Accident Analysis and Prevention 122 (2019) 25–35

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Reducing intercity bus crashes through driver rescheduling

Shuo-Yen Wang, Kun-Feng Wu*

Department of Transportation and Logistics Management, National Chiao Tung University, Taiwan, ROC

ARTICLE INFO

Keywords:
Intercity bus safety
Driver scheduling
Case-study approach

ABSTRACT

Intercity bus crashes often involve driver fatigue, which itself is usually the result of sleep deprivation, long driving hours, a maladjusted circadian rhythm, or some combination of the above. And driver scheduling has long been suspected as the root cause affecting sleepiness and fatigue. As such, a fundamental question for intercity bus carriers is how to reduce crashes associated with driver schedules, while maintaining a nonstop service? This research seeks to develop a paradigm to minimize overall fleet crash risk by rescheduling. In this study, we first identified those driving schedules associated with the highest crash risks, and a rescheduling scheme is then proposed to reduce fleet crashes overall. A case-study approach was employed to identify driver scheduling associated with higher crash risk, and a mathematical program was then formulated to minimize fleet crash risk. Our results showed that several types of driver schedules would lead to higher crash risk; for example: (1) working in the afternoon or early hours in the morning for two consecutive days; and (2) commencing a driving shift in the mornings, the afternoon or the early hours of the morning after being off-duty for more than 24 h. To meet the challenge of maintaining a nonstop service while simultaneously minimizing the crash risk associated with these risk patterns, a mathematical program was developed, and it was found that rescheduling based on our algorithm could reduce the incidence of crashes by approximately 30 percent.

1. Introduction

Driver fatigue and sleepiness have been found to be closely correlated with crash risk; indeed, they are often cited as the major crash contributing factor for intercity bus drivers who are often required to drive long hours and do night-shift driving. More and more research has demonstrated that a driver's shift schedule (referred to as working pattern in the rest of this study) is associated with driver fatigue, and hence crash risk (e.g.Jovanis et al., 2012). The reasons why driver working pattern is associated with driver fatigue and sleepiness include sleep deprivation, long driving hours, and circadian rhythm (Pack et al., 1995; Feyer and Williamson, 1995; Williamson et al., 1996; Williamson et al., 2011). Hence, a well-designed driver schedule was found to be beneficial in lowering crash risk (e.g. Shinar, 2007; Crum et al., 2001; Crum and Morrow, 2002). A well-designed schedule will obviate the need for overlong working or irregular working patterns, and reduce the likelihood of driver fatigue and sleepiness.

Nevertheless, despite the obvious benefits, there have been limited studies into how to manage intercity bus drivers' crash risk through proper scheduling. Clearly, the greatest constraint on intercity bus carriers is their need to maintain a nonstop service 24 h a day, seven days a week. In other words, although there may be some time slots associated with a higher crash risk (e.g. 2 AM to 4 AM) the bus carriers have no choice but to assign drivers to run a service at these times in order to operate a 24/7 timetable. Moreover, as crash risk is determined not only by the driving environment but also by the time of day (or night), it is challenging to quantify crash risk for different working patterns.

1.1. Driver scheduling patterns and crash risk

Past research suggested that fatigue would lower drivers' overall level of attention, or disrupt matching of effort to task demand, and hence result in slower information processing and higher crash risk (e.g. Kahneman, 1973; Desmond and Matthews, 1997). Many factors related to driver scheduling have been argued to be associated with driver fatigue, including sleep deprivation, time-on-task/distance driven, and circadian rhythm (Shinar, 2007). The effects of sleep deprivation and time-on-task/distance driven on fatigue and fatigue-related crashes are straightforward and have been well-established to some extent (e.g. Rogé et al., 2003; Hulst et al., 2001). Studies on circadian rhythm indicate that most people experience a slight drop in their wakefulness from approximately 2 to 4PM and a significant drop from approximately 2 to 6 AM. In fact many drivers report they are more likely to

* Corresponding author.
E-mail addresses: jack2pp12345@gmail.com (S.-Y. Wang), kfwu@nctu.edu.tw (K.-F. Wu).

https://doi.org/10.1016/j.aap.2018.09.019
Received 23 April 2018; Received in revised form 20 August 2018; Accepted 18 September 2018
Available online 07 October 2018
0001-4575/

experience fatigue during these hours than during other hours (Feyer and Williamson, 1995).Wylie et al., (1996), who used alertness tests and instrumented truck measures rather than crashes, found a strong correlation between fatigue and time of day, and little correlation between fatigue and driving hours. Williamson et al. (2011) found that the speeds of driver reaction time, divided attention, short-term memory, and vigilance vary according to the drivers' circadian rhythm. Pack et al. (1995) reported that for all age groups, peak crash involvement is in the two time periods of the circadian rhythm. In particular, older drivers have more difficulty in altering the cycle, and they are more prone to have single-vehicle crashes in early afternoons. In summary, circadian rhythm is likely to play an important role affecting driver fatigue and crash risk.

Although much research has associated increased crash odds (or relative risk) with total hours spent driving, particularly when exceeding five to six hours of driving (e.g. Jovanis et al., 2012; Soccolich et al., 2013; Sparrowa et al., 2016), some studies have highlighted driver scheduling patterns and multi-day driving patterns as also correlated with higher crash risk (e.g. Kaneko and Jovanis, 1992; Jovanis et al., 2012). Kaneko and Jovanis (1992) and Jovanis et al. (2012) used data logs from less-than-truckload carrier operations to estimate the probability of a crash after a certain amount of time spent driving. Driver logs for 7 days before each crash were used and compared with a random sample (two drivers) of drivers who did not crash and were selected from the same company, terminal, and month. Certain multi-day driving patterns were reported to have a higher crash risk, especially for those returning from extended periods off duty (Jovanis et al., 2012). Not surprisingly, research also showed that breaks from a particular driving task, defined as a period within a driving trip when the driver was off duty or took a nap, are beneficial in lowering crash risk (Akerstedt and Gillberg, 1990; Bonnefond et al., 2004; Jovanis et al., 2012; Soccolich et al., 2013; Chen and Xie, 2014; Torregroza-Vargas et al., 2014; Sparrowa et al., 2016). Taken as a whole, ad hoc driver scheduling patterns, referred to as driving patterns in the rest of this study, are likely to be associated with the onset of fatigue while driving, and hence of a higher crash risk.

1.2. Bus driver scheduling and traffic safety management

Current practice regarding bus driver scheduling and safety management focuses mostly on meeting hour-of-service regulations, national and local labor union rules, the balance of work load among drivers, number of days off work, and operational costs (Torrance et al., 2009), but little thought is given to the impact of multi-day driving patterns on overall crash risk. To meet these requirements while minimizing associated operating costs, most research on driver scheduling was conducted in the operations research field (e.g. Wren and Rousseau, 1995; Banihashemi et al., 2000; de Matta and Peters, 2009; Torrance et al., 2009). Although driver and vehicle scheduling are related, vehicle and driver scheduling problems were dealt with separately in the past due to their complexity. With increased computing power, current research is now able to account for both of them at the same time (Fischetti and Matteo, 2001).

For bus driver scheduling, the objective is to minimize the total costs associated with drivers while considering contractual constraints. The total costs associated with drivers generally include a driver's salary, compensation package, and overhead costs (Torrance et al., 2009). The contractual constraints are specified by bus carriers, including, but not limited to, national and local labor rules; total time worked per day, total time worked per week/month, and/or number of days off per week/month, etc. (Wren and Rousseau, 1995; Yaoyuenyong and Nanthavani, 2005; Al-Yakoob and Sherali, 2007).

Although bus driver scheduling should be based on minimizing both operational costs and crash risks, most research focuses on minimizing operational costs under existing contractual constraints. However, from a traffic safety management perspective, minimizing crash risks should be considered part of the objective functions rather than part of the contractual constraints, as those constraints generally do not promise minimum crash risk (e.g. Jovanis et al., 2012). More research is needed to account for the effects of driver scheduling on crash risk.

1.3. Research objectives

This study seeks to first identify bus driver scheduling patterns that are associated with higher crash risk, and to use the information identified to minimize a fleet's overall crash risk through rescheduling. The idea of minimizing a fleet's overall crash risk by rescheduling is illustrated in Fig. 1. In Fig. 1, the number in each node indicates a given driver's working pattern for a given day. A working pattern may be composed of multiple shifts in a day. As an example, suppose an intercity bus driver only needs to drive one round trip per day. Driver A's working pattern on the first day is "Pattern 1" (say he leaves at 8 A M,

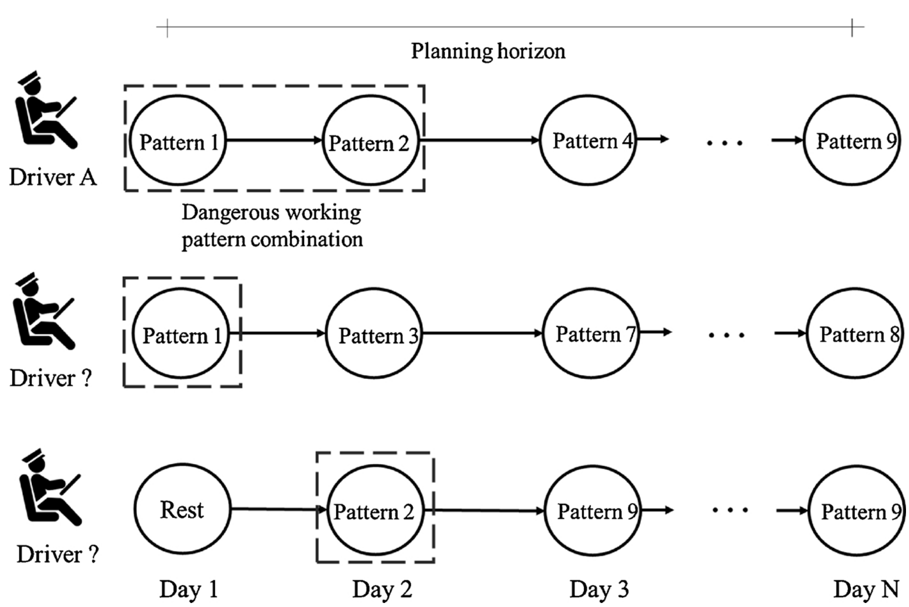

Fig. 1. Illustration of rescheduling.

and returns at 2 PM), and "Pattern 2″ (say he leaves at 12 PM, and returns at 6 PM) on the second day, giving a combination of working patterns, Patterns "1″ and "2″ for the first two days. Hypothetically, when "combination 1-2″ is identified to be associated with a high crash risk, rescheduling can then be carried out to split Pattern 1 on the first day and Pattern 2 on the second day between different drivers to avoid "combination 1-2″, and thus reduce crash risk. This way of rescheduling is not only simple and intuitive, but flexible in maintaining the same operational timetable.

2. Methodology

The object of this study is to identify combinations of daily working patterns that are associated with higher crash risk. Identifying combinations associated with higher crash risk is more valuable and useful than identifying individual working patterns associated with higher crash risk. For example, even if we find that driving in the middle of the night is associated with high crash risk, this information is unlikely to lead us to "avoid driving in the middle of the night." Instead, if we take a driver's working patterns before he drives the night shift into account, we may be able to find a combination in which although drivers still need to drive through the middle of the night, they will experience a lower crash risk when this midnight shift is combined with another working pattern the previous day (e.g. rest a day before driving at midnight). This is particularly useful for intercity bus companies wishing to provide 24/7 service.

This study seeks to first quantify the crash risk for each combination of working patterns, and then minimize a fleet's overall crash risk by rescheduling. Three tasks are required: (1) analyzing and grouping drivers' working patterns, (2) quantifying the crash risk of the working patterns, and (3) rescheduling through an optimization model.

2.1. Step1: group combinations of working patterns

Since there are many possible driving patterns over multiple days one has the challenge of identifying combinations of working patterns associated with higher crash risk. Cluster analysis has been successfully used to group combinations of working patterns into relatively consistent multi-day driving patterns for manageable statistical analysis in previous studies (e.g. Jovanis et al., 2012).

This clustering is determined by the time of a driver's first shift of the day. The goal is to divide the planning horizon into several "units" (e.g. a 30-day schedule can be divided into 30, 15, or 10 units, according to the length of a unit). We use 3 units to form combinations; units are 1 day, 2 days and 3 days, and the combinations are "1 day-1 day combinations", "2 day-2 day combinations" and "3 day-3 day combinations". We assume that time before a combination is independent, so a 1 day-1 day combination can only observe a driver's working pattern over 2 days. Using longer units allows us to observe drivers working on more days and rescheduling with the same optimization model. The units should be in pairs in order to successfully apply the optimization model.

2.2. Step2: quantify crash risk of working pattern combinations

A logistic regression model is first employed to identify and quantify the crash risk of each combination of working patterns. As there are some unobserved factors that have effects on crash risk, but are not controlled for in this study, a case-control study is utilized to further verify the significance of the effects of those combinations on crash risk. In each case-control study, cases were trips involving a crash, and controls were trips not involving crashes. The controls were matched from previous trips of those drivers who had trips involving a crash (Jovanis et al., 2012). As the controls were randomly matched from previous trips of those drivers who had trips involving a crash, a bootstrapping method was applied to enhance the robustness of the results (Cassidy et al., 2009; Laumon et al., 2005). The case-control study combined with the bootstrapping method is described below:

- Step 1: Number each trip in the case group.
- Step 2: For each trip in the case group, randomly select N trips in the control group with the same driver, and make these N + 1 trips as panel data.
- Step 3: Employ a fixed-effect logit regression model with panel data. The independent variables are whether this trip is taken as which combination such as combination *1*, combination *2*..., and combination *k*. The dependent variable is whether there was a crash or not. Recording the coefficients of independent variables, $\beta_1, \beta_2, \ldots, \beta_k$, this set is called T^b.
- Step 4: Put randomly selected trips back into the control group.
- Step 5: Repeat Step 2 ~ Step4 B times, and we get $T^1, T^2, \ldots, T^B$.
- Step 6: Calculate $p = \left\lfloor \frac{\alpha}{2}(B+1) \right\rfloor$ *and* $q = \left\lfloor (1-\frac{\alpha}{2})(B+1) \right\rfloor$, α is the cutoff for significance. Make $k \longleftarrow 1$.
- Step 7: Sort β_k of all $T^1, T^2, \ldots, T^B$. Take p th β_k as lower bound of confidence interval and q th β_n as upper bound of confidence interval to test the following hypothesis:

$$H_0: \beta_k = 0 \tag{1}$$

$$H_1: \beta_k > 0 \tag{2}$$

- Record the result, and make $k \longleftarrow k+1$. Repeat Step 7 until all β_k' s hypotheses are tested.

2.3. Step3: minimize fleet's crash risk by rescheduling

The combination of working patterns of drivers can be expressed by a network shown in Fig. 2. The optimization is solved as a minimum cost flow problem. The minimum cost flow problem is an optimization and decision problem to find the cheapest possible way of sending a certain amount of flow through a flow network. The amount of flow is equal to total drivers, nodes in this network mean different working patterns in units, arcs connecting two nodes representing this schedule in two units are acceptable (regarding local work-hour regulation), and the flow on an arc representing the quantity of drivers are arranged as nodes linked together. The cost of sending a flow across the arc is the probability of crash calculated by the previous logit regression model. It should be noted that only those combinations passing through the case-control study would be assigned a cost. The optimization model based on this network minimizes the expected number of crashes. So the solution to this minimum cost flow problem is the safest scheduling for drivers. And in this problem we also consider remaining the sum of drivers, and the departure schedule should remain the same. The mathematical model is as below:

2.3.1. Sets

H :Set of *unit* s in the time horizon. $H = \{1, 2, \ldots, |H|\}$
P: Set of working patterns in each unit. $P = \{1, 2, \ldots, |P|\}$
Q: Set of time periods in a *unit*. $Q = \{1, 2, \ldots, |Q|\}$

V^S: *Set of start and end nodes.* $V^H = \{0, |H||P| + 1\}$

V^H: *Set of all nodes without start and end node in planning horizon*

V: *Set of all nodes in the shift network.* $V = V^H \cup V^S$

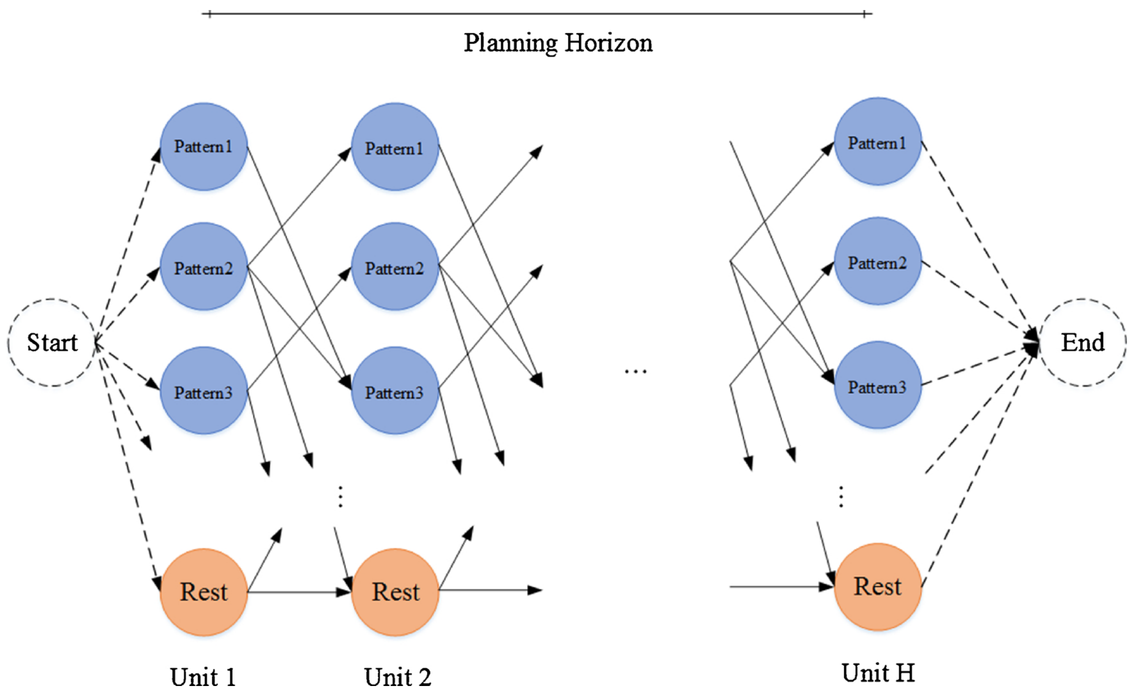

Fig. 2. Directed Network for Rescheduling.

E^S: *Set of arcs linking source or sink node to other nodes.*

$E^S = \{(0 \times V^1) \cup (V^H \times |H| \bullet |P| + 1)\}$

E^F: *Set of arcs representing working schedule break local work −hour regulation .*

E^H: *set of arcs which represent possible working schedule in two consecutive unit.*

$E^H = \{V^h \times V^{h+1}\}\backslash E^F,\ h \in \{1,2,\ \ldots, |H|-1\}$

E: *Set of all arcs in the network*

$E = E^H \cup E^S$

δ_u^+: *Set of arcs which link out from node u.*

$\delta_u^+ = \{(u,\ v) \in E,\ v \in V\}$

δ_u^-: *Set of arcs which links to node u.*

$\delta_u^- = \{(v,\ u) \in E,\ v \in V\}$

2.3.2. Parameter

N: *Number of total drivers*

D_{hq}: *Number of departures desired in h th unit , time period q,* $h \in H,\ q \in Q$

A_{uq}: *The working pattern of node u with departure in time period q,* $u \in V^H,\ q \in Q$

r_e: *Probability of crash when scheduling as arc e,*can be obtained from logit regression model of previous step, $e \in E^H$

2.3.3. Decision variable

x_e: *Volumn on arc e, represeting how many drivers are scheduled as these two nodes connected by arc e.* $e \in E,\ x_e \geq 0\ ,\ x_e \in Integer$

2.3.4. Mathematical model

$$Min \sum_{e \in E^H} r_e x_e \tag{3}$$

s. t.

$$\sum_{e \in \delta_0^+} x_e = N \tag{4}$$

$$\sum_{e \in \delta^-_{|P| \bullet |H|+1}} x_e = N \tag{5}$$

$$\sum_{e \in \delta_u^-} x_e - \sum_{e \in \delta_u^+} x_e = 0\ \ \forall\ \ u \in V^H \tag{6}$$

$$\sum_{u \in V^H} \sum_{e \in \delta_u^-} A_{up} \bullet x_e = D_{hp}\ \ \forall\ \ h \in H,\ p \in P \tag{7}$$

$$x_e \geq 0,\ x_e \in integer,\ e \in E \tag{8}$$

The objective function (3) minimizes the sum of probability of crash, i.e. it minimizes crash risk of all drivers. Constraints (4)-(5) limit the drivers in this network equal to N, and the flow begins at source node, and ends at sink node. Constraint (6) ensures flow conservation in this network. Constraint (7) ensures the new schedule must meet the bus departure timetable. And finally constraint (8) ensures each arc corresponds to a non-negative integer, which represents the volume on the arc.

Table 1
Types of crashes included in this study.

Type of crash	Number of crashes
Rear-end collision	13
Fender-bender	62
Crashes while parking	48
Total	123

Table 2
Crash frequency for the drivers included in this study.

Crashes caused by same driver	Number of drivers
5	1
4	1
3	3
2	19
1	67
0	219
Total	310

3. Data description

3.1. The data

The study data was provided by an intercity bus carrier, and includes one year of bus departure information (departure time, arrival time, driver's name and involving a crash or not) for a route. This study includes a total number of 129,009 trips, involving 123 crashes and 310 drivers. The average driving time of this route is about 4 h. Although the driving time of each trip depends on traffic conditions, the variation is small. There is no rest time provided for drivers during this trip. An identity verification and alcohol test is required by law and is conducted before the departure of each trip to ensure the driver is capable and responsible for driving duty.

The details of crash data are shown in Tables 1 and 2. The types of crashes included rear-end collision, fender-bender, and crashes while parking (e.g. scratching another vehicle while reversing). The crashes included were at-fault crashes, crashes due to drivers' negligence or fault (judged by the police and the bus company after they reviewed the video clips at the scene and interviewed the drivers involved). It is hypothesized that these at-fault crashes are more likely to be associated with driver fatigue or drowsiness, and hence are more likely to be associated with driver scheduling. Table 2 shows the crash frequency of the drivers included in this study. 70 percent of drivers were not involved in a crash, and only five drivers had three or more crashes in a year.

3.2. Driving patterns

To comply with a pre-determined master schedule, a bus carrier would often need to first determine a weekly, bi-weekly, or monthly run package (planning horizon), and find individual drivers to fulfill each shift (Torrance et al., 2009). A driver's working pattern may be composed of multiple shifts in a day (or a number of days). A multi-day working pattern is referred to as a "combination" of working patterns (or simply as driving patterns).

Since there are many possible driving patterns over multiple days, one has the challenge of identifying combinations of working patterns associated with higher crash risk. Cluster analysis has been successfully used to group combinations of working patterns into relatively consistent multi-day driving patterns for manageable statistical analysis in previous studies (e.g. Jovanis et al., 2012). The clustering will be first conducted on the basis of working patterns, and the optimization discussed in the previous section will be based on the combinations of working patterns. The idea of reducing a fleet's crash risk is to replace one combination related to high crash risk with another two combinations related to lower crash risk (as discussed in the hypothetical example in Fig. 1, when "combination 1-2″ is identified to be associated with a high crash risk, rescheduling can then be carried out to split the combination with another two combinations 3-2 and 1–4 by different drivers or the same driver but on different days).

To be able to execute the optimization, a planning horizon (e.g. driver's 30-day schedule) can be broken down into a number of combinations, each combination including a number of working patterns, depending on the length of a "unit" (a day or a number of days). Sensitivity analyses were carried out to test the impact of length-of-planning horizons, combinations, and length-of-units on the results of optimization. Three different units were tested to form combinations, and they are 1, 2, and 3 days. The planning horizons are 6, 14, 15, and 30 days, corresponding to commonly used one week, two weeks, and one month planning horizons in practice by intercity bus carriers. Without loss of generality, the combinations considered in this study were "1 day-1 day combinations", "2 day-2 day combinations" and "3 day-3 day combinations" (1-1 combination, 2-2 combination 3-3 combination in short).

Table 3 presents all working patterns in one day (the unit is one day). For example, Pattern 12 indicates a driver's first trip departs between 9:00 and 12:00, and then returns between 15:00 and 18:00. This type of working pattern accounted for 15.22 percent of all driving patterns in one day. It should be noted that we used three hours to categorize shifts because (1) as the average driving hours are about four hours, a time window longer than four hours may result in two shifts being undertaken but only one shift being recorded; and (2) a shorter time window would lead to a greater number of working patterns, and hence become more difficult to cluster.

Fig. 3 shows an example of a driver's driving schedule in a six-day planning horizon. There are six working patterns in these six days. The solid circles (first row) indicate the driver's shift in each day, along with the exact departure times for each trip. The second row hollow circles indicate the numbering of the working patterns. In this example, the driver departed at 9:00 for the first trip, and returned at 15:30 for the second trip on day one, labeled as Pattern 12. The working pattern on day two is labeled as Pattern 15, etc.

A longer length of units would allow us to observe longer driving patterns, but the longer the unit, the more difficult it is to cluster driving patterns. As an example, if we use two days as a unit and if the time window is still three hours, the number of types of working patterns will be over 1000, and hence leads to a small sample size in many working patterns. The time window is specified as six and eight hours for the units of two and three days, respectively, to capture the departure time of a driver's first trip in a day (for this intercity bus carrier, a driver is asked to drive a roundtrip, and each trip is usually four hours of driving). The results of clustering for the units of two and three days are shown in the Tables A1 and A2 in Appendix.

4. Results

In addition to lowering crash risk, intercity bus carriers are also interested in reducing operational costs and avoiding unequal workloads among drivers, so three performance measures were utilized in this study, as follows:

Table 3
Working patterns in terms of a unit of one day.

1 day	00:00	03:00	06:00	09:00	12:00	15:00	18:00	21:00	Percentage(%)
Pattern_1									1.42
Pattern_2									2.34
Pattern_3									2.76
Pattern_4									10.13
Pattern_5									0.41
Pattern_6									5.86
Pattern_7									2.53
Pattern_8									12.86
Pattern_9									1.58
Pattern_10									0.99
Pattern_11									1.02
Pattern_12									15.22
Pattern_13									1.84
Pattern_14									1.09
Pattern_15									9.71
Pattern_16									0.45
Pattern_17									0.37
Pattern_18									6.31
Pattern_19									0.53
Pattern_20									0.65
Pattern_21									5.14
Pattern_22									4.93
Pattern_23									3.7
Pattern_24									2.05
Pattern_25									
		Departure in time period						Sum	93.89%

1 Percentage of crash risk reduction: Difference in crash risk between the original and after rescheduling. Indication of how much crash risk can be reduced.
2 Reduction in number of working drivers: Difference in number of working drivers between the original and after rescheduling. The greater the reduction the more a company can reduce its financial cost.
3 Deviation of workload: Deviation of all drivers' workloads. The lower the deviation the more equal a driver's work load.

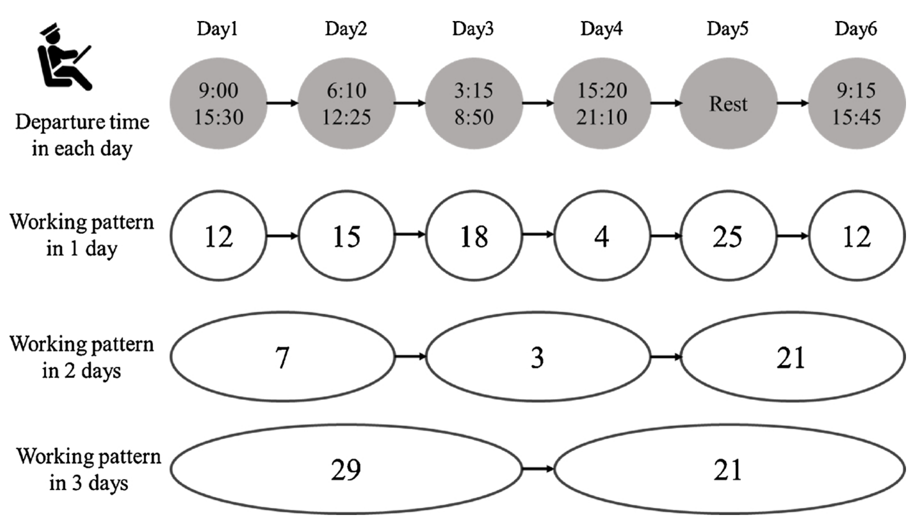

Fig. 3. Example of working patterns for different length of units.

Table 4
Results of a logistic regression based on 1-1 combination.

					First Day								Second Day							
Model 1	Odds Ratio	Std. Err.	# of crashes	# of samples	00:00	03:00	06:00	09:00	12:00	15:00	18:00	21:00	00:00	03:00	06:00	09:00	12:00	15:00	18:00	21:00
Combination1-1	5.628	4.040	2	226						■		■								■
Combination1-2	6.174	4.435	2	206							■			■						
Combination1-3	13.825	14.050	1	46	■						■			■						
Combination1-4	7.571	7.657	1	84					■		■						■			■
Combination1-5	5.191	3.726	2	245					■		■							■	■	
Combination1-6	10.175	7.330	2	125					■		■							■		
Combination1-7	4.702	2.411	4	541				■		■								■		■
Combination1-8	8.833	6.357	2	144				■		■									■	
Combination1-9	6.115	6.178	1	104				■		■								■	■	
Combination1-10	15.899	16.183	1	40				■		■							■			
Combination1-11	7.310	7.392	1	87			■		■									■		■
Combination1-12	16.736	17.045	1	38			■		■									■		
Combination1-13	13.249	13.458	1	48	■		■										■		■	
Combination1-14	8.050	8.145	1	79			■								■			■		
Combination1-15	7.227	7.307	1	88	■									■		■				
Combination1-16	6.523	4.686	2	195																■
Combination1-17	7.709	5.543	2	165															■	
Combination1-18	13.531	13.748	1	47														■	■	
Combination1-19	8.652	6.226	2	147														■		
Combination1-20	9.784	9.912	1	65													■			
Combination1-21	12.979	13.181	1	49												■				
Constant	0.002	0.000	91	57872	Pseudo R2 = 0.0388; Number of obs = 60,764								■	Departure from the station in this time period						

4.1. Identification of working patterns associated with high crash risk

The clustering analysis resulted in 443 combinations of working patterns with a unit of 1-day, 455 combinations with a unit of 2-day, and 1257 combinations with a unit of 3-day. A logistic regression model is first employed to identify and quantify crash risk of each combination of working patterns, as shown in Table 4. But it should be noted that only those combinations passing through the case-control study would be assigned a cost in the optimization process. For the case-control study, the numbers of controls matched to a case were also tested (2, 4, 6, and 10), and it was found that 10 controls would be the most suitable because not only is the sampling distribution closer to a normal distribution, it is also applicable given the data available.

As expected, the coefficients of the three logistic regressions for combinations with different units are different. Nevertheless, the combinations that are identified to be associated with higher crash risk for all three models are consistent. As an example, Fig. 4 shows three combinations associated with high crash risks in terms of the three units (1-1, 2-2, and 3-3 combinations), and it is apparent that the working patterns in the last two days are the same, i.e. drivers work in the afternoons for two consecutive days.

Ultimately, two combinations of working patterns were identified as the combinations having the greatest crash risk, and they are:

1 **Driving in the afternoon or early hours of the morning for two consecutive days**: The afternoon (2:00pm ~ 4:00pm) and the early hours of the morning (2:00am ~ 6:00am) are considered to have a higher crash risk than other time periods due to our circadian rhythm. It is assumed our body will become drowsy during these two time periods (Williamson et al., 2011; Pack et al., 1995). And driving during these two periods for two consecutive days will only exacerbate driver fatigue. The crash risk is 6.57 times that of other driving patterns.
2 **After resting at least 24 h, commence driving in the morning, afternoon or the early hours of the morning**: Chen and Xie (2014) and Jovanis et al. (2012) reported that crash risk will increase after a driver has rested for a prolonged period. This may be because a driver needs a little time to re-adjust his physical and psychological

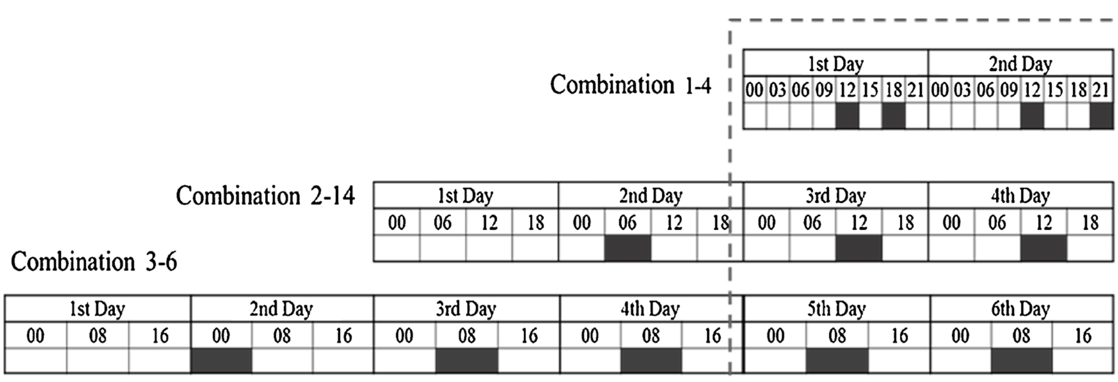

Fig. 4. An example of combinations with similar working patterns but different units.

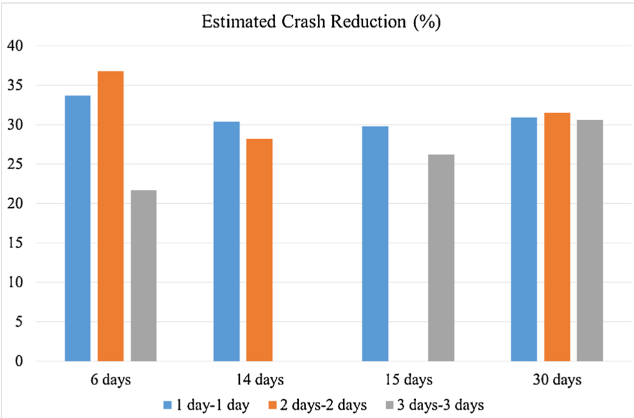

Fig. 5. Comparison of rescheduling results with different units and planning horizon.

Table 5
Results of sensitivity analyses.

Time horizon	Combinations	Performance Measures	
		Percentage of crash risk Reduction	Deviation of workload (times/day/driver)
6 days	1 day-1 day	34%	0.275
	2 days-2 days	37%	0.545
	3 days-3 days	22%	0.399
14 days	1 day-1 day	30%	0.319
	2 days-2 days	28%	0.251
	3 days-3 days	N/A	N/A
15 days	1 day-1 day	30%	0.246
	2 days-2 days	N/A	N/A
	3 days-3 days	26%	0.243
30 days	1 day-1 day	31%	0.204
	2 days-2 days	32%	0.306
	3 days-3 days	31%	0.311

equilibrium. The crash risk is 7.68 times compared to other driving patterns. Interestingly, we found that after resting for at least 24 h, it was safest to return to work in the evening (6:00pm ~ 12:00pm).

4.2. Crash risk reduction after rescheduling

After applying the algorithm developed in the methodology section, the results of rescheduling showed a significant crash reduction of as much as 30 percent in the 30-day planning horizon, given the carrier maintained the existing operational timetable, as shown in Fig. 5. In Fig. 5, the results of different planning horizons, 6, 14, 15, and 30 days, were also reported. The crash reduction is most robust and stable (30 percent crash reduction) when using a single day as an analysis unit (the blue bars).

In addition to lowering crash risk, intercity bus carriers are also interested in reducing operational costs and avoiding unequal workloads among drivers, i.e. the number of working drivers needed to run a master schedule and the deviation of all drivers' workloads. Workload deviation is the standard deviation of drivers' daily workload (number of trips per day). The higher the deviation, the higher the workload difference between drivers. Although the results also show that all the drivers in the fleet are still required to provide cover for all shifts, as shown in Table 5, the deviations in workload varied. Most of them are less than the original deviation in workload, which is 0.36, meaning workloads are shared more evenly. The workload deviation of 15 and 30 days planning horizon with 1-1 combinations have the lowest deviation, i.e. the most even workload.

Overall, the longer the time horizon, the more robust the results for crash reduction and deviations in workload. The 15 and 30 days planning horizon with 1-1 combinations have both greater crash reduction as well as the most even workload distribution.

5. Summary and discussion

Driver scheduling has long been suspected as one of the most influential factors affecting sleepiness and fatigue. As such, a practical question for intercity bus carriers is how to reduce the number of crashes associated with driver schedules while maintaining a non-stop service? This research seeks to develop a paradigm to minimize overall fleet crash risk by rescheduling.

In this study, we first identified those driving schedules associated with higher crash risk, and a rescheduling scheme is then proposed to reduce fleet crashes overall. It was found that an intercity bus company should avoid having its drivers continuously driving in the afternoon or in the early hours of the morning for two consecutive days, and likewise should arrange an evening shift for drivers who have rested at least 24 h (because the crash risk in the evening is lower than other time periods following a 24-hour rest). By avoiding high crash risk working pattern combinations, rescheduling can reduce as much as 30 percent of crashes.

The most important feature of the method we proposed is to piece together and quantify the crash risk associated with each combination of working pattern, and then minimize a fleet's crash risk by rescheduling their working patterns. The method developed in this study can be readily applied to different routes, different carriers, different analysis units (e.g. change 1 day to 8 h), or even to other considerations of scheduling (e.g. drivers have to rest at least 2 days a week). It should be noted that the method proposed in this study does not eliminate all crash risk, but does lower the crash risk per day by avoiding dangerous working patterns and reassigning shifts to different drivers or different days to avoid risk.

Although the findings are promising, there are still several limitations, as follows. First, this study is based on the assumption that at-fault crashes are more likely to be associated with driver scheduling. But this may not always be the case since at-fault crashes may also be due to other factors causing driver impairment, e.g. drugs or illness. Secondly, as a case-control study does not allow calculation of absolute risk, the crash probabilities estimated in this study are the "relative" crash risks compared to other working patterns. Therefore, care should be taken when applying the results of this study if one is interested in formulating a bus driver schedule. Thirdly, technical issues concerning the complexity of optimization arise when broader driver contractual agreements are considered; for example, some drivers work full time (40-hour work weeks), while others may be part-time or drive split shifts. Lastly, although a case-control study was designed along with the bootstrapping method to mitigate potential biases due to model misspecification, there are a couple of ways to further limit the potential biases: (1) more data should be included and a cross validation (using half of the data for model building and the other half for testing) could be conducted to test the robustness of the estimation; (2) estimates based on the use of intermediate performance measures to capture risky driving behaviors that are associated with crash occurrence while driving may well be more accurate than the use of actual crashes as a performance measure.

The implications of the findings are twofold. First, consistent with past research, it was found that driver scheduling in terms of multi-day driving patterns indeed has an effect on a driver's crash risk. This is also supported by the fact that rescheduling could help reduce overall fleet crash risk. Secondly, future research is recommended to further examine why these multi-day driving patterns are associated with higher crash risk, and hopefully, the results of this study could then be incorporated in future service regulations for driver hours.

Acknowledgement

The authors thank the Ministry of Science and Technology, Republic of China (Taiwan), for supporting this research (106-2221-E-009-069-MY2).

Appendix

Table A1
Working patterns in terms of a unit of two days.

2 day	First day				Second day				
	00:00	06:00	12:00	18:00	00:00	06:00	12:00	18:00	Percentage
Pattern_1	■				■				7.37
Pattern_2	■					■			3.69
Pattern_3	■						■		0.49
Pattern_4	■							■	0.07
Pattern_5	■								1.16
Pattern_6		■			■				0.96
Pattern_7		■				■			17.08
Pattern_8		■					■		7.8
Pattern_9		■						■	0.51
Pattern_10		■							4.24
Pattern_11			■		■				0.02
Pattern_12			■			■			1.35
Pattern_13			■				■		12.25
Pattern_14			■					■	4.85
Pattern_15			■						5.46
Pattern_16				■		■			0.24
Pattern_17				■			■		0.24
Pattern_18				■				■	6.93
Pattern_19				■					6.49
Pattern_20					■				4.38
Pattern_21						■			7.96
Pattern_22							■		3.19
Pattern_23								■	1.2
Pattern_24									
	■	Begin to work in time period						Sum	97.93%

Table A2
Working patterns in terms of a unit of three days.

3 day	First day			Second day			Third day			
	00:00	08:00	16:00	00:00	08:00	16:00	00:00	08:00	16:00	Percentage
Pattern_1									■	0.58
Pattern_2								■		2.28
Pattern_3							■			1.08
Pattern_4						■				0.83
Pattern_5						■			■	1.15
Pattern_6					■					1.19
Pattern_7					■				■	0.67
Pattern_8					■			■		4.79
Pattern_9					■		■			0.78
Pattern_10				■						0.58
Pattern_11				■				■		2.25
Pattern_12				■			■			4.6
Pattern_13			■							1.13
Pattern_14			■						■	0.49
Pattern_15			■					■		1.23
Pattern_16			■				■			4.48
Pattern_17			■			■				4.24
Pattern_18			■			■			■	7.28
Pattern_19		■								2.2
Pattern_20		■							■	0.75
Pattern_21		■						■		2.78
Pattern_22		■					■			0.9
Pattern_23		■				■				2.36
Pattern_24		■				■			■	2.88
Pattern_25		■			■					4.37
Pattern_26		■			■				■	4.44
Pattern_27		■			■			■		13.29
Pattern_28		■		■				■		0.95
Pattern_29		■		■			■			0.19
Pattern_30	■									0.8
Pattern_31	■							■		0.67
Pattern_32	■						■			0.45
Pattern_33	■				■					0.64
Pattern_34	■				■				■	0.54
Pattern_35	■				■			■		5.01
Pattern_36	■				■		■			0.75
Pattern_37	■			■						0.82
Pattern_38	■			■				■		3.31
Pattern_39	■			■			■			7.7
Pattern_40										
	■	Begin to work in time period							Sum	95.43%

References

Akerstedt, T., Gillberg, M., 1990. Subjective and objective sleepiness in active individuals. Int. J. Neurosci. 52, 29–37.

Al-Yakoob, S.M., Sherali, H.D., 2007. Mixed-integer programming models for an employee scheduling problem with multiple shifts and work locations. Ann. Oper. Res. 155, 119–142.

Banihashemi, Mohamadreza, Ali Haghani, 2000. "Optimization model for large-scale bus transit scheduling problems". Transp. Res. Rec. J. Transp. Res. Board No. 1733. Transportation Research Board of the National Academies, Washington, DC.

Bonnefond, A., Tassi, P., Rogé, J., Muzet, A., 2004. A critical review of techniques aiming at enhancing and sustaining worker's alertness during the night shift. Ind. Health 42, 1–14.

Cassidy, J.D., Boyle, E., Côté, P., He, Y., Hogg-Johnson, S., Silver, F.L., Bondy, S.J., 2009. Risk of vertebrobasilar stroke and chiropractic care: results of a population-based case-control and case-crossover study. J. Manipulative Physiol. Ther. 32 (2), S201–S208.

Chen, C., Xie, Y., 2014. Modeling the safety impacts of driving hours and rest breaks on truck drivers considering time-dependent covariates. J. Saf. Res. 51, 57–63.

Crum, M., Morrow, P., 2002. The influence of carrier scheduling practices on truck driver fatigue. Transp. J. 42 (1), 20–41.

Crum, M., Morrow, P., Olsgard, P., Roke, P., 2001. Truck driving environments and their influence on driver fatigue and crash rates. Transp. Res. Rec. 1779, 125–133.

de Matta, R., Peters, E., 2009. Developing work schedules for an inter-city transit system with multiple driver types and fleet types. Eur. J. Oper. Res. 192 (3), 852–865.

Desmond, P.A., Matthews, G., 1997. Implications of task-induced fatigue effects in vehicle countermeasures to driver fatigue. Accid. Anal. Prev. 29 (5), 15–523.

Feyer, A.M., Williamson, A.M., 1995. The influence of operational conditions on driver fatigue in the long distance road transport industry in Australia. Int. J. Ind. Ergon. 15, 229–235.

Fischetti, Matteo, et al., 2001. "A Polyhedral Approach to Simplified Crew Scheduling and Vehicle Scheduling Problems", vol. 47 Management Science. The Institute for Operations Research and the Management Sciences (INFORMS) 6, June.

Hulst, Van der M., Meijman, T., Rothrngatter, T., 2001. Maintaining task-set under fatigue a study of time-on-task effects in simulated driving. Transport. Res. F 4, 103–108.

Jovanis, P.P., Wu, K.-F., Chen, C., 2012. Effects of hours of service and driving patterns on motor carrier crashes. Transp. Res. Rec. J. Transp. Res. Board 2281, 119–127.

Kahneman, D., 1973. Attention and Effort. Prentice Hall, Englewood Cliffs, NJ.

Kaneko, T., Jovanis, P.P., 1992. Multiday driving patterns and motor carrier accident risk: a disaggregate analysis. Accid. Anal. Prev. 24 (5), 437–456.

Laumon, B., Gadegbeku, B., Martin, J.L., Biecheler, M.B., 2005. Cannabis intoxication and fatal road crashes in France: population based case-control study. Bmj 331 (7529), 1371.

Pack, A.I., Pack, A.M., Rodgman, E., Cucchiara, A., Dinges, D.F., Schwab, C.W., 1995. Characteristics of crashes attributed to the driver having fallen asleep. Accid. Anal. Prev. 27 (6), 769–775.

Rogé, J., Pébayle, T., Hannachi, S.E., Muzet, A., 2003. Effect of sleep deprivation and driving duration on the useful visual field in younger and older subjects during simulator driving. Vis. Res. 43 (13), 1465–1472.

Shinar, D., 2007. Traffic Safety and Human Behavior.

Soccolich, S.A., Blanco, M., Hanowski, R.J., Olson, R.L., Morgan, J.F., Guo, F., Wu, S.-C., 2013. An analysis of driving and working hour on commercial motor vehicle driver safety using naturalistic data collection. Accid. Anal. Prev. 58, 249–258.

Sparrowa, A.R., Mollicone, D.J., Kanb, K., Bartels, R., Satterfielda, B.C., Riedya, S.M., Unice, A., Van Dongena, H.P.A., 2016. Naturalistic field study of the restart break in US commercial motor vehicle drivers: truck driving, sleep, and fatigue. Accid. Anal. Prev. 93, 55–64.

Torrance, K., Haire, A.R., Machemehl, R.B., 2009. "Vehicle and Drivers Scheduling for Public Transit,". University of Texas, Austin.

Torregroza-Vargas, N.M., Bocarejoa, J.P., Ramos-Bonilla, J.P., 2014. Fatigue and crashes: The case of freight transport in Colombia. Accid. Anal. Prev. 72, 440–448.

Williamson, A., Feyer, A., Friswell, R., 1996. The impact of work practices on fatigue in long distance truck drivers. Accid. Anal. Prev. 28 (6), 709–771.

Williamson, A., Lombardi, D.A., Folkard, S., Stutts, J., Courtney, T.K., Connor, J.L., 2011. The link between fatigue and safety. Accid. Anal. Prev. 43 (2), 498–515.

Wren, A., Rousseau, J.M., 1995. Bus Driver Scheduling- An Overview. University of Leeds Report, pp. 173–187 93.31, 1993, Computer-Aided Transit Scheduling.

Wylie, C.D., Shultz, T., Miller, J.C., Mitler, M.M., Mackie, R.R., 1996. Commercial Motor Vehicle Driver Fatigue and Alertness Study: Technical Summary. Publication FHWA-MC-97-001. FHWA, U.S. Department of Transportation.

Yaoyuenyong, K., Nanthavani, Sj, 2005. Energy-based workforce scheduling problem: mathematical model and solution algorithms. Sci. Asia 31, 383–393.