchmark.

In the order to evaluate and compare the results of both recovery system methods in different operating conditions, the problem was solved for two stores. Our purpose here is to compare both recovery system methods under various absenteeism scenarios and levels of priority for schedule planning stability. In each store, we used to 30 employees at the same 3 departments or areas and their actual restrictions. We consider four absenteeism scenarios: 0%, 5%, 10%, and 20%. For the last three levels tested it was assumed the chance an employee would not turn up for a planned shift was Bernoulli distributed with a probability of $p = 5\%$, 10%, and 20%, respectively. The recovery resources tested were the transfer multiskilled employees between store areas, the modification or elimination of existing shifts and the creation of new ones. Three recovery scenarios were considered: (1) Without recovery; (2) Without transfers; and (3) With transfers. The transfers can be only for whole shifts or the whole part of a shift before or after the lunch break. A sensitivity analysis was conducted on benefits in relation to the number of shift modifications using multiple cost levels for transfers and modifications in order to capture a variety of situations in which implementation of the modifications ranged from very easy to very difficult.

A total of 15 instances were simulated for each store and absenteeism level. Each instance was solved using the iterative MIP system algorithm implemented in C# on an ILOG CPLEX 12.4.0.1 optimization solver. Altogether, more than 1.000 different instances were solved on a workstation with a 3.0 GHz Core 2 Duo processor.

For the scenarios in which transfers were permitted, the areas were ordered heuristically from large to small according to the size of their benefit loss percentages due to contingencies. The column generation algorithm was configured to consider up to the 5 best schedules for each possible workday. The algorithm was stopped after 8 generation stages or when the maximum number of daily schedules considered had been reached without adding new columns. In each stage of the generation, a maximum of 30 schedules per employee was added. The solution retained was the best one delivered by the MIP model in the first 150 s; the remaining stages of the algorithm were not subject to any time limit.

Table 2
Best recovery averages for all instances.

Absenteeism	Store 1			Store 2		
(%)	Without recovery (%)	Without transfers (%)	With transfers (%)	Without recovery (%)	Without transfers (%)	With transfers (%)
0	100.00	100.00	100.29	100.00	100.00	101.67
5	97.00	97.65	97.92	97.32	97.89	100.35
10	93.67	94.76	95.16	91.73	93.30	95.97
20	88.15	89.82	90.46	86.88	88.82	91.29

To make the desired comparisons between both recovery system methods, all of the instances were run on the greedy algorithm with absenteeism levels of 5% and 10%.

5.3. Case study results

The results and discussion of this case study are divided into the following two subsections: Results with the iterative MIP system and Comparison of the two recovery system methods.

5.3.1. Results with the iterative MIP system

The averages for the best recovery achieved in the 15 instances are set out in Table 2 for each store and absenteeism level. The graphs in Fig. 1 show how the benefits of the best generated solution varied from the store's benchmark best benefit level (i.e., no absenteeism and with transfers). In each graph, the top curve indicates the recovery achieved with transfers, the middle curve the solutions without transfers and the bottom curve the benefits with neither transfers nor modifications.

The system recovery levels, expressed as the percentage of benefits lost due to the effect of a contingency on the store's original shift planning that was recovered, are displayed in Fig. 2. The percentage of original shifts that were modified is on the horizontal axis and the curves represent the different absenteeism levels. Note that these results refer only to expected benefits and penalties for employee deficits.

For both stores, these graphs reveal that the contribution of multiskilling was intimately related to a poor initial distribution of resources between departments. As absenteeism levels rose, the additional benefits derived from transferring multiskilled employees grew only slightly, the increase for Store 2 being superior at 1.67% to the Store 1 increase of just 0.29%. Without multiskilling the recovery capacity improved steadily as absenteeism worsened, closing the gap with solutions that included transfers. This occurred mainly because with more contingencies arising there were more and better opportunities for reassigning existing shifts. In both cases, maximum recovery capacity grew as contingencies increased in number. Also, as Fig. 2 shows, the benefit recovery percentage declined as absenteeism rose because the increase in net losses outweighed the increase in recovery capacity. The figure further reveals that benefit recovery exhibited decreasing returns to the number of modifications.

We also note that for both stores, maximum recovery was not achieved from the solution with the most modifications. Though this may seem contradictory, it must be kept in mind that the curves were constructed using solutions generated by assigning different penalty levels to the modifications rather than the best solution that could be obtained with a certain maximum number of modifications. Although the solutions found using different penalty levels were feasible in both cases, it was discovered that when the penalties were very small the column generation routine heuristics were unable to find the columns of the best solution. This was so because the heuristics had to explore a larger space and thus did not succeed in constructing weekly schedules that were among the best solutions. Furthermore, the MIP problem became more complex and the time limits imposed on the experiments in many cases prevented the model from finding the optimal solution. As was explained Hur et al. (2004b), there clearly exists a certain number of modifications beyond which no further benefits are obtained. In our case the best solution was always found when the number of modified shifts was no more than 40%, counting transfers as shift modifications.

In general terms, without multiskilling the loss recovery levels stabilized when 15% to 25% of the shifts were modified, loss recovery itself ranging between 13% and 23% depending on the absenteeism level. With multiskilling, however, the results for the two stores were very different. At Store 1, with approximately 30% of shifts modified the recovery levels oscillated between 18% and 30% while at Store 2, recovery stabilized with about the same proportion of modifications but recovery levels reached 100% with 5% absenteeism and in the other cases stayed above 30%. Thus, results show that in the worst scenario, the proposed model's schedule adjustments reduced lost profits due to unexpected variations by 18%.

5.3.2. Comparison of the two recovery system methods

Recovery capacity for both methods with 10% absenteeism is shown in Fig. 3. As can be seen, the greedy algorithm was outperformed in every case by the iterative MIP system. This occurred because given the way the greedy algorithm constructs its solutions, the number of modifications made is always fewer than with MIP. Table 3, which shows the averages for solution times and objective function value loss recovery (the latter after deduction of modification costs), confirms that

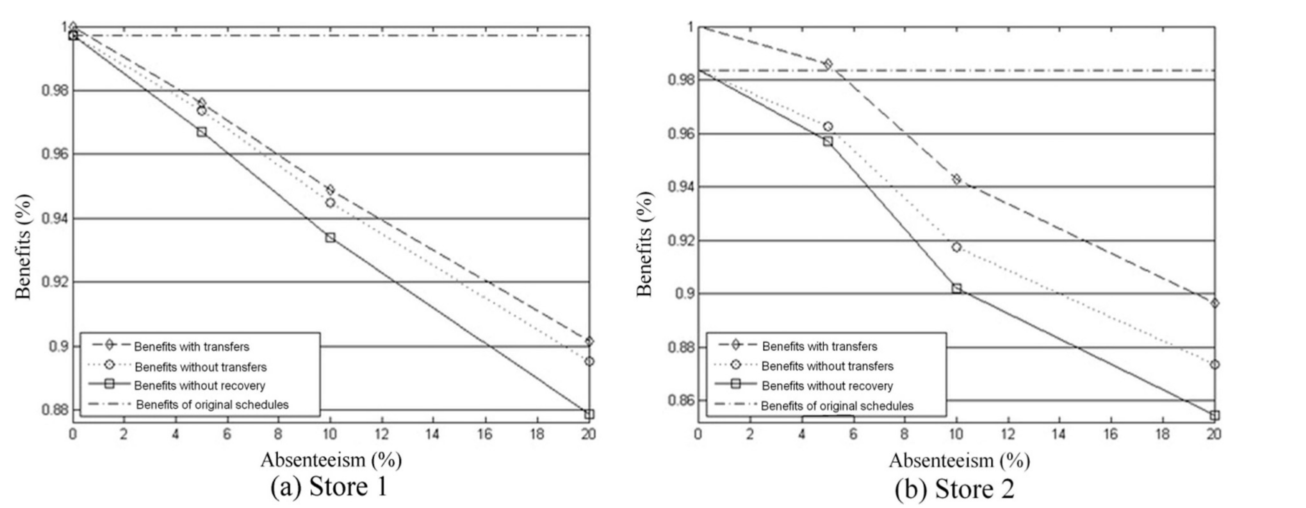

Fig. 1. Recovery capacity for different absenteeism levels.

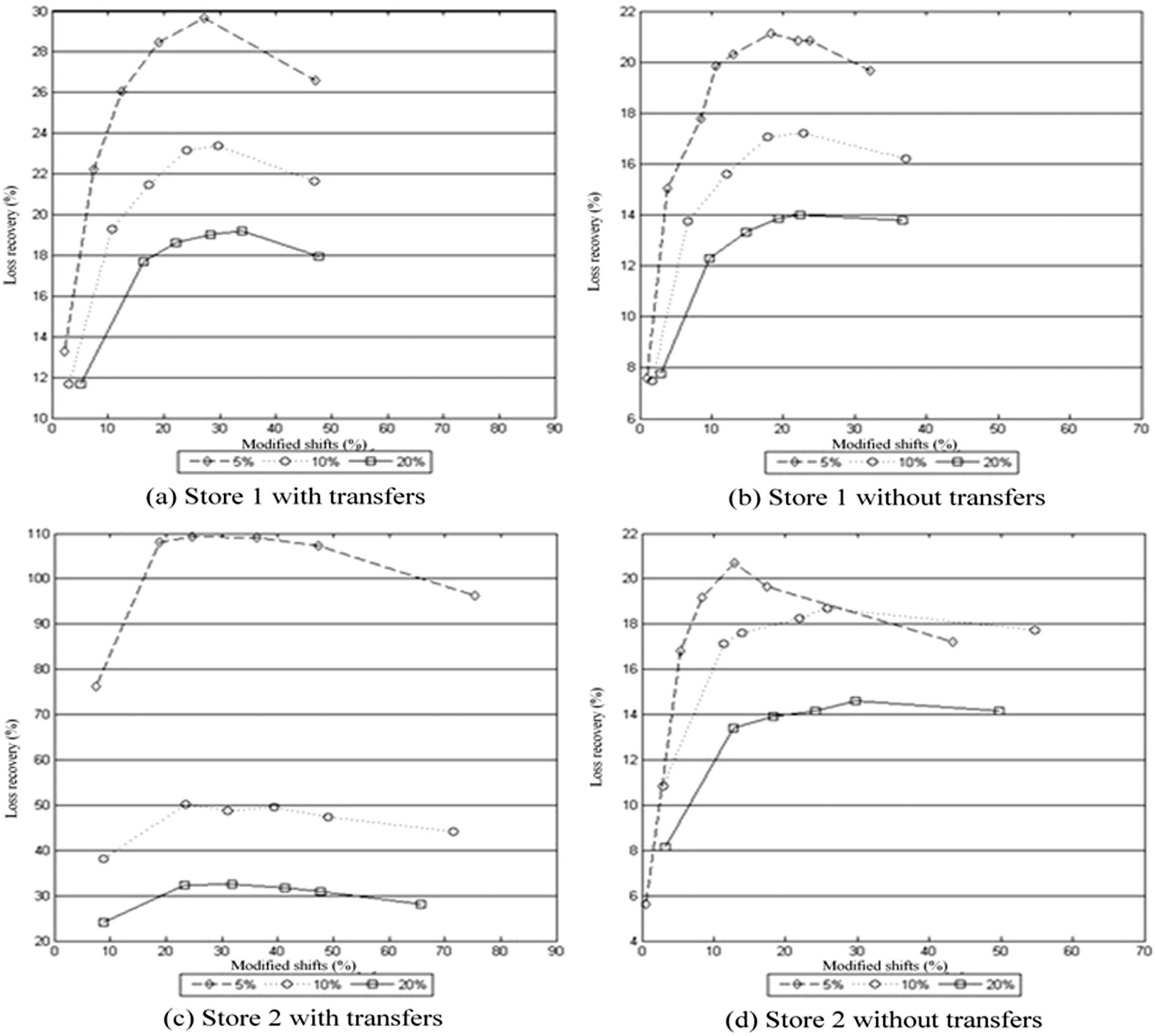

Fig. 2. Recovery percentage by number of modifications and absenteeism level.

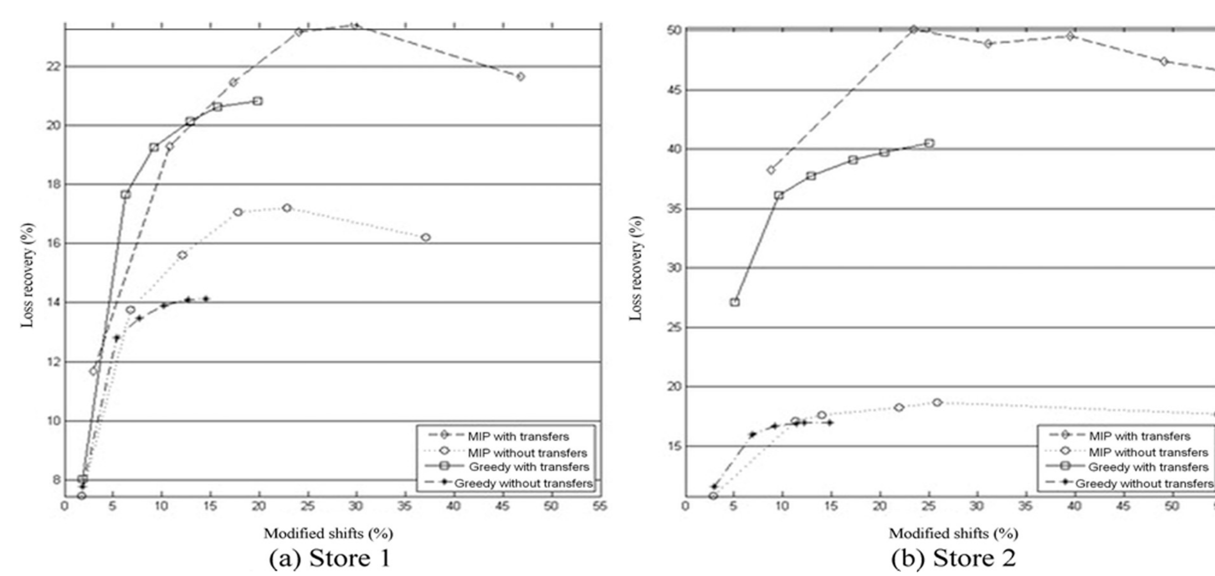

Fig. 3. Performance of the two algorithms with 10% absenteeism.

Table 3
Average solution times and objective function (OF) value loss recovery for both algorithms.

Store	Absenteeism (%)	Time (s)		OF value recovery (%)	
		MIP	Greedy	MIP	Greedy
Store 1	5	464.92	0.49	22.95	19.41
	10	452.80	2.47	19.39	17.21
Store 2	5	680.45	0.23	122.41	33.58
	10	606.86	1.75	46.90	35.56

although the iterative MIP system was much slower, its results were far superior. Also note that, none of the methods were able to find a better solution without resorting to multiskilled employees.

6. Conclusions and future research

Two methods for solving the contingency problem in workforce scheduling were developed and evaluated. Both make use of all the standard recovery resources identified in this study. The first method is built around an iterative algorithm that solves a relatively large mixed integer programming (MIP) model by column generation. The second method takes a more practical approach, using a greedy algorithm to apply the most beneficial shift modifications in response to contingencies identified previously. Experiments were carried out to compare the results of the two methods using real data from two stores belonging to a Chilean retail chain.

The results of the experiments demonstrated that much of the benefits lost due to unexpected events could be recovered using a contingency control tool. The transfer of multiskilled employees to store areas affected by a contingency proved to be an excellent recovery tool in cases where there was a potential benefit not exploited in the original workforce planning. In every case, such transfers mitigated the impact of the contingency on the various store departments and achieved equal or better recovery levels with few modifications to existing shifts.

Comparisons of the two methods indicated that if the same cost was attributed to all possible modifications, the most obvious actions represented by the greedy algorithm led to solutions with lower recovery levels than those of the iterative MIP system. Even when greedy was two orders of magnitude faster than MIP, the latter solved every instance in a reasonable length of time. In real-world situations the possibilities of an employee are a mere fraction of those that were assumed in the experiments, suggesting that the range of problems solvable with the iterative MIP system would include much bigger cases than the one used here. Since in practice a solution involving modification of more than 40% of the shifts is unlikely to be implementable, the heuristics performed well over real operating ranges.

The proposed iterative MIP system is thus an effective tool for contingency control and recovery problems, using all of the standard shift scheduling recovery resources and integrating optimization with a process in which individual actions must be validated and selected for application by a decision maker. Since such interaction is necessary between the decision maker and the affected employees, reducing the time required by the iterative algorithm to find and present a series of shift modifications is of prime importance. Four possible approaches to reducing solution times involve implementing a branch and price methodology for the MIP problem, generation heuristics, the treatment of multiskilling, and the use of initial solutions. The first one, would be a hybrid method where the column generation is coupled with a branch-and-bound method. On the second approach, weekly shift schedule heuristics could be improved to find good solutions more quickly. Including randomness in the generation process should also be explored as a way of increasing variety in the shifts selected. A further possibility would be to search in larger spaces during generation with heuristics based on constraint programming using techniques of the sort described in Demassey, Pesant, and Rousseau (2005) and Demassey, Pesant, and Rousseau (2006).

As regards multiskilling, this adds much complexity to the MIP problem. To speed execution, this factor could be treated during the column generating phase by incorporating the search for good employee transfers into the schedule selection process. Finally, using solutions found by faster algorithms such as the greedy routine could also help in lowering solution times.

As was noted earlier, when the solutions are restricted to those with a limited number of modifications, the greedy algorithm performs similarly to the iterative MIP system. If we assume that in practice only such solutions are feasible, improving solution quality using the greedy algorithm may be a good alternative. One possible enhancement would be to include randomness in the choice of modifications during construction of the solution, thus generating multiple solutions during the search and thereby increasing the chance of finding a good one. Also, the algorithm's fast solution times together with the properties of the problem could be exploited to explore actions that involve more than one modification at a time, choosing the best series of modifications at each step.

Finally, the user interface is a key factor if it intends to achieve a successful real implementation of the proposed system. Thus, under the implementation of our real-time system for adjusting personnel schedules, decision makers must make improvements on their actual user interface of the system. Such improvements should allow an effective and fast interaction between the decision maker and the affected employees.

Acknowledgements

The authors would like to thank the anonymous referees for their helpful comments and insights on this paper. They are also grateful to their colleagues at SHIFT SpA for providing the data used to solve the problem for the retail firm.

References

Alfares, H. (2004). Survey, categorization, and comparison of recent tour scheduling literature. *Annals of Operations Research, 127*(1–4), 145–175. http://dx.doi.org/10.1023/B:ANOR.0000019088.98647.e2.

Bard, J., & Wan, L. (2005). Weekly scheduling in the service industry: An application to mail processing and distribution centers. *IIE Transactions, 37*(5), 379–396. http://dx.doi.org/10.1080/07408170590885288.

Campbell, G. (2011). A two-stage stochastic program for scheduling and allocating cross-trained workers. *Journal of the Operational Research Society, 62*(6), 1038–1047. http://dx.doi.org/10.1057/jors.2010.16.

Cornejo, G. (2007). *Formulación e implementación de un modelo de programación lineal entera para la asignación de turnos a empleados con jornada de trabajo variable (Formulation and implementation of an integer linear programming model for assigning shifts to employees with variable workday).* Unpublished master's thesisPontificia Universidad Católica de Chile.

Cuevas, R., Ferrer, J. C., Klapp, M., & Muñoz, J. C. (2016). A mixed integer programming approach to multi-skilled workforce scheduling. *Journal of Scheduling, 19*(1), 91–106. http://dx.doi.org/10.1007/s10951-015-0450-0.

Demassey, S., Pesant, G., & Rousseau, L.-M. (2005). Constraint programming based column generation for employee timetabling. *Proc. 2th int. Conf. on integration of AI and OR techniques in constraint programming for combinatorial optimization problems – CPAIOR-05* (pp. 140–154). Springer-Verlag LNCS 3524.

Demassey, S., Pesant, G., & Rousseau, L.-M. (2006). A cost-regular based hybrid column generation approach. *Constraints, 11*(4), 315–333. http://dx.doi.org/10.1007/s10601-006-9003-7.

Easton, F. F. (2011). Cross-training performance in flexible labor scheduling environments. *IIE Transactions, 43*(8), 589–603. http://dx.doi.org/10.1080/0740817X.2010.550906.

Easton, F., & Goodale, J. (2002). *Labor scheduling with employee turnover and absenteeism.* Technical reportSyracuse University.

Easton, F., & Goodale, J. (2005). Schedule recovery: Unplanned absences in service operations. *Decision Sciences, 36*(3), 459–488. http://dx.doi.org/10.1111/j.1540-5414.2005.00080.x.

Easton, F., & Rossin, D. (1991). Sufficient working subsets for the tour scheduling problem. *Management Science, 37*(11), 1441–1451. http://dx.doi.org/10.1287/mnsc.37.11.1441.

Ernst, A., Jiang, H., Krishnamoorthy, M., Owens, B., & Sier, D. (2004). An annotated

bibliography of personnel scheduling and rostering. *Annals of Operations Research, 127*(1–4), 21–144. http://dx.doi.org/10.1023/B:ANOR.0000019087.46656.e2.
Ernst, A., Jiang, H., Krishnamoorthy, M., & Sier, D. (2004). Staff scheduling and rostering: A review of applications, methods and models. *European Journal of Operational Research, 153*(1), 3–27. http://dx.doi.org/10.1016/S0377-2217(03)00095-X.
Henao, C. A., Ferrer, J. C., Muñoz, J. C., & Vera, J. (2016). Multiskilling with closed chains in a service industry: A robust optimization approach. *International Journal of Production Economics, 179*, 166–178. http://dx.doi.org/10.1016/j.ijpe.2016.06.013.
Henao, C. A., Muñoz, J. C., & Ferrer, J. C. (2015). The impact of multi-skilling on personnel scheduling in the service sector: A retail industry case. *Journal of the Operational Research Society, 66*(12), 1949–1959. http://dx.doi.org/10.1057/jors.2015.9.
Hur, D., Mabert, V., & Bretthauer, K. (2004a). Real-time schedule adjustment decisions: A case study. *Omega, 32*(5), 333–344. http://dx.doi.org/10.1016/j.omega.2004.01.002.
Hur, D., Mabert, V., & Bretthauer, K. (2004b). Real-time work schedule adjustment decisions: An investigation and evaluation. *Production and Operations Management, 13*(4), 322–339. http://dx.doi.org/10.1111/j.1937-5956.2004.tb00221.x.
Kabak, Ö., Ülengin, F., Aktas, E., Önsel, S., & Topcu, Y. I. (2008). Efficient shift scheduling in the retail sector through two-stage optimization. *European Journal of Operational Research, 184*(1), 76–90. http://dx.doi.org/10.1016/j.ejor.2006.10.039.
Menezes, M., Kim, S., & Huang, R. (2006). Optimal workforce size and allocation for urban retail chains. *European Journal of Operational Research, 175*(2), 1151–1163. http://dx.doi.org/10.1016/j.ejor.2005.06.038.
Milner, J., & Pinker, E. (2001). Contingent labor contracting under demand and supply uncertainty. *Management Science, 47*(8), 1046–1062. http://dx.doi.org/10.1287/mnsc.47.8.1046.10233.
Nissen, V., & Günther, M. (2010). Automatic generation of optimised working time models in personnel planning. *7th international conference on swarm intelligence, Brussels, Belgium, lecture notes in computer science 6234* (pp. 384–391). .
Orsoni, A. (2004). GAs and simulation techniques for dynamic resources sharing and reallocation across workgroups. *Second IEEE international conference on intelligent systems, June 2004*.
Pinker, E., & Larson, R. (2003). Optimizing the use of contingent labor when demand is uncertain. *European Journal of Operational Research, 144*(1), 39–55. http://dx.doi.org/10.1016/S0377-2217(01)00378-2.
Van den Bergh, J., Belien, J., De Bruecker, P., Demeulemeester, E., & De Boeck, L. (2013). Personnel scheduling: A literature review. *European Journal of Operational Research, 226*(3), 367–385. http://dx.doi.org/10.1016/j.ejor.2012.11.029.
Zhang, X., Chakravarthy, A., & Gu, Q. (2009). Equipment scheduling problem under disruptions in mail processing and distribution centres. *Journal of the Operational Research Society, 60*(5), 598–610. http://dx.doi.org/10.1057/palgrave.jors.2602581.