Detection of High-Speed red-aspect approaches using a multi-data approach

R.A.H. EL Rashidy[1], P. Hughes[1], C. Van Gulijk[1].

[1]*Institute of Railway Research, University of Huddersfield, Huddersfield, UK*

Abstract

This paper proposes a proactive safety indicator for Signals Passed at Danger (SPAD) using multiple data sources. The proposed technique integrates data sources using a graph database and R software to store, process and analyze the train services and signals data to yield a key performance indicator for high-speed red-aspect approaches. The method is illustrated using a case study where three data sources were used *viz*., On Train Data Recorders (OTDR), Red Aspect Approach to Signals (RAATS) data and the railway infrastructure manager's signal database. The proposed approach aims to shift safety management for SPADs from 'avoiding things going wrong' to 'ensuring that everything goes right'. The approach is complementary to the current driver competency performance system, potentially allowing the driver supervisors to evaluate drivers' braking style without any subjectivity. It also indicates which signals on the infrastructure may be particularly prone to SPADs, even if a SPAD has not occurred there yet.

Keywords: Proactive safety management; Red aspect; Driver performance; Graph database

1 Introduction

The exponential growth and wide availability of digital data offers transformative potential for proactive safety management of the railway. The UK Government's strategic vision for rail (DfT, 2017) has emphasized the need for the digital railway, which includes: the exploitation of data sources to allow a better understanding of safety-related issues; to maximize efficiency gains; to improve asset management; and to improve travel experience. Despite these potential benefits, use of big-data approaches to support the operational railway safety management has, to date, been imperceptible (The European Union Agency for Railways, 2016). Safety learning has remained solely with safety experts often by performing incident analysis after an accident has occurred. These accident analysis reports provide recommendations to prevent similar incidents occurring in the future (for example Rail Accident Investigation Branch 2016a; 16b). Whilst such a retrospective approach is clearly meaningful, it is an aspiration of railway safety staff to be able to identify potential causes of accidents prior to accidents occurring. Identifying high-risk situations that could, but did not, result in accidents is one of the ways forward for proactive safety management. This work describes such a proactive approach for a major risk on the railways: Signals Passed at Danger.

Proactive safety management has a a wide range of applictions in verious areas such as the health sector (Hollnagel, 2014) and construction management (Carbonari et al., 2011). It emphasizes the role of day-to-day performance on safety, rather than focusing on incidents or accidents. Continuous monitoring of performance can work as a layer of defense that detects deviations in performance that may lead to adverse safety consequences (Sumwalt et al., 2002). Real-time performance monitoring can identify divergences from best practice rules and enables timely corrective measures to be taken.

This paper proposes a non-intrusive data-driven approach based on multiple data sources to monitor how a train driver approaches a red aspect. The approach identifies high risk signal approaches which could lead to a serious breach of safety: a SPAD. This kind of information can also be used to improve driver competency performance systems. It provides the opportunity to replace time-consuming, obtrusive monitoring of drivers based on manually marking-up OTDR (On Train Data Recorders) download, potentially freeing up a significant amount of driver managers' time to undertake more targeted support of driver competence. Furthermore, the proposed approach will allow the driver supervisors to evaluate the braking style without any subjectivity as it is based on well-defined calculations.

2 Signals Passed at Danger

Signals Passed at Danger (SPAD) is an event where a train passes a stop signal and proceeds onto a section of track where it does not have authority. A SPAD is a serious breach of safety that can lead to trains colliding with other trains or road vehicles on level crossings, derailing, or striking workers and equipment. Accidents following SPADS, such as Ladbroke Grove which resulted in 31 fatalities in 1999, still ring clear in the railway's safety conscience (Health and Safety Executive, 2000). After a series of fatal accidents, the Train Protection and Warning System (TPWS) was introduced in Britain in the 1990s; it helped reduce the number of SPADs significantly (Nikandros and Tombs, 2007). A risk-based approach was used to identify locations where TPWS should be installed, aiming to cover all by the lowest risk signals (RSSB, 2015). TPWS helps identify near-miss SPAD events by recording emergency brake demand activation/intervention[1] but the system does not provide details about events preceding the SPAD. In addition to TPWS, the automatic warning system (AWS) provides drivers with an audible and visual indication of whether the signal they are approaching is clear (showing a green aspect) or at caution (either red, yellow or double yellow). When the signal is clear, the AWS system sounds a bell, signifying a clear signal approach, and no response is required. If the signal is at caution signal, the AWS system sounds a horn and drivers need to respond by pressing an 'acknowledge button'. If a driver fails to react within 2.2 seconds, the brakes automatically halt the train. Train drivers must avoid over-speeding and control the train's speed to avoid interventions from the TPWS

[1] A TPWS intervention occurs when the TPWS applies the brakes in the absence of (or prior to) the driver doing so whereas, activation occurs when a driver has already applied the brakes before the TPWS operates (RSSB, 2013).

and AWS, at the same time they are expected to avoid driving the train so slowly that timetable requirements are not met.

According to Nikandros and Tombs (2007), human error is largely responsible for SPADs occurring. Within the literature, there are a number of methods that could be used to monitor drivers' responses to signals, one example is to record drivers' eye movements (Groeger et al., 2003 and Luke et al., 2006). These studies focused on the time spent looking at signals and how different attributes such as signal aspect, signal type and signal complexity affect the visual strategies of train drivers. A simulation technique (Li et al., 2006) was also used to investigate drivers' responses to lineside signals and signs at various speeds. The study showed a non-linear relationship between drivers' responses to signals/signs and approach speed. Whilst the results of such studies may demonstrate how drivers use different visual strategies under different circumstances, they do not contribute to train driver performance evaluation since they consider only the driver's observation activities and do not describe what action was taken as a result of the observation.

Driver performance, which addresses human error in a more general sense, is routinely observed by in-cab riding by driver managers. A number of train operating companies focus on the use of in-cab safety techniques such as monitoring drivers' use of the Driver's Reminder Appliance (DRA) (McCorquodale et al., 2002). In some cases, digital cameras have been utilized to record driver's action (RSSB, 2004); however, these techniques are not widely used in Britain. Whilst these techniques have their merits in assessing the driver's safety performance as they supply comprehensive details about the driver's actions, , drivers may act differently under observation, limiting the potential for independent driver assessment. Another technique for monitoring drivers' performance is to occasionally review OTDR data, e.g. once every six months, and manually mark it up to identify sharp braking events or over speed intervals. The OTDR data gives a clear description of the driver actions but without any details of the signal status, consequently it is difficult to understand how a driver approached certain signals. The time and cost of traditional methods hinder their use for continuous monitoring. A particular situation, where the drivers' actions are examined, is after accident Accident investigations are often supported by extensive analysis of data from OTDR (DfT, 2007) to help scrutinize safety-related actions rigorously. Fortunately accidents are relatively rare on the railways and this kind of data analysis does not routinely take place.

There are a number of research studies that consider use of OTDR data to systematically improve safety (Balfe, 2017; El Rashidy & Van Gulijk, 2016; Walker & Strathie, 2015; Green et al., 2011). OTDR data is used in different areas such as station duties, driver assessment and interaction with warning systems; Green et al. (2011) proposed a number of performance indicators such as the speed at which the maximum power setting is selected when accelerating, and the mean speed when AWS horns are activated. El Rashidy & Van Gulijk (2016) introduced a number of indicators to assess drivers' use of the emergency bypass switch, the DRA as well as drivers' reaction times.

Data systems enable a systematic and complete detection of high-speed red-aspect approaches. Using different data sources, such as OTDR (On Train Data Recorders) and RAATS (Red Aspect Approach to Signals), provides a greater insight into understanding how drivers approach signals showing red aspects. Quantifying how drivers approach red aspects will aid to identify where the safety operational envelope is being pushed (Hale et al., 2007). This paper proposes a near-miss indicator (NMI) that can be used not only to continuously detect high-risk red-aspect approaches but also to understand how drivers approach a red aspect.

3 Method

3.1 Data sources

The data sources used in this study were OTDR, RAATS and the national signal database. OTDR is performed with an on-board device, which is used within the GB Railways to collect data relating to train operations and the state of various train systems throughout a journey such as power and brake controller position and driver acknowledgment of signaling system warnings. In this investigation 20 OTDR data files: the files form pairs with two OTDR files representing one day's data for one train, i.e. one file from the driving cab at either end of the train.

The RAATS data is used to determine the status of the signal approached by a service. RAATS combines two separate parts of the Train Describer (TD) data feeds from Network Rail on train movements and signal states, namely C-class to track the train movements and S-class messages to time the signal aspect changing; more details about RAATS and the way in which it stores data can be found in Zhao et al. (2017). One month's worth of operational data for the railway-line overlapping with the 20 OTDR files (14,978,800 records) were used in this paper.

The signals database contains 40,007 records of signal data covering the British railway network in its entirety. The database includes a range of information such as the signal ID, type, description, signal box and location. The availability of such information facilitates the synchronization between OTDR and RAATS; some records were incomplete for this route so a manual check to identify unrealistic records (such as the entering berth time is less than the exit time of the previous berth) was required for some records.

3.2 Graph database

To facilitate data integration, a graph database was used to store, manage and query the data. The fundamental units of a graph database are nodes and edges. Nodes typically contain chunks of data and edges create meaningful relationships between the nodes, making the database flexible, scalable and relatively easy to work with. The advantages of using a graph database for safety purposes are explained elsewhere (El Rashidy et al., 2018).

RAATS and OTDR files were imported into a graph database (Neo4J Community Edition version 3.3.5) and different labels were used to identify the data included in the nodes. Table 1 shows the content of a “Service” node and a “Service Instance” node; in total, 177,121 service instances, were loaded into the database. A parser algorithm was developed to deposit three RAATS types into three node types “Approaching NRA”, “Approaching RA” and “Error” which represent “not red aspect approach”, ”red aspect approach” and “error”, respectively (see table 1); more details about RAATS classification can be found in Zhao et al. (2017). RAATS records, of which there were 14,978,800, were loaded to the graph database and classified into these three categories. The signal database contains many parameters of the signal, including type, location and comments but only the signal reference was used; 34,989 signals were entered into the database.

Table 1: Node labels for OTDR and RAATS data.

Node labels	Data included
Service	Name of OTDR source record Start time End time Service ID
Service Instance	Service ID Service instance timestamp (UTS timestamp) Travel distance (in relation to start of the service) Speed (at the instance of the timestamp) AWS/TPWS data (active, not active) Data about the state of technical systems (not relevant to this work)
RAATS_RECORD_NRA	Train describer(TD) Service ID Berth entering time, Berth exiting time, Previous, current and proceeding berths Signal ID RAATS[2]
RAATS_RECORD_RA	Idem
RAATS_RECORD_ERR	Idem

3.3 Data logging for moving trains

Two data systems log a train as it traverses a railway line. On train device records many parameters every 1 to 10 seconds and known as OTDR. Figure 1 shows the data points as a string of dots below the train: OTDR data for service (i). For this work only the data of the velocity of the train, the AWS system and the Service identification number (also known as head-code) are relevant. AWS horns and bells are recorded in the OTDR data system as they occur on a train service. Figure 1 illustrates how this works. In either case, the activation of the horn/bell is captured in a single service instance data point alongside a timestamp.

Independently, the RAATS system collects data about the same train from the signaling system. It records when a train enters a berth $t_{in(RAATS)}$, leaves it $t_{clear(RAATS)}$ and when it changes status from caution to clear $t_{cleared(RAATS)}$ and vice versa. It also records the train's head-code. Figure 1 shows the data points as a string of dots above the train (RAATS data).

[2] The classification of red aspect state when the train enter the berth according to Zhao et al., 2017.

Information about each signal, like signal(j), is found network rail Signal Database; location and reference number being the key information for this work.

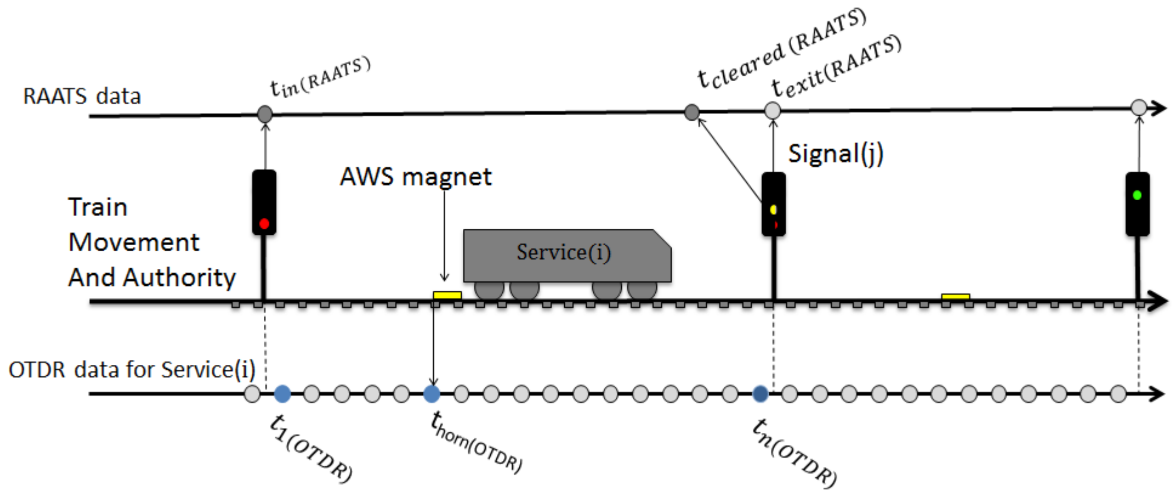

Figure 1: train movement and data-recording.

3.4 Synchronization of data sources

The OTDR and RAATS data-files are not easily linked; both OTDR and RAATS data contain the train's the head-code but this service number is not unique. A head-code typically repeats itself every day on a route (e.g. Manchester to London Route and vice versa); it can appear on different routes at the same time (e.g. on the Manchester to London Route and on a service to Reading at the same time); it can even occur twice at the same time on the same route. In this case, the OTDR data did not store the geographical location, complicating the linkage even further. Figure 1 shows that OTDR-RAATS integration contains information about several events that take place at different times, i.e. from $t_{1(OTDR)}$ to $t_{n(OTDR)}$. It also shows that OTDR data generates many service instance data points in between the time that the train enters and leaves the berth; each of which has a timestamp between entering and leaving the berth. In practice, a hundred or more service instances data-points could be recorded in between signals due to the fact that the service instances data-point is created at any change in one of OTDR channels, which could lead to a data-point every second. In the absence of a link based on unique identifiers, three criteria were used to establish whether data from OTDR and RAATS linked to one another and the same real-life instance of a train approaching a red-aspect as listed below and presented graphically in Figure 1:

1. Service criterion: both data sources have to refer to the same Service ID (head-code).
2. Time criterion: the train travel time recorded by OTDR should be in the range between the entering and exiting berth time recorded by RAATS:

$$t_{in(RAATS)} \leq t_{(OTDR)} \geq t_{exit(RAATS)}$$

where $t_{in(RAATS)}$ is the time when the service entered the berth recorded by RAATS, $t_{(OTDR)}$ is the service travel time recorded by OTDR and $t_{exit(RAATS)}$ is the time when the service exited the berth recorded by RAATS.

3. Logical sequence criterion: the timestamps in RAATS data for entering the berth, exiting the berth, to satisfy the logical sequence of events of a real-life signal approach; i.e. the exiting time for a record must be less than the entering time for the next record.

In case of red aspect approaches two additional conditions have to be met:

a) The AWS Horn occurs before the signal changes, i.e.

$$t_{horn(OTDR)} < t_{cleared(RAATS)}$$

b) The AWS bell occurs after the signal changes, i.e.

$$t_{bell(OTDR)} > t_{cleared(RAATS)}$$

where $t_{horn(OTDR)}$ is the instance travel time when the driver received AWS horn for Service (i), $t_{bell(OTDR)}$ is the instance train travel time when the driver received AWS bell for Service (i) and $t_{cleared(RAATS)}$ is the time when signal (j) changed to proceeding aspect as shown in Figure 1.

The data related to signal approaches were extracted according to the above rules.

3.5 Identification of high-speed red-aspect approaches

To assess how a driver approaches a red aspect, near-miss indicator (NMI) is proposed. The NMI is defined in Equation 1 considering the deceleration rate required to stop a service in front of a red signal compared with the emergency brake of the train class used by the service. NMI incorporates the impact of train travel speed and the time difference between the service exit time and signal cleared time. The value of NMI should be less than or equal to 1 for the driver to be able to stop the train in case the signal continuing to show a red aspect. Values of NMI greater than 1 indicate a near-certain risk of a SPAD.

$$NMI_{ij} = a_{ij}/e \qquad (1)$$

where NMI_{ij} is the near-miss indicator for a service (i) passed a signal (j), a_{ij} is the deceleration rate of service (i) approaching a signal (j) and e_m is the emergency deceleration rate of a train used in this service.

In fact, there are two scenarios where NMIs are applicable.

<u>Scenario I:</u> A service had to stop at the signal, the deceleration rate calculated using Equation (2):

$$a_{ij} = v_{ij\_horn(OTDR)}/(t_{v=0(OTDR)} - t_{horn(OTDR)}) \qquad (2)$$

where $v_{ij\_horn(OTDR)}$ is the train travel speed of service (i) approached a signal(j) at AWS horn as showed in Figure 1.

Scenario II: A service approaches a red aspect which is cleared before the train comes to halt and service proceeds. The train's deceleration is calculated using Equation (3):

$$a_{ij} = v_{ij(OTDR)}/(t_{exit(RAATS)} - t_{cleared(RAATS)}) \tag{3}$$

where $v_{ij(OTDR)}$ is the train travel speed of service (i) when a signal(j) aspect changes from a red aspect to a proceed aspect. $t_{cleared(RAATS)}$ is the time when signal (j) changed to proceeding aspect as shown in Figure 1. $t_{exit(RAATS)}$ is the time when the service exited the berth recorded by RAATS.

In this work, NMI investigates a 'what if condition', i.e. the ability of the train driver to stop the train before passing a signal at danger; a high-risk condition that is encountered too often on the railways (especially under low-adhesion conditions).

4 Results

Out of 14,978,800 RAATS records, 3957 records were synchronized with 67,637 OTDR data points. The majority of signals approaches recorded by RAATS (3926 approaches) were "not red aspects signals" and 31 signals showed a red aspect.

Figure 2 and Figure 3 show the speed profile of services A and B, respectively, with the signals that were passed or stopped at during their journeys. Red aspect signals are shown as in red lines, non-red aspect signals are show as grey notches. Figure 2 shows a service A that encountered 191 signals of which 5 displayed red aspects. In one red aspect approach, the train halted. In service B, the driver encountered 1 red aspect that cleared prior the trains arrival at that point as depicted in Figure 3.

Figure 4 shows a magnified scale of the speed profile of service A in front of different signals that showed a red aspect when the service entered the berth. The color code is used to differentiate the train travel speed before and after the approach signal changes its aspect, i.e. red color represents the travel speed during red aspect phase whereas green color represents the travel speed after the signal clearing.

Table 2 presents the NMI indices of red aspect approaches for 27 services where OTDR, RAATS and Signals database have successfully synchronized, whereas Figure 5 presents the frequency of NMI indices of all services used in this paper.

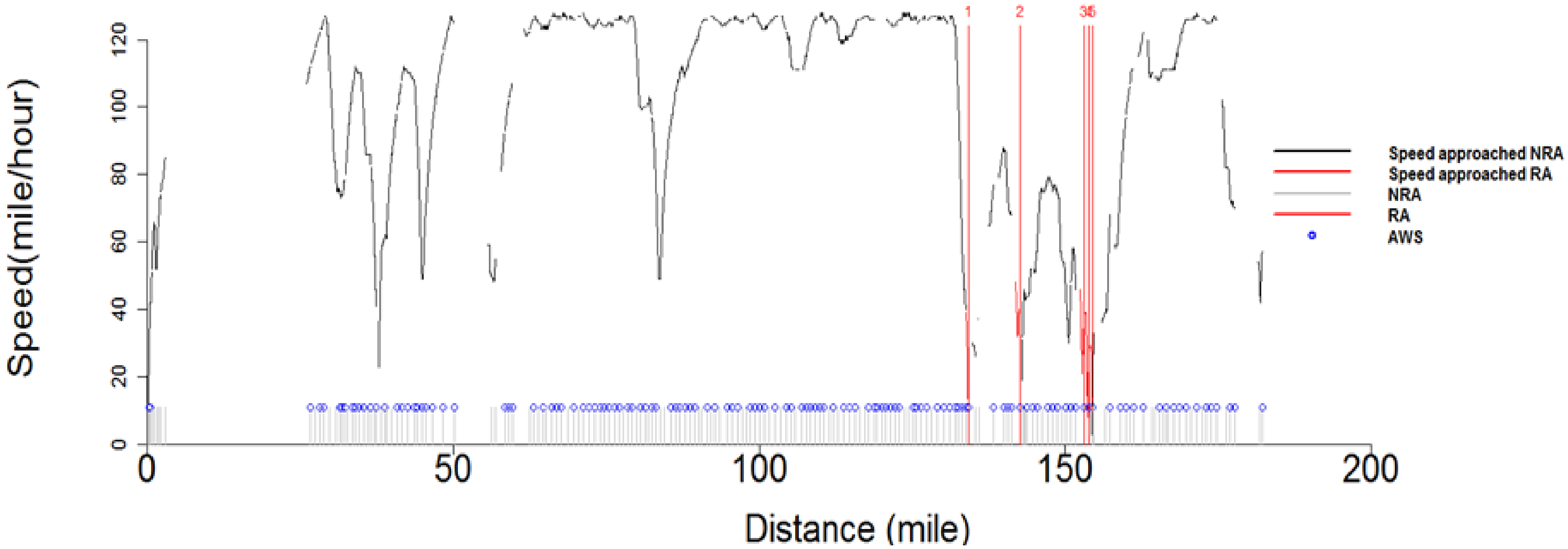

Figure 2: Service A train travel speed profile and signals the driver experienced.

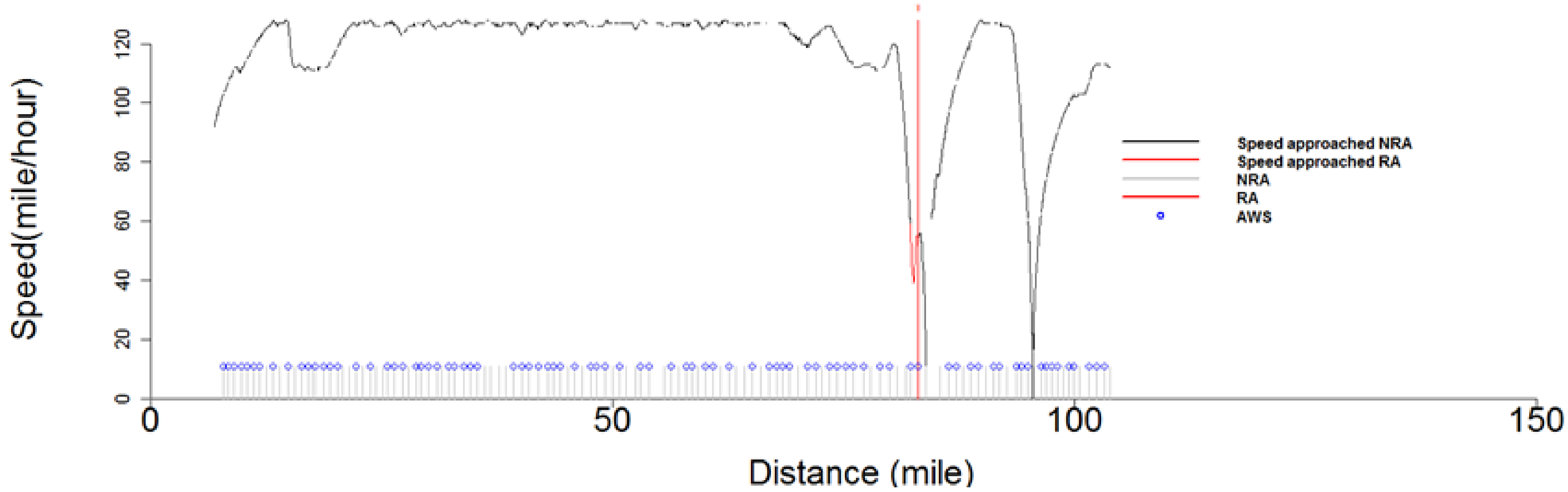

Figure 3: Service B train travel speed profile and signals the driver experienced.

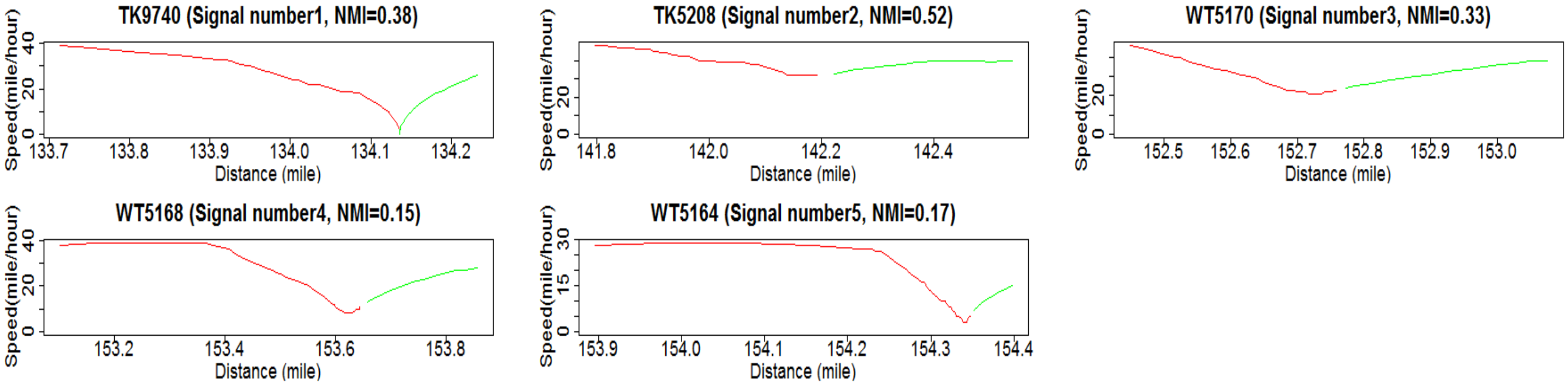

Figure 4: Service A speed profile in front of different signals.

Table 2: NMI for a number of red approaches.

Signal ID	NMI	$v_{ij(OTDR)}$(train travel speed at signal cleared) [mph]	$t_{exit(RAATS)} - t_{cleared(RAATS)}$ [s]	$v_{horn(OTDR)}$ [mph]	$t_{v=0(OTDR)} - t_{horn(OTDR)}$ [s]	Train service Status at signal change
WT5170	0.33	**24**	**37**	21	-	Proceeding
WT5164	0.17	**6**	**18**	27	-	Proceeding
SOT275	0.56	**25**	**23**	22	-	Proceeding
WK423	0.70	**47**	**33**	41	-	Proceeding
TK5208	0.52	**33**	**32**	32	-	Proceeding
WT5168	0.15	**12**	**40**	20	-	Proceeding
M398	0.41	**12**	**15**	14	-	Proceeding
MP302	0.32	**12**	**19**	15	-	Proceeding
WS4302	0.05	**5**	**49**	9	-	Proceeding
KR3345	0.62	**40**	**43**	53	-	Proceeding
MP346	0.12	**11**	**46**	NA	-	Proceeding
OS7732	0.01	**1**	**39**	NA	-	Proceeding
EL262	0.06	**37**	**137**	15	-	Proceeding
TK9740	0.38	0	55	**22**	**29**	halted
BW9284	0.21	0	34	**14**	**33**	halted
BW6280	0.21	0	87	**13**	**32**	halted
EL350	0.15	0	26	**41**	**143**	halted

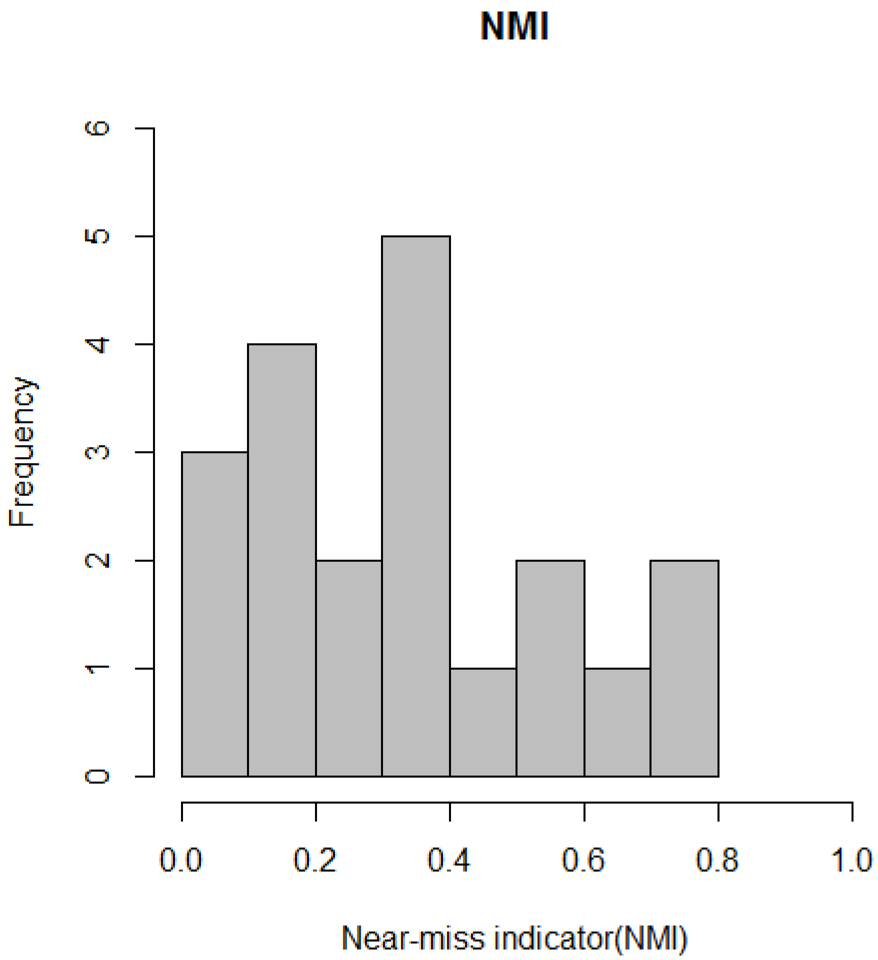

Figure 5: NMI frequency for the analyzed OTDR data.

5 Discussion

The NMI calculation presented in this paper compares the deceleration rate at a certain point, for example at AWS horn for Scenario I and at signal clearing for Scenario II. These measures only assess the final stage of braking as a driver should start braking prior to these points. The value of NMI should be between 0 to 1 to reflect the ability of the train to stop at that signal. Figure 5 shows that there is a variation among red-aspect approaches, the highest NMI value was 0.72 whereas the lowest value was 0.05, but NMI values remain below 1. The variation in NMI values may help us to understand the different ways that drivers respond to signals and whether some signals are particularly prone to high-risk approaches. It makes sense to introduce an alarm system for NMI values higher than 1; however lower values of NMI may be considered under low adhesion conditions. In addition, the analysis of

NMI is easily extended to include related information that is already in the database. Table links signal types and signal box to the NMI indicator which potentially identify SPAD-risk hot spots.

Table 3: Signal type and NMI.

Signal ID	NMI	Signal type	Signal Box
WT5170	0.33	EZ220 - SIG HEAD - COLOUR LIGHT - LED	RUGBY - Rugby SCC
WT5164	0.17	EZ220 - SIG HEAD - COLOUR LIGHT - LED	RUGBY - Rugby SCC
SOT275	0.56	EZ104 - SIGNAL HEAD - COLOUR LIGHT - 4 ASPECT	STOKE - Stoke SCC
WK423	0.72	EZ220 - SIG HEAD - COLOUR LIGHT - LED	WOKIN - Woking
TK5208	0.52	EZ220 - SIG HEAD - COLOUR LIGHT - LED	RUGBY - Rugby SCC
WT5168	0.15	EZ220 - SIG HEAD - COLOUR LIGHT - LED	RUGBY - Rugby SCC
M398	0.41	EZ104 - SIGNAL HEAD - COLOUR LIGHT - 4 ASPECT	MOTHE - Motherwell
MP302	0.32	EZ104 - SIGNAL HEAD - COLOUR LIGHT - 4 ASPECT	MANPI - Manchester Piccadilly
WS4302	0.052	EZ220 - SIG HEAD - COLOUR LIGHT - LED	WSCC – WMSC

Different data-sources do not provide unique identifiers to the same train, making the synchronization of data relatively complicated. This introduces a possibility of data source synchronization error, mainly due to the non-unique signal ID and train ID, e.g., there are a number of signals and train services that have the same ID; these errors can be identified, to some extent, by manually checking the speed profile. When inconsistencies were found, they were excluded from this study. In real-life operation, however, they can still be used for safety monitoring purposes. In addition, we found that, data-streams may be interrupted which leads to missing RAATS records. Also, the coverage of data-recording is not complete for the railways in Britain which renders the method impossible on parts of the network. Furthermore, the timestamp in both data sources (OTDR and RAATS) need to be correlated by identifying the discordancy between timestamp in both sources.

Looking toward the future; a multitude of alternative uses is foreseen for this method. Particularly when it comes to in-depth analysis of drivers' actions. For example, Kohls and Watson (2010) suggested that the more times trains approach signals at red, the more the possibility of them having SPAD. It would be relatively straightforward to add data about slippery rail-head condition in which case high-speed red-aspect approaches almost invariably lead to SPADS.

More importantly, the technique paves the way for the shift of safety management from 'avoiding things going wrong' to 'ensuring that everything goes right'. This method demonstrates the way forward for proactive safety management with data.

6 Conclusion

This paper has provided an example of how day-to-day operational data from different sources can be integrated and used to monitor frontline train operations related to safety. Such integration will encourage a mind shift of safety management to monitor safety critical scenarios of everyday performance and evaluate its risk level rather than relying on information gained after an accident has occurred. The method is very precise by identifying a scenario where several conditions coincide to create a particularly hazardous situation: hard braking AND approaching a red signal. This avoids a broad brush approach (e.g. arbitrarily looking for hard braking) and offers opportunities to add even more conditions (e.g. slippery rail-head conditions). The work demonstrates that existing data from day-to-day operation offers a chance to improve safety by measuring successes alongside errors. Trains approaching red aspects with high speed are a precursor of SPADs. Therefore, by monitoring how drivers approached red aspects, a data-driven technique provides insights in SPAD risks at individual signals.

The technique introduces the basis that can be extended to add more data sources, and to uncover additional factors affecting how the driver approaching a red aspect. Hypothetically further data sources such as localized weather conditions including sun angle, or even factors such as timetable data and train on-time running data could be included to broaden our understanding.

Acknowledgments

This work was funded by the Rail Safety and Standards Board (RSSB).

References

Balfe, N. (2017). A framework for human factors analysis of railway on-train data. In D. de Waard, A. Toffetti, R. Wiczorek, A. Sonderegger, S. Röttger, P., Bouchner, T. Franke, S. Fairclough, Noordzij, M., and Brookhuis , K. (Eds.) (2017). Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2016 Annual Conference.

Carbonari , A., Giretti, A. & Naticchia, B. (2011) . A proactive system for real-time safety management in construction sites. Automation in Construction, 20(2011), pp 686–698

DfT, Department for Transport. (2007). Autumn Adhesion Investigation Part 1: Signals WK338 and WK336 Passed at Danger at Esher 25 November 2005. Report 25 (Part 1)/ 2006-January 2007.

Health and Safety Executive, (2000). The Ladbroke Grove Rail Inquiry, Part 1 - Report. HSE BOOKS, Norwich.

El Rashidy, R. & Van Gulijk, C. (2016). Driver competence performance indicators using OTMR. In Proceedings of CIT2016 Congreso de Ingeniería del Transporte, XII Congress of Transport Engineering, pp. 354-361.

El Rashidy, R., Huges, P., Figueres, M., Harrison, C. and Van Gulijk, C. (2018) A big data modeling approach with graph databases for SPAD risk. Safety Science,110 (Part B), pp. 75-79.

European Union Agency for Railways. (2016). Big data in railways: Common occurrence reporting programme. [Online]. [Accessed 2-01-2017]. Available from: http://www.era.europa.eu/Document-Register/Documents/COR%20-%20Big%20Data.pdf

Groeger, J. A., Bradshaw, M. F., Everatt, J., Merat, N. & Field, D. (2003). A pilot study of train-drivers' eye-movements. Department of Psychology, University of Surrey, Surrey.

Green, S.R., Barkby, S., Puttock, A. & Craggs, R. (2011). Automatically assessing driver performance using black box OTDR data. Railway Condition Monitoring and Non-Destructive Testing (RCM 2011), 5th IET Conference, pp.1-5, 29-30 Nov. 2001.

Hale, A. R., Ale, B. J. M., Goossens, L. H. J., Heijer, T., Bellamy, L. J., Mud, M. L., Roelen, A., Baksteen, H., Post, J., Papazoglou, I. A., Bloemhoff, A. & Oh, J. I. H. (2007). Modeling accidents for prioritizing prevention. Reliability Engineering & System Safety, 92, pp. 1701-1715.

Hollnagel, E. (2014). *Safety-I and Safety-II: The Past and Future of Safety Management.* Farnham, UK, Ashgate.

Kohls, H.H. & Watson, R. (2010). SPAD – reducing timetable related risk. WIT Transactions on State of the Art in Science and Engineering, Vol 45, pp. 129 -136.

LI, G., Hamilton, W. I., Morrisroe, G. & Clarke, T. (2006). Driver detection and recognition of lineside signals and signs at different approach speeds. *Cognition, Technology & Work,* 8**,** pp. 30-40.

Luke, T., Brookcarter, N., Parkes, A. M., Grimes, E. & Mills, A. (2006). An investigation of train driver visual strategies. *Cognition, Technology & Work,* 8**,** pp. 15-29.

Nikandros, G. & Tombs, D. (2007). Measuring Railway Signals Passed At Danger. Proceedings of the Twelfth Australian Conference on Safety, Critical Systems and Software (2007), pp. 41-46.

McCorquodale, B., Chissick, C., McGuffog, A., Rowley, I. Bunting, A. & Page, H. 2002. Driver's Reminder Appliance (DRA) effectiveness study: Fi-nal report, Qinet-iq/KI/CHS/CAP/CR020937/2.0/2.0, QinetiQ, Farn-borough.

Miller, J.J. (2013). Graph database applications and concepts with Neo4j, in: Proceedings of the Southern Association for Information Systems Conference, Atlanta, GA, USA. p. 36.

Panzarino, O. (2014). Learning Cypher. Packt Publishing, Birmingham, UK.

RAIB, Rail Accident Investigation Branch, (2016a). Two signal passed at danger incidents, at Reading Westbury Line Junction, 28 March 2015, and Ruscombe Junction, 3 November 2015. Rail Accid. Rep.

RAIB, Rail Accident Investigation Branch, (2016b). Signal passed at danger on approach to Wootton Bassett Junction, Wiltshire 7 March 2015. Rail Accid. Rep.

RSSB. (2015). AWS and TPWS Handbook. RS/522 Issue 3 December 2015.

RSSB. (2013). Category A SPAD and TPWS activity report. RSSB.

RSSB. (2015). AWS and TPWS handbook. RSSB.

RSSB. (2004). The Rail Safety and Standards Board, Driver Error Data Collection Project: Final Report, RSSB.

RULE BOOK: GE/RT8000/TW1 (2015). Train Driver Manual. Retrieved December 14, 2017,from: https://www.rssb.co.uk/rgs/rulebooks/GERM8000-traindriver%20Iss%202.pdf.

Sumwalt, R. L., Thomas, R. J. & Dismukes, K. (2002). Enhancing Flight-crew Monitoring Skills Can Increase Flight Safety. *55th International Air Safety Seminar.* Dublin, Ireland: Flight Safety Foundation.

Walker, G. and Strathie, A. (2015). Leading indicators of operational risk on the railway: A novel use for under utilized data recordings. Safety Science, 74, pp. 93-101

Zhao, Y., Stow, J. & Harrison, C. (2016). Estimating the frequency of trains approaching red signals: a case study for improving the understanding of SPAD risk. IET Intell. Transp. Syst. 10(7), pp. 579:586.