CASPT 2015

An Offline Framework for Reliability Diagnosis by Automatic Vehicle Location Data

Benedetto Barabino · Massimo Di Francesco · Roberto Murru

Abstract Time reliability problems are unavoidable because of the stochastic environment where bus services are operated. The characterization of reliability and the comprehension of its possible sources may help keep buses on schedule and/or maintaining planned headways, as public transport companies may be put in the position of selecting the most suitable strategies and passengers are expected to receive higher-quality service. This paper presents a framework which aims to (a) characterize the reliability over all bus stops and time periods of each route, (b) quantify the occurrence of unreliability sources and (c) make quantification-based links to the most appropriated strategies for the case at hand. The experimentation of this framework is performed in a real environment by easy-to-read control dashboards in order to provide practical insights on unreliability for the local bus operator.

Keywords Automatic Vehicle Location data · Service Quality · Time reliability measure · Transit network · Wealth of data

1 Introduction

Generally speaking, the reliability is the capability of Public Transport Companies (PTCs) to provide the service as promised in terms of multidimensional aspects, such as time, passenger loads, vehicle quality and so on (Ceder (2007);

B. Barabino
Technomobility Srl, Cagliari, Italy
E-mail: bbarabino@gmail.com

M. Di Francesco
Department of Mathematics and Computer Science, University of Cagliari, Italy
E-mail: mdifrance@unica.it

R. Murru
CTM SpA, Cagliari, Italy
E-mail: roberto.murru@ctmcagliari.it

Chen et al (2009); Kimpel (2001); Van Oort (2011)). Relevant reliability elements in terms of time are regularity for medium - high frequency services and punctuality for medium low frequency services.
Time reliability problems are unavoidable due to the stochastic environment where bus services are operated. Therefore, investigating the reliability is crucial for PTCs in order to select the most suitable strategies and passengers are expected to receive higher service quality . According to Abkowitz et al (1978) and Cham (2006), the analysis of reliability can be organized as follows:

- Characterizing the reliability, i.e. including key data inputs, calculating additional attributes from inputs, describing service measures, setting bus-operator-dependent thresholds and generating performance reports.
- Identifying possible unreliability problem sources which can clustered into four groups namely: (i) Improper Service Design (ISD), (ii) Drivers and/or Supervisors Failures (D&SF), (iii) Uncertainties in Passengers Volumes (UPV) and (iv) Uncontrollable External Factors (UEF) (Ceder (2007)).
- Selecting strategies which can be classified according to their type (e.g. priority, operational and control) and their applications (preventive and corrective).

Nowadays, measuring time reliability is technologically feasible by AVL systems, which can collect abundances of disaggregated data on the delivered service. The proper processing by automatic methods of archived Automatic Vehicle Location (AVL) data can disclose information on the service time reliability to PTCs. This information is of tremendous interest to them because of the advantage of rapidly alerting them where attention is needed, while avoiding that PTC planners scrutinize where reliability performance is low.
However, additional work must be done to set up a practical framework supporting PTCs in facing unreliability sources. Therefore, this study aims to present a framework, which (a) characterizes the reliability over all bus stops and time periods of each route, (b) quantifies the occurrence of unreliability sources and (c) selects the most fitting strategies for the case at hand. Steps (b) and (c) extend step (a) as described in (Barabino et al (2013a), Barabino et al (2013b),Barabino et al (2015)) who already addressed AVL data anomalies such as Bus Overtaking (BO), Techical Failures (TF) and Incorrect Operation in the Service (IOS). Moreover, the framework is expected to provide more detailed results with respect to Horbury (1999), Cham (2006), Mandelzys and Hellinga (2010), Feng and Figliozzi (2011), because they focused on a limited subset of bus stops and time periods.
In addition, continuous advances on ITS can provide information on all bus stops and improve the links between the occurrences of unreliability sources and strategies. Results are presented by easy-to-read Control Dashboards (CDs), which are tables organized in time and space attributes. Unlike space-time trajectory graphs, these CDs do not suffer from the abundance of represented data.
This paper is organized as follows. Section 2 presents prior studies on the

reliability characterization, sources of unreliability and strategies to improve service reliability. Section 3 proposes a practical framework to deeply analyze time reliability at all bus stops and time periods for each route. Its experimentation on a route of a medium-sized PTC is illustrated in Section 4. Finally, in Section 5, conclusions and research perspectives are outlined.

2 Literature Review

According to Abkowitz et al (1978)and Cham (2006), reliability analysis can be organized as follows:

- *Characterizing the reliability*, i.e. including key data inputs, calculating additional attributes from input, describing service measures, setting bus-operator-dependent thresholds and generating performance reports.
- *Identifying possible unreliability problem sources.* Although several unreliability problem sources exist and are not clear-cut, they can be listed to describe service unreliability phenomena and clustered into four groups: (a) ISD, (b) D&SF, (c) UPV and (d) UEF (Ceder (2007)). Moreover, despite their possible links, (i) and (ii) are more directly under the control of PTCs.
- *Selecting strategies*, which are classified according to their type and application.

Early experiences in previous steps of reliability analysis were performed by PTCs using manually collected data. Due to economic constraints and a lack of technology, data could not typically be collected at each bus stop of a route during service hours. As a result, measurements were performed at a few random (Nakanishi (1997)) or selected (Henderson et al (1990); Trompet et al (2011)) check points, such as terminals, owing to the ease of control, and mid or maximum load sections, in which the peak demand must be served (Cham (2006)). In order to have representative data, observations are aggregated in time intervals representing slack and peak hours in the morning and in the evening. These activities usually exhibit a significant level of empiricism and unpredictability, as PTCs are forced to operate with little, if any, data. Therefore, *ad hoc manual* data handling procedures have been adopted, but this *modus operandi* requires small data sets and results in too local analysis and narrow conclusions. In fact, single check points and/or aggregated time intervals do not describe the detailed reliability of the whole route. Currently, the analysis of reliability over all bus stops and time intervals is of preminent interest because passengers are concerned mostly with adherence to the headways or to the schedule at their particular stop (e.g. Abkowitz et al (1978); Kimpel (2001); Koffman (1992)). Moreover, this type of analysis can be supported by Automated Data Collection Systems such as AVL and APC technology, which can collect huge amounts of raw data for different bus stops and time intervals (e.g. Mendes-Moreira et al (2015); Moreira-Matias et al (2015); Furth et al (2004), Furth et al (2006); Hounsell et al (2012)). For example, some applications are discussed in Strathman et al (1999); Tétreault and El-Geneidy

(2010); Barabino et al (2014)), but additional challenges must be faced in order to effectively handle, accurately process and represent AVL data in a user-friendly way. In what follows we review some studies on the previously listed organization of reliability analysis, as well as their links with adopted AVL technologies.

2.1 Characterizing the reliability

We briefly review existing studies on the previous components of reliability characterization:

- The input is represented by observed data, which contain details on time and location. Data input was mainly available through manual surveys (e.g. Nakanishi (1997), Strathman et al (1999), Liu and Sinha (2007), Kimpel (2001)) or automated data collection system (Camus et al (2005), Cham (2006), Lin et al (2008), Liu and Sinha (2007), Lin and Ruan (2009), Mandelzys and Hellinga (2010), Feng and Figliozzi (2011), Barabino et al (2013a), Barabino et al (2013b), Ma et al (2014),Barabino et al (2015)).
- Output calculations are outcomes calculated from a single or a pair of data inputs. Typically, they were schedule deviation (e.g. Nakanishi (1997), Strathman et al (1999), Cham (2006), Chen et al (2009), Mandelzys and Hellinga (2010)), actual and scheduled headways (Nakanishi (1997), Strathman et al (1999), Camus et al (2005), Cham (2006), Lin et al (2008), Liu and Sinha (2007), Lin and Ruan (2009), Chen et al (2009), Feng and Figliozzi (2011), Barabino et al (2013a), Barabino et al (2013b)), dwell time (Cham (2006), Mandelzys and Hellinga (2010), Feng and Figliozzi (2011)) and running times (e.g. Strathman et al (1999), Camus et al (2005), Cham (2006), Lin et al (2008), Liu and Sinha (2007), Chen et al (2009), Mandelzys and Hellinga (2010), Feng and Figliozzi (2011)).
- Service measures are a set of aggregated metrics used to characterize the overall bus service, measure performance and evaluate the provided service. Typically, service measures include on-time performance (Nakanishi (1997), Strathman et al (1999), Associates et al (2003), Cham (2006), Mandelzys and Hellinga (2010)), headway adherence (Associates et al (2003), Barabino et al (2013a), Barabino et al (2013b)), running time distribution (Cham (2006), Lin et al (2008), Liu and Sinha (2007), Ma et al (2014)), early and late departures (Cham (2006)), and occurrence and distribution of bunching, (Cham (2006), Feng and Figliozzi (2011)), etc. (Camus et al (2005), Strathman et al (1999), Liu and Sinha (2007), Lin and Ruan (2009), Chen et al (2009), Ma et al (2014)). Synthetic metrics outputs can be expressed in terms of mean values and variance, coefficient of variation, and the percentage of observation. In fact, they are (a) well understood by PTC managers; (b) widely adopted in PTCs; and (c) clearly represented by quantification.
- Thresholds are values used to set the output acceptability (e.g. the calculation of a punctual bus arrival at stop) and the acceptability of the

service measures such as LoS (e.g. the level of OTP). Despite thresholds being PTC-dependent, they were largely used for punctuality analysis in the range between 3 minutes early and 5 minutes late (Nakanishi (1997), Strathman et al (1999), Associates et al (2003), Camus et al (2005), Cham (2006), Chen et al (2009), Mandelzys and Hellinga (2010)). Threshold settings for regularity diagnosis can be found in (e.g. Nakanishi (1997), Cham (2006),Feng and Figliozzi (2011)).
– Performance reports are dashboards visualizing final results and possible intermediate outcomes. They can be appropriate tables (e.g. Lin et al (2008), Cortés et al (2011)) or graphs (e.g. Chen et al (2009), Cortés et al (2011)).

Data collection affects the overall quality of previous steps: in the case of manual collection, one has typically little input data which result in problematic inferences on conclusions; in the case of AVL data collections, despite data abundance, additional pre-processing is required to correct data anomalies such as BO, IOS and TF.
Most of previous studies do not completely characterize bus reliability at all bus stops and time periods and neglects data anomalies, which may result in incorrect reliability characterization. For example, only human-collected data were used in Chen et al (2009). All anomalies in AVL data were disregarded in Horbury (1999), Camus et al (2005) and Lin and Ruan (2009), whereas missing data points were not investigated in Lin et al (2008) and Feng and Figliozzi (2011). All previous drawbacks were investigated in Barabino et al (2013a), Barabino et al (2013b), Barabino et al (2015), but they focused on characterization only, without any analysis of unreliability sources and related strategies.

2.2 Identifying unreliability sources

To our knowledge, lower attention has been devoted to identify possible unreliability sources from AVL/APC data. Cham (2006) investigated unreliability mainly at terminals, because occurring problems may spread down the route. The author investigated a route whose headway varies between 3 and 15 minutes and inferred that unreliability sources may be drivers and/or supervisors failing in the execution of planned services. Mandelzys and Hellinga (2010) presented a methodology to identify time points, where standard schedule adherence was not achieved, and causes of inadequate performances. Their experimental results showed that unreliability causes were not dependent on the departure terminal. Hammerle et al (2005) analysed bunching causes at the start terminal and at other bus stops by plotting time-space trajectory graphs, and showed that problems depend on irregularity headways at the terminal. Feng and Figliozzi (2011) proposed a method to identify some bus-bunching causes, organized in combinations of attributes of the front bus and the following one. While previous studies aimed to observe and quantify unreliability problems, this paper aims to link unreliability problems to whomever

Table 1 Summary of strategies to improve service reliability

Type of strategy	Sub-type of strategy	Preventive	Corrective
Priority	Exclusive lanes (Bus only streets, busways, with and contra flow bus lane)	•	
	Route Design	•	
	Signal Priority	•	•
Operational	Reserve vehicle and operators	•	•
	Operator training	•	
	Operator incentives and penalties	•	
	Schedule adjustments	•	
	Supervision	•	
	Improve vehicle access (e.g. fare collection, device for boarding/alightings)	•	•
Control	Holding (Scheduled-based or Headway-based)		•
	Overtaking		•
	Expressing (Full expressing, Limited stops, Alighting only)		•
	Short- Turning		•
	Deadheading		•
	Exchanging vehicle shift		•
	Adding a reinforce shift		•
	Providing in-vehicle message		•
	Operator self-regulation		•

is in charge of their correction. In order to perform this analysis, unreliability phenomena are clustered into the four groups described in Ceder (2007).

2.3 Selecting strategies

In line with Abkowitz et al (1978)and Cham (2006), strategies can be classified according to their type and application. The types of strategies are:

– *Priority strategies*, where buses receive proper care to reduce the influence of external factors.
– *Control strategies*, where buses receive real-time advices on their service operation.
– *Operational strategies*, where buses receive changes in the planning of routes, schedules and resource allocations.

Their application can be divided into preventive and corrective: the first case aims to reduce the expected occurrence of reliability problems *a priori*, and focuses on reducing the variability of running and dwell times; and the second case aims to reduce their spread and focuses on reducing the negative impact to passengers.

According to previous classification, strategies and related sub-strategies are summarized in Table 1, which is self-explanatory. Since our analysis is based on archived AVL data, this study refers to preventive strategies only.

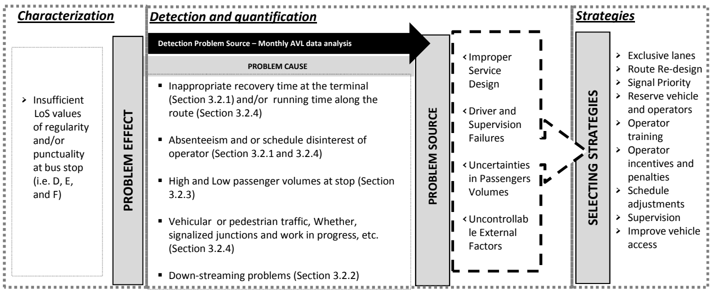

Fig. 1 Effects, causes and sources of unreliability problems

3 Methodology

In this section, a new framework is presented to characterize service time reliability, recognize possible unreliability sources, measure their magnitude and systematically select some strategies for their mitigation. The methodological framework is organized in three steps:

– Characterizing service reliability;
– Detecting and quantifying unreliability problems sources;
– Systematization of strategies.

These steps interplay each other according to the scheme illustrated in Figure 1.

This figure shows how to detect problem sources from problem effects by processed AVL monthly data, which pointed out when and where unreliability problems occur in the characterization stage. Next, one switches to the analysis of possible unreliability causes. Then, one infers information on the possible sources generating the problem. Finally, some possible links are established between sources and strategies.

3.1 Realiability characterization

First the AVL data of the considered route are picked up from a database on the provided service. The most relevant attributes for this characterization are: date, route, vehicle-block, trip number, bus stop code and order, actual and scheduled arrival (or departure) times and, finally, the time spent in a pre-defined area around each bus stop, or the dwell time, depending on the specific AVL architecture. Second, AVL data need proper handling to account for BOs and recognize missing data points, which are either TFs or IOSs. BOs are addressed by ordering actual arrival (or departure) times chronologically at bus stops. Missing data points are recognized by building a monthly report of IOSs, which is merged with the original schedule and, next, with AVL data. In this method, any missing data is considered as a TF unless it is a reported

IOS. Third, the route is classified as high or low frequency, according to its scheduled headway. In the first case, the route is evaluated in terms of regularity, otherwise in terms of punctuality. Since the separation is not clear-cut, a route can be also analysed in both ways. Fourth, the reliability measure is computed. In the case of high frequency routes, the actual and scheduled headways are computed as the difference between two consecutive arrival (or departure) times. It is important to note that actual headways need different analysis, depending on the occurrence of TFs and IOSs. TFs cannot be used to compute the real headways, owing to the missing information on arrivals (or departures) occurring for real. Conversely, actual headways are computed in the case of IOSs, because these temporal gaps are suffered by passengers for real. Next, the coefficient of variation of headway Cvh is computed for all bus stops and time periods, and can be linked to the appropriate LoS (e.g. Associates et al (2003)). In the case of low frequency routes, deviations from the original schedule are computed as the difference between actual and scheduled arrival (or departure) time. In this case, TFs are disregarded. Conversely, IOSs are penalized, because passengers are actually conditioned by these missing bus arrivals (or departures). Next, the OTP is computed for all bus stops and time periods, and can be linked to the appropriate LoS (e.g. Associates et al (2003)). Fifth, if LoS values report a sufficient mark denoted by A, B or C, the service is considered as acceptable and no further analysis is required. Conversely, if LoS values report an insufficient mark denoted by D, E or F, the service needs further investigation to understand the possible unreliability sources. Sixth, AVL processed data are represented effectively using CDs organized in space and time attributes in order to show which parts of the route contain most of the problems and deserve further analysis.

3.2 Unreliability sources identication

3.2.1 Sources at terminals

Let A be the set of terminals and J the set of runs. For each terminal $a \in A$, the arrival and departure of a generic run $j \in J$ are considered in case of run conclusions or beginning, respectively. Since real and scheduled departure time of a run $j \in J$ may be different, one may compute their difference as:

$$T_a^j = RDT_a^j - SDT_a^j \qquad \forall a \in A, \forall j \in J \tag{1}$$

where,

- T_a^j represents the time deviation of run $j \in J$ at terminal $a \in A$.
- RDT_a^j represents the real departure time of $j \in J$ at terminal $a \in A$.
- SDT_a^j represents the scheduled departure time of $j \in J$ at terminal $a \in A$.

If $T_a^j > 0$, the run departs after the schedule; if $T_a^j \approx 0$, the run is on time; if $T_a^j < 0$ the run departs early. Since early and late departures represent unreliability problems, in what follows we focus on their analysis. In order

Table 2 Possible unreliability sources at the terminal

ART_a^j / T_a^j	< 0	≈ 0	> 0
< 0	n/a	n/a	ISD
≈ 0	ok	ok	D&SF and ISD
> 0	D&SF	ok	D&SF

to begin the next trip as scheduled and give operators a short break, drivers are provided with recovery times, but actual recovery times may be different from scheduled ones. Therefore, it is important to derive the available recovery time, which is computed as

$$ART_a^j = SDT_a^j - RAT_a^{j-1} \qquad \forall a \in A, \forall j \in J \tag{2}$$

where,

- ART_a^j represents the available recovery time of run $j \in J$ at terminal $a \in A$.
- SDT_a^j represents the scheduled departure time of run $j \in J$ at terminal $a \in A$.
- RAT_a^{j-1} represents the real arrival time of run $j-1 \in J$ at terminal $a \in A$.

The analysis of the deviation at the terminal and that of the available recovery time are crucial for detecting unreliability sources, because they may result in additional irregularity in the next run. Their combined analysis helps understand if unreliability depends mainly on ISD or D&SF. Nine different cases can be obtained, as shown in Table 2, according to the sign of T_a^j and ART_a^j. The notation ≈ 0 must be read as:

$$\alpha \leq T_a^j \leq \beta \qquad \forall a \in A, \forall j \in J \tag{3}$$

$$\gamma \leq ART_a^j \leq \delta \qquad \forall a \in A, \forall j \in J \tag{4}$$

where α, β, γ and δ are numerical thresholds set up by the PTC, and represent the minimum and maximum acceptable T_a^j and ART_a^j values for all terminals $a \in A$ and runs $j \in J$, respectively. Each entry in Table 2 shows the type of possible unreliability source, or when there is not an applicable condition (n/a), or a good performance is obtained (ok), respectively. For each entry in Table 2, one can compute the related magnitude in terms of percentage values. Table 2 shows that four critical combinations of T_a^j and ART_a^j may occur when:

- $ART_a^j < 0$ and $T_a^j > 0$, the average driver usually does not have the available recovery time for the new run, then they start a late run and the problem may be ISD.
- $ART_a^j > 0$ and $T_a^j > 0$, the average driver usually has sufficient recovery time to start the new run on time, but they start a later-than-scheduled run. Therefore, in this case the problem may be (D&SF).

– $ART_a^j > 0$ and $T_a^j < 0$, the average driver usually has sufficient recovery time to start the new run in time, but they start an earlier-than-scheduled run. Even in this case, the problem may be (D&SF).
– $ART_a^j \approx 0$ and $T_a^j > 0$, the average driver usually has little or no recovery time for the new run, so they may start their trips late, if they want to have a break. However, in this case there may also be a problem in ISD or D&SF.

3.2.2 Down-Streaming sources

Let I be the set of all bus stops (including terminals) and J the set of runs. In the case of bus stop $i \in I/A$, a run $j \in J$ can arrive (depart) before the scheduled time, on-time or after the scheduled time. In order to recognize these cases, one can compute the difference between the real arrival (or departure) times and the scheduled arrival (or departure) times as:

$$T_i^j = RAT_i^j - SAT_i^j \qquad \forall i \in I/A, \forall j \in J \tag{5}$$

or

$$T_i^j = RDT_i^j - SDT_i^j \qquad \forall i \in I/A, \forall j \in J \tag{6}$$

where,

– T_i^j represents the time deviation of run $j \in J$ at bus stop $i \in I/A$.
– RAT_i^j represents the real arrival time of run $j \in J$ at bus stop $i \in I/A$.
– SAT_i^j represents the scheduled arrival time of run $j \in J$ at bus stop i$i \in I/A$.
– RDT_i^j represents the real arrival time of run $j \in J$ at bus stop $i \in I/A$.
– SDT_i^j represents the scheduled arrival time of run $j \in J$ at bus stop i$i \in I/A$.

If $T_i^j > 0$, the run arrives (departs) after the schedule; if $T_i^j \approx 0$, the run is on time; if $T_i^j < 0$ the run arrives (departs) before. As early and late transits represent unreliability problems, in what follows we focus on their analysis. Since reliability problems tend to spread along the route, it is also important to consider for each bus stop $i \in I/A$ the previous one, which is denoted by $i-1 \in I/A$, and compute T_{i-1}^j and compare T_{i-1}^j and T_i^j. In fact, early and late arrivals (or departures) at bus stop $i \in I/A$ may be generated by early and late transits at bus stop $i-1 \in I/A$. To conclude, the comparison between T_{i-1}^j and T_i^j helps understand if problems are clustered at a specific bus stop, or if they are propagated as a result of upstream causes, according to the following cases:

– early arrivals (or departures) at bus stop $i-1 \in I/A$ and $i-1 \in I/A$;
– late arrivals (or departures) at bus stop $i-1 \in I/A$ and $i \in I/A$.

The magnitude of clustered problems and upstream causes is expressed in terms of percentage values.

3.2.3 Time spent at stops

The time spent may depend on passenger volumes if the dwell time is not available. In fact, the dwell time can provide information on the volumes of boarding and alighting passengers, in the case of less advanced PCTs, which are not APC-equipped, and may not be well-informed about passengers (typically, the longer this time, the greater the number of boarding and alighting passengers). Moreover, to our knowledge, many AVL architectures are conceived to record the time spent in the proximity of bus stops. Then, in what follows, we refer to the time spent instead of the dwell time. As a result, the method tries to understand if problems at stop $i \in I/A$ depend on passenger volumes (UPV). The magnitude of these causes is evaluated by percentage values. More precisely, two different types of time spent are considered: the real time spent rts_i^j by run $j \in J$ at bus stop $i \in I/A$; and the scheduled mean time spent at bus stop $i \in I/A$, which is computed as follows:

$$smts_i^j = \frac{\sum_{j=1}^{N} sts_{i,j}}{N} \qquad \forall i \in I \tag{7}$$

where:

- N represents the maximum number of scheduled times spent at bus stop $i \in I/A$.
- $smts_i^j$ represents the scheduled time mean spent at stops $i \in I/A$.
- $sts_{i,j}$ represents the scheduled time spent by run $j \in J$ at bus stop $i \in I/A$.

Some problems can be detected by the value of rts_i^j and appropriate threshold parameters $\epsilon < 1$ and $\zeta > 1$:

- if $rts_i^j < \epsilon * smts_i^j$,the volume of passenger boarding and alighting is probably lower than expected and, thus, UPV may occur.
- if $\epsilon * smts_i^j \leq rts_i^j \leq \zeta * smts_i^j$, the volume of passenger boarding and alighting is as expected and there is no problem.
- if $rts_i^j > \zeta * smts_i^j$, the volume of passenger boarding and alighting is probably greater than expected and, thus, UPV may occur.

3.2.4 Speed between bus stops

In order to understand the causes of early and late transits at bus stop $i \in I$, the proposed method performs an analysis of the speed along any leg from $i-1 \in I$ to $i \in I$, because speed can provide information about the running time. Next, the method tries to understand if problems along any leg from $i-1 \in I$ to $i \in I$ depend on ISD, (D&SF) or UEF. The magnitude of these causes is evaluated by percentage values. Two different speeds are considered: the real speed between stops $i-1 \in I$ and $i \in I$ for each run $j \in J$; and the scheduled mean speed between stops $i-1 \in I$ and $i \in I$. If we consider a fixed time period t, they are computed as:

$$rs^j_{i-1,i} = \frac{l_{i-1,i}}{rrt^j_{i-1,i}} \qquad \forall i \in I, \forall j \in J \tag{8}$$

$$sms_{i-1,i} = \frac{l_{i-1,i}}{\frac{\sum_{j=1}^{N} srt_{(i-1,i)j}}{N}} \qquad \forall i \in I, \forall j \in J \tag{9}$$

where:

- N represents the maximum number of scheduled running times, which are recorded on $j \in J$ on the leg between stops $i-1 \in I$ and $i \in I$.
- $l_{i-1,i}$ represents the length of the leg between stops $i-1 \in I$ and $i \in I$.
- $rs^j_{i-1,i}$ represents the real speed of run jJ between stops $i-1 \in I$ and $i \in I$.
- $sms_{i-1,i}$ represents the scheduled mean speed between stops $i-1 \in I$ and $i \in I$.
- $rrt^j_{i-1,i}$ represents the real running time between stops $i-1 \in I$ and $i \in I$.
- $srt_{(i-1,i)j}$ represents the scheduled running time of run $j \in J$ between stops $i-1 \in I$ and $i \in I$.

Some problems can be detected by the value of $rs^j_{i-1,i}$ and appropriate threshold parameters $\iota < 1$ and $\kappa > 1$.

- if $rs^j_{i-1,i} \leq \eta$ (i.e.the minimum acceptable speed), the unreliability source is probably UEF, because, if it did not occur, buses would run faster.
- if $\eta < rs^j_{i-1,i} \leq \iota * sms_{i-1,i}$ the unreliability source is probably ISD, because, even if buses run beyond the minimum acceptable speed, they cannot reach the planned speed.
- if $\iota * sms_{i-1,i} < rs^j_{i-1,i} \leq \kappa * sms_{i-1,i}$ there is no problem disclosed by the speed analysis, because the real speed is close to the planned one.
- if $rs^j_{i-1,i} > \kappa * sms_{i-1,i}$ or larger than the urban speed limit θ, the cause is probably D&SF, owing to the slack between the expected driving style and the executed one, which is too sporty.

3.3 Strategies Selecting

Since this framework is intended to operate offline, we focus only on preventive strategies, which can be divided into priority and operational ones. A possible link between unreliability sources and strategies is reported in Table 3, although additional analysis is required by the PTC when sources do not occur systematically.

4 Application in a real case

The experimentation was performed with CTM, which is a bus operator providing public transport services in eight communalities of the metropolitan area of Cagliari (Sardinia, Italy), which has about 400,000 inhabitants. CTM

Table 3 Summary of possible preventive strategies which can be implemented

Unreliability source	Type of strategies	Sub-Type of strategies
UEF	Priority	Exclusive lanes
UEF		Route re-design
UEF		Signal Priority
ISD	Operational	Reserve vehicle and operators
D		Operator training
D		Operator incentives and penalties
ISD, UPV		Schedule adjustments
D&SF		Supervision
UPV		Improving vehicle access (e.g. fare collection, device for boarding/alightings)

manages 264 buses, serving around 35,500,000 passengers a year on 30 routes. For the sake of synthesis, the proposed method is tested on the estbound direction of a route 5.45 km long with 18 bus stops, which links a suburban area from the city centre through the historical district. The route has been chosen because of these heterogeneous characteristics, which are supposed to point out different problem sources in its parts:

- Its headway is 12 minutes from 07:00 to 19:59.
- Vehicles deployed on this route have the same typology from 7.00 to 19.59 (capacity = 58 passengers, length = 8 m, low-floor).
- The route is close to regional government offices, schools, hospitals, and shopping centres. It can be divided into five parts, depending on heterogeneous traffic components. In Part 1 (from bus stop 1 to 3) buses move along larger streets with few pedestrian flows. In Part 2 (from bus stop 4 to 7), buses approach the historical district and move in mixed traffic. The street is narrow, and high numbers of pedestrians and vehicles may interfere with buses. As a result, buses may perceive several disturbances during their service. In Part 3 (from bus stop 8 to 11), buses move through the historical district, and only pedestrian and bus movements are allowed. In Part 4 (from bus stop 12 to 14) buses move along one-way streets in mixed-traffic conditions. In this more central part, increased pedestrian and vehicular flows occur. Several vehicles are likely to be looking for parking spaces at low speed, because no parking slots are available in Part 3. Finally, in Part 5 (from bus stop 15 to 18) buses leave the city-centre area along two-way streets in mixed-traffic conditions.
- The scheduled running time is 22 minutes on average, and amounts to 24 minutes from 18:00 to 19:59. Recovery time is scheduled to be, on average, 3 minutes at the departure terminal on weekdays, but it amounts to 0 minutes from 18:00 to 19:59.

Since 2007, all buses have been equipped with a specific AVL architecture which records various data, such as actual and scheduled arrival times at every bus stop, measured in minutes and seconds. Around 100,000 AVL data records are available on a daily basis over the CTM network. In addition, control room operators follow buses in real time at the Automatic Vehicle Monitoring centre

Eastbound direction

Bus stop	7.00 7.59	8.00 8.59	9.00 9.59	10.00 10.59	11.00 11.59	12.00 12.59	13.00 13.59	14.00 14.59	15.00 15.59	16.00 16.59	17.00 17.59	18.00 18.59	19.00 19.59
1	nd	nd	nd	nd	nd	nd	D	nd	nd	nd	nd	nd	D
2	A	A	A	A	A	B	D	nd	nd	A	A	C	E
3	A	A	A	A	A	B	D	nd	nd	A	A	B	D
4	A	A	A	A	B	C	D	A	A	A	A	B	D
5	A	A	A	A	A	C	D	B	A	A	A	B	D
6	nd	A	nd	nd	nd	B	D	nd	nd	A	A	B	D
7	A	A	B	A	B	C	D	B	A	A	A	B	D
8	A	A	B	A	B	C	D	A	A	A	A	B	E
9	A	A	B	A	B	C	D	B	A	A	A	B	E
10	A	A	B	A	B	C	D	B	A	A	A	B	D
11	A	A	C	A	B	C	D	C	A	A	B	C	D
12	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
13	A	B	C	A	B	C	D	C	A	A	B	C	D
14	A	B	C	B	B	C	D	C	A	A	B	C	nd
15	A	B	C	B	B	C	D	C	A	A	B	C	nd
16	A	B	C	A	B	C	D	C	A	B	B	C	D
17	A	B	C	A	B	C	D	C	A	B	B	C	E
18	A	B	C	B	B	C	D	C	A	A	B	C	E

Fig. 2 The first CD which characterizes the reliability in terms of regularity in all bus stops and time periods (n/a means data not available)

Eastbound direction

Bus stop	7.00 7.59	8.00 8.59	9.00 9.59	10.00 10.59	11.00 11.59	12.00 12.59	13.00 13.59	14.00 14.59	15.00 15.59	16.00 16.59	17.00 17.59	18.00 18.59	19.00 19.59
1	A	A	A	A	B	C	C	B	A	A	A	E	E
2	A	A	B	B	A	B	C	nd	nd	B	B	D	E
3	C	C	E	E	E	D	D	nd	nd	F	E	E	F
4	F	E	F	F	F	E	F	D	E	F	F	F	F
5	F	F	F	F	F	F	F	E	F	F	F	F	F
6	nd	F	nd	nd	nd	F	F	nd	nd	F	F	F	F
7	F	F	F	F	F	F	F	F	F	F	F	F	F
8	F	F	F	F	E	F	F	F	F	F	F	F	F
9	F	F	F	F	E	F	F	F	F	F	F	F	F
10	F	F	F	E	D	F	F	F	F	F	F	F	F
11	F	F	F	F	D	F	F	F	F	F	F	F	F
12	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
13	F	F	E	C	D	E	F	F	E	F	F	F	F
14	F	nd	E	nd	D	nd	F	nd	E	nd	F	nd	nd
15	F	nd	E	nd	E	nd	F	nd	D	nd	F	nd	nd
16	F	E	E	D	D	E	F	F	E	E	F	F	F
17	F	F	E	D	D	E	F	F	E	F	F	F	F
18	F	F	E	D	D	E	F	F	F	F	F	F	F

Fig. 3 The first CD which characterizes the reliability in terms of punctuality in all bus stops and time periods (n/a means data not available)

and inform drivers of online actions to improve time reliability. As a vehicle terminates its service, it moves back to the depot where data recorded during the daily shift are downloaded by a wireless connection. Daily AVL data are stored in a central database. The AVL data of the route were collected during weekdays of July 2014.

The method in Section 3 was developed and implemented on MS Access and MS Excel on a standard PC (Pentium 4, CPU 2.80 gHz, RAM 1.00 gb).

Table 4 Terminal analysis at time period 19.00 19.59

Bus stop	D&SF	ISD	ISD and or D&SF	ok
1	0 %	49 %	1%	50%

Currently, owing to its offline use, this framework is not yet implemented in a single tool, but the calculation of all outcomes is automatic and can be performed in a few seconds. Results on Section 3.1 of the method can be represented as CDs consisting of tables, where lines represent bus stops and columns time periods. Each entry represents the LoS at that bus stop and at that time period. Since this route can be analysed in terms of either punctuality or regularity, the related CDs are both reported. Figure 2 reports the CD on regularity, which is measured according to Associates et al (2003).
Figure 3 reports the CD on punctuality, which is measured in terms of OTP under the assumption of punctual transits ranging from -1 early and 3 minutes late. For the sake of practice, an *ad hoc* scale was arranged as follows: LoS F means less than 50 % of punctual transits, LoS E between 50% and 60%, and so on.

Figure 2 provides evidence that the route is generally well performed in terms of regularity, and problems are clustered only in two time periods. On the other hand, Figure 3 shows that punctuality is a critical features of this route. A possible explanation for these contrasting results is the online supervision of the regularity route by control room operators located at the dispatch centre. However, the PTC is satisfied with the service in this route, because it is managed in terms of regularity only.
In order to carry out the experimentation of Section 3.2 in the method, the selection of the time period is required. Since both punctuality and regularity disclose criticalities in the last time period, it was selected and analysed in what follows. Moreover, since CTM does not have data on scheduled time spent at bus stop $i \in I/A$, some minor adjustments were made to (7), (8) and (9): (a) the schedule mean time spent (*smts*) was calculated by averaging the rts for each time period; and (b) the scheduled running time (*srt*) was calculated by the difference between the scheduled arrival time at bus stops $i \in I/A$ and $i-1 \in I/A$, respectively. For the sake of consistency, this calculation was performed for the real running time as well. Table 4 reports the analysis of the departure terminal. The table shows that about half of problem sources depend on ISD, and hence the driver does not have the usual available recovery time for the new run, so trips usually start late. Yet, the route seems to be affected by low reliability from its beginning.
Figure 4 reports the analysis of the unreliability in the following bus stops and in each leg, where the black background shows occurrences larger than 50%. The down-streaming analysis shows that, in Part 1, buses usually run late, but the delay tends to decrease owing to high speed and lower-than-expected passenger volumes. When buses enter Part 2, they have already corrected the delay, but they maintain high speeds, which result in the occurrence of

Time period [19.00 - 19.59]	Down-streaming analysis				Time spent analysis			Leg_Code	Speed analysis			
Bus stop	E_E	L_L	Other	OK	L_UPV	U_UPV	OK		D&SF	ISD	UEF	OK
								T1A	**99%**	0%	0%	1%
2	0%	42%	14%	44%	**100%**	0%	0%					
								T2A	**100%**	0%	0%	0%
3	9%	34%	27%	30%	**79%**	1%	20%					
								T3A	**85%**	0%	6%	9%
4	26%	27%	21%	26%	**59%**	12%	29%					
								T4A	**90%**	0%	1%	9%
5	39%	25%	10%	26%	23%	19%	58%					
								T5A	**96%**	0%	2%	3%
6	46%	23%	6%	25%	38%	11%	51%					
								T6A	42%	0%	25%	32%
7	49%	22%	1%	28%	5%	**65%**	30%					
								T7A	7%	0%	**80%**	13%
8	45%	20%	8%	26%	5%	**71%**	24%					
								T8A	6%	0%	**82%**	12%
9	38%	23%	15%	23%	**93%**	2%	5%					
								T9A	6%	0%	**77%**	17%
10	30%	27%	22%	21%	8%	**58%**	34%					
								T10A	*nd*	*nd*	*nd*	*nd*
11	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*					
								T11A	*nd*	*nd*	*nd*	*nd*
12	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*					
								T12A	*nd*	*nd*	*nd*	*nd*
13	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*					
								T13A	*nd*	*nd*	*nd*	*nd*
14	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*					
								T14A	*nd*	*nd*	*nd*	*nd*
15	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*					
								T15A	*nd*	*nd*	*nd*	*nd*
16	*nd*	*nd*	*nd*	*nd*	**53%**	2%	46%					
								T16A	**100%**	0%	0%	0%
17	27%	38%	8%	27%	**75%**	0%	25%					
								T17A	**99%**	0%	0%	1%
18												

Fig. 4 Analysis of unreliability sources per down-streaming, time spent and speeds

early arrivals. At the beginning of Part 3, the high passenger volumes at stops and the lower speeds owing to UEF (i.e. high pedestrian flows) decrease the percentage of early bus transits. Despite some missing figures, in Part 4 there is a lower role of UEF, but it is problematic to maintain the scheduled trip, and the percentage of early arrivals tends to stabilize. In Part 5, the reduced passenger volumes and the higher-than-scheduled speed do not help increase the percentage of early run arrivals.
According to Section 3.2.2 of the method, some strategies can be taken from the analysis of the route in this time period, as described in Table 3. The lack of recovery times at terminals exhibits ISD, whereas along the route the difference between scheduled and actual speed results in D& SF. Since in Part 3 of the route, it is not possible to increase the speed, because of large volume of pedestrian flows, no action can be taken to correct drivers. In addition it is not possible to add exclusive bus lanes owing to the topological nature of the street. As a result, the recommended strategy is to adjust the schedule in order to provide drivers with longer recovery times, slower running times in Parts 1 and 2, and larger ones in the remaining parts of the route.

5 Conclusions

Although AVL data are commonly used in PTCs for the real-time monitoring of buses, little attention has been devoted to the use of archived data for understanding reliability in detail. However, analysing unreliability on bus routes is crucial for the quality and efficient operational planning of PCTs. This paper sheds light on the use of AVL archived data to characterize unreliability and detect its possible sources. We have proposed a framework to:

- automatically generate a mainstream source of AVL archived data;
- include streams of AVL data in the framework using a single data source and integrating procedures to measure the magnitude of problem sources at terminals, bus stops and legs between consecutive bus stops;
- provide details on bus route unreliability sources at all bus stops and time periods.

The proposed framework is currently being tested in a real bus route by the bus operator CTM in Cagliari (Italy), and their experience will be used for further research in the tuning of thresholds.

Acknowledgements The authors are grateful to Eng. Sara Mozzoni for help support in preparing data per analysis.

References

Abkowitz M, Slavin H, Waksman R, Englisher R, Wilson N (1978) Transit service reliability. Tech. rep., USDOT Transportation Systems Center and Multisystems, Cambridge, MA. NTIS, no. UMTA/MA-06-0049-78-1

Associates K, Administration USFT, Program TCR, Corporation TD (2003) Transit Capacity and Quality of Service Manual, vol 100. Transportation Research Board

Barabino B, Di Francesco M, Mozzoni S (2013a) Regularity diagnosis by automatic vehicle location raw data. Public Transport 4(3):187–208

Barabino B, Di Francesco M, Mozzoni S (2013b) Regularity analysis on bus networks and route directions by Automatic Vehicle Location raw data. IET Intelligent Transportation Systems 7(4):473–480

Barabino B, Di Francesco M, Mozzoni S (2014) An offline framework for handling automatic passenger counting raw data. IEEE Transaction on Intelligent Transportation Systems 15(6):2443– 2456

Barabino B, Di Francesco M, Mozzoni S (2015) Rethinking bus punctuality by integrating automatic vehicle location data and passenger patterns. Transportation Research Part A: Policy and Practice 75:84–95

Camus R, Longo G, Macorini C (2005) Estimation of transit reliability level-of-service based on automatic vehicle location data. Transportation Research Record: Journal of the Transportation Research Board 1927(1):277–286

Ceder A (2007) Public transit planning and operation: theory, modelling and practice. Springer

Cham LC (2006) Understanding bus service reliability: a practical framework using AVL/APC data. Master's thesis, Massachusetts Institute of Technology

Chen X, Yu L, Zhang Y, Guo J (2009) Analyzing urban bus service reliability at the stop, route, and network levels. Transportation Research Part A: Policy and Practice 43(8):722 – 734

Cortés CE, Gibson J, Gschwender A, Munizaga M, Zúñiga M (2011) Commercial bus speed diagnosis based on gps-monitored data. Transportation Research Part C: Emerging Technologies 19(4):695–707

Feng W, Figliozzi M (2011) Using archived avl/apc bus data to identify spatial-temporal causes of bus bunching. In: Compendium of papers of 90th Transportation Research Board, Washington, D.C.

Furth PG, Muller TH, Strathman JG, Hemily B (2004) Designing automated vehicle location systems for archived data analysis. Transportation Research Record: Journal of the Transportation Research Board 1887(1):62–70

Furth PG, Hemily B, Muller TH, Strathman JG (2006) Using archived AVL-APC data to improve transit performance and management. Project H-28

Hammerle M, Haynes M, McNeil S (2005) Use of automatic vehicle location and passenger count data to evaluate bus operations. Transportation Research Record: Journal of the Transportation Research Board 1903(1):27–34

Henderson G, Adkins H, Kwong P (1990) Toward a passenger-oriented model of subway performance. Transportation Research Record (1266)

Horbury AX (1999) Using non-real-time automatic vehicle location data to improve bus services. Transportation Research Part B 33(8):559–579

Hounsell N, Shrestha B, Wong A (2012) Data management and applications in a world-leading bus fleet. Transportation Research Part C: Emerging Technologies 22:76–87

Kimpel TJ (2001) Time point level analysis of transit service reliability and passenger demand. PhD thesis, Portland State University,Oregon

Koffman J (1992) Automatic passenger counting data: Better schedules improve on-time performance. In: Computer-Aided Transit Scheduling, Springer, pp 259–282

Lin J, Ruan M (2009) Probability-based bus headway regularity measure. IET intelligent transport systems 3(4):400–408

Lin J, Wang P, Barnum DT (2008) A quality control framework for bus schedule reliability. Transportation Research Part E: Logistics and Transportation Review 44(6):1086–1098

Liu R, Sinha S (2007) Modelling urban bus service and passenger reliability.

Ma Z, Ferreira L, Mesbah M (2014) Measuring service reliability using automatic vehicle location data. Mathematical Problems in Engineering 2014

Mandelzys M, Hellinga B (2010) Identifying causes of performance issues in bus schedule adherence with automatic vehicle location and passenger count data. Transportation Research Record: Journal of the Transportation Research Board 2143(1):9–15

Mendes-Moreira J, Moreira-Matias L, Gama J, de Sousa JF (2015) Validating the coverage of bus schedules: A machine learning approach. Information

Sciences 293:299–313
Moreira-Matias L, Mendes-Moreira J, de Sousa JF, Gama J (2015) Improving mass transit operations by using avl-based systems: A survey. IEEE Transaction on Intelligent Transportation Systems (forthcoming) 99(1):100–101
Nakanishi YJ (1997) Part 1: Bus: Bus performance indicators: On-time performance and service regularity. Transportation Research Record: Journal of the Transportation Research Board 1571(1):1–13
Strathman JG, Dueker KJ, Kimpel T, Gerhart R, Turner K, Taylor P, Callas S, Griffin D, Hopper J (1999) Automated bus dispatching, operations control, and service reliability: Baseline analysis. Transportation Research Record: Journal of the Transportation Research Board 1666(1):28–36
Tétreault PR, El-Geneidy AM (2010) Estimating bus run times for new limited-stop service using archived avl and apc data. Transportation Research Part A: Policy and Practice 44(6):390–402
Trompet M, Liu X, Graham DJ (2011) Development of key performance indicator to compare regularity of service between urban bus operators. Transportation Research Record: Journal of the Transportation Research Board 2216(1):33–41
Van Oort N (2011) Service reliability and urban public transport design. PhD thesis, Delft University of Technology