

Back to basics: train protection

Questions for the Certificate (Module A) will not be set on the text which has been struck through

Effective from October 2023 exam



David Fenner



This, the fifth article on 'back to basics' themes, covers the subject of train protection. The fundamental requirement of a railway signalling system is to ensure a train movement can be completed with the minimum of risk from human error, while at the same time allowing as many trains as possible to run on a given infrastructure. Much of the focus of a signalling engineer is on the design of the infrastructure control system that ensures safe routing of the trains and supports the person responsible for that activity, the signaller. But the other key operative is the driver.

The history of railway signalling is very often about the development of systems to prevent the wrong signal being given to a driver and the evolution of the technology that underpins that function. Relatively little relates to supporting the driver, although it would be wrong to assume there is none. It is worth considering the different roles to see if there is any underlying reason for this. The signaller is dealing with a fixed set of equipment that has a limited range of states and thus a limited set of options. The requirement is to set specific routes at given times to permit the safe passage of a train, and to be reminded which of those actions have yet to be completed. The driver on the other hand is constantly making judgments about the safe movement of the train based on information from their own route and train performance knowledge, signs at the trackside and the signals seen. And in the case of signals seen these can and often will be changing minute by minute as other traffic moves. So the computational task of assisting a driver is significantly more sophisticated than that required to support a signaller. Maybe this is why train protection systems seem to have evolved more slowly, with the sophistication of train protection requiring computing power akin to a processor not a logic box.

"The driver is constantly making judgements about the safe movement of their train"

Supporting the driver starts with providing clear and concise signal indications so there is no confusion in the meaning and limited opportunity to read the wrong one. One can wonder how drivers could see a dim oil lamp flickering behind a coloured lens as they approached many of the mechanical signals still in service, at least in the UK until as little as 15 years ago. Many years ago, drivers had the advantage of a dark sky with the general absence of town and street lighting. Now of course most signals are colour lamps but often against a much brighter background. The problem is how do we help the driver respond to those signals and what other support do we need to give them to limit the number of accidents. There are two main errors a driver can make:

- Failure to control the speed of the train correctly.
- Failure to stop at the defined End of Authority (EoA).

It is these functions a train protection system must address.

The remainder of this article will start by examining the various types of train protection and how those may affect the performance of the railway. It will then provide a generalised review of the development of systems to outline the key requirements with examples of systems from the UK and other countries.

Warning or protection

Within the overall scope of train protection there are two main sub divisions, warning systems and protection systems. A warning system will normally be associated with the location at which braking should commence whereas a protection system will aim to stop the train at or soon after the End of Authority (EoA). A warning system will alert the driver to take note of information presented from the trackside and may initiate



Supporting the driver by relaying accurate information in the cab and reducing the possibility of mistakes being made has been a long running focus of the command, control and signalling industry. The view may have improved from the steam era to the modern day (as shown above) but increased speed and traffic complicates the situation.

Photos Westinghouse B&S Archive/
Chippenham Museum and Shutterstock/aapsky.

"Train stops reduce capacity because extended overrun distance increases the headway between trains"



braking action if the driver fails to acknowledge the warning. After acknowledgement the train is usually solely under driver control. A protection system will enforce a brake application should the system detect the train is being driven outside safe parameters. There are several sub-divisions of protection systems. It should be noted there is overlap between some of these classifications.

Train stops

Probably the most basic but effective form of train protection, train stops are typically mechanical devices that trigger a brake application if passed when in the 'protecting' position. They are usually associated with a stop signal. A train stop will ensure the train comes to a halt at braking distance beyond. To be effective in eliminating accidents they require an extended protected length of track beyond the signal, which limits application to lines with relatively low top speeds (around 80km/h) and good brake performance – typically metros. There are two fundamental weaknesses with such systems; firstly, they reduce capacity because the extended overrun distance increases the headway between trains, and secondly, they do not address the issue of overspeed at other locations. There have been applications where timed release of train stops is used to monitor speed, especially approaching buffer stops or high-risk junctions, but they are difficult to implement successfully.

Intermittent train protection

Intermittent train protection is provided at those locations assessed to pose significant risk. These are usually junctions with substantial traffic or a particular layout that creates additional risk, for example where there is a chance of head-on collision. As a train approaches such sites it passes through a speed trap and if above a pre-set speed will experience a forced brake application.

In more recent systems it will be fed information about the current conditions and then calculates the safe speed profile to be followed.

Modern systems of this type usually require the fitment of an Automatic Train Protection (ATP) system to the trains but avoid expensive

investment in a comprehensive trackside ATP installation. Many of the systems introduced prior to the turn of the century (2000) can be considered to be of this form.

Train protection with intermittent update

These systems continuously monitor the movement of the train to ensure it is within the safe envelope. The driver is provided with an indication of the current maximum safe speed but otherwise drives to the signals and signs observed at the trackside. The train picks up information about the infrastructure ahead including the traffic conditions at intermittent intervals, usually from a balise or beacon or from inductive loops laid in the track. The intermittent intervals are usually associated with a signal, to facilitate the inclusion of information about the extent of any current movement authority and any speed limits associated with that authority.

These systems provide full supervision of the driver at all times and will intervene to reduce the train speed back within the safe envelope prior to returning control to the driver. The exception to this is in the event of a Signal Passed at Danger (SPAD) when a 'trip' state is entered and the train is brought to a halt. The systems normally include roll back protection, a function of most ATP systems, to prevent the train rolling away when brakes are released but traction power is not applied.

Because the system only receives intermittent updates, the system may restrict track capacity because the train must be driven according to the last update received and may thus continue to enforce a reducing speed even though the signal ahead has changed to a less restrictive aspect. A means of partially overcoming this constraint is to have 'infill' in the form of loops or additional balises which augment the update frequency, especially approaching critical signals.

One advantage of such ATP systems is, because the train is being driven on conventional signals, the system can be overlaid on the railway to enhance safety. It is thus possible to partially equip either or both track and trains with only

selected portions of the route and chosen trains being protected. One disadvantage with this is that an unfitted train which SPADS, may encroach on the path of a fitted train, resulting in a serious accident. Thus, this arrangement would ideally only be used as a migratory configuration.

Train protection with continuous update

These systems have the facility to provide updated movement authority information at any time. Consequently, they are often part of an in-cab signalling system because the driver is continually fed information on the current safe speed and is no longer required to watch for lineside signals or signs. There may be other reasons, outside the scope of the signalling system, for the driver to watch the passing trackside and one of the challenges is to ensure the right balance is struck between watching the displays in the cab and watching the line.

Another challenge with these systems is that the driver is given advice on the safe speed profile based on the train parameters encoded in the system. If these parameters are wrong the driver may be advised to brake later in the journey than is desirable, with the risk of passing the End of Authority (EoA) – the equivalent of a SPAD. This is of particular concern for variable formation trains, especially locomotive-hauled freight trains, but is also relevant at times when adhesion is poor.

The major advantage of such systems is the ability to enhance the capacity of the infrastructure. Because the train and infrastructure are in regular contact, and with a safe speed display in the cab, the limitations of fixed lineside signalling with a limited number of aspects can be removed. Thus, a fixed block railway is no longer limited to say four or five aspect signals, meaning closer running is possible and, more importantly, release of trains held at a converging junction can happen significantly earlier with very short block sections. This form of protection is a corner stone of moving block systems and the heart of Communication Based Train Control (CBTC). Metro railways have been able to gain significant uplift in service frequency as a result of such application, for example the Victoria line in London increased from around 30 to 36 trains per hour in the peak as a result of implementing CBTC. Continuous update train protection thus offers enhanced performance and options for reduced trackside equipment as well as improved safety, making the business justification measurably easier.

The advantages of in cab signalling and increased capacity by removal of lineside signals usually means either all trains must be fitted with a suitable system or the configuration of existing signals needs to be amended to allow changes in operational rules. This poses a significant cost hurdle before such benefits can be gained.

Warning systems

Warning systems are usually designed and positioned to alert the driver to the presence of a signal or sign at the trackside. The Automatic Warning System (AWS) in the UK and used

elsewhere is a typical example. If a distant signal is approached with a clear indication the driver receives a bell (or gong) sound and is free to continue. Should the signal be at caution the sound is a buzzer (or horn) and the driver must acknowledge this both to turn off the sound and prevent a brake application. After acknowledgement the driver is responsible for the continued motion of the train. These basic functions are almost identical to the French "Crocodile" system.

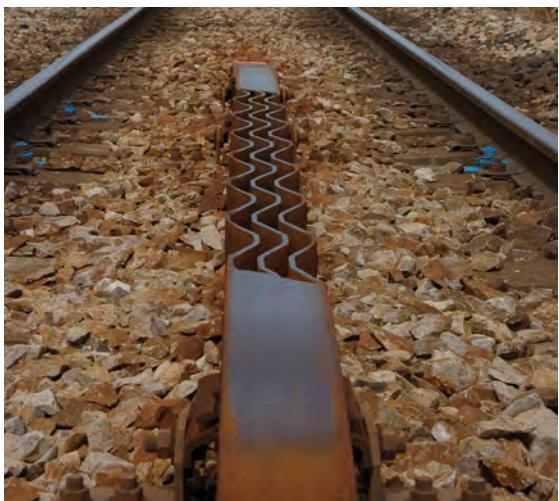
~~It is interesting to note how the technology of this system has changed. The basic system was developed by the Great Western Railway, trialled in 1906 and implemented shortly afterwards across their network. Initially the communication between track and train was a ramp in the track with an electrical feed to indicate clear and no electrical charge when at caution. This follows the "fail safe" principle because the absence of electricity could arise in several failure situations whereas the mechanical lifting of the contact shoe would fail much less often. The problem with a mechanical contact system is wear and tear especially as speed and frequency of use increases as was inevitable when Multiple Aspect Signalling (MAS) became common. The form of AWS implemented by British Railways from the 1950s uses magnetic induction to convey the signal status to the train. During the 1950s when the system was being installed, semiconductor devices were new so the equipment on the train used relays. Today the logic is performed by a PLC and the magnetic field detected by a Hall Effect sensor.~~

~~AWS uses a permanent magnet centrally located between the rails to initiate the onboard system and provide the driver with a warning if the signal is at caution or stop. An electromagnet is positioned immediately afterwards where a signal may show a line clear aspect. This second magnet is energised when the signal is clear (green) and effectively cancels the warning, instead sounding a bell in the cab (see Figure 1). The French "Crocodile" system conveys the two messages from the infrastructure to the train by different voltages applied to the "Crocodile" ramp collected by a wire brush on the driving vehicle mechanically contacting the ramp.~~

~~To ensure the two successive magnets are detected at low speed the system delays output to the driver for approximately one second after passing the permanent magnet. This provides time for a slowly moving train to detect the electromagnet prior to initiating the warning sequence. It then gives the driver around three seconds to acknowledge a warning and observe the signal. If no acknowledgement is received the brakes will be applied. On passing the signal displaying the caution aspect the train should stop at or before the associated stop signal. Thus the AWS trackside installation is placed around 4 seconds before the signal, which is normally interpreted as about 180m.~~

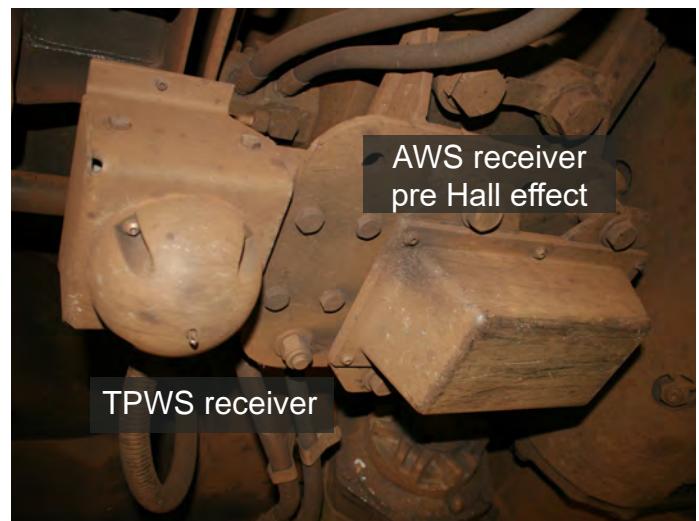
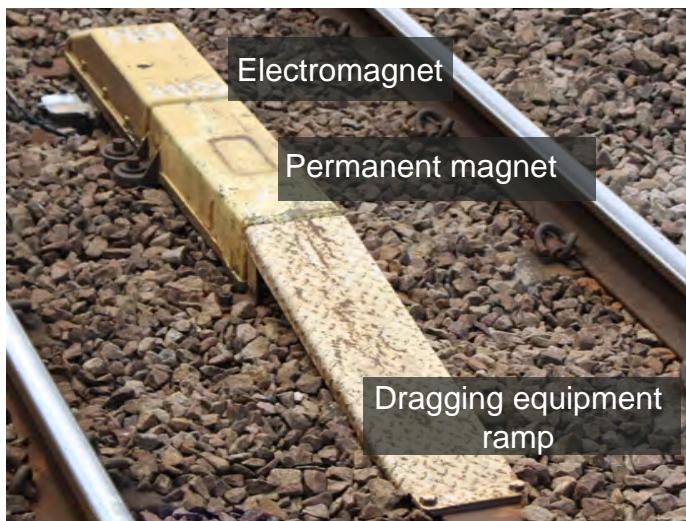
~~A key technical challenge for AWS arose with its application to third rail electrified infrastructure.~~

"If these parameters are wrong the driver may be advised to brake later than is desirable"



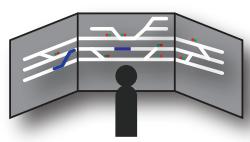
The "Crocodile" system conveys two messages to the train by different voltages fed to the ramp. The right hand photo shows Eurobalises, TBL antenna and Crocodile together in a Belgian application.

Photos Wikimedia CC-BY-SA Charlotte Noblet (left) and François Melchior.



AWS. Left trackside magnets, right, train-carried equipment.
Photos David Fenner.

- 1) Signal green. Train responds to both magnets, sounds bell to driver, no action.
- 2) Signal NOT green. Train responds to one magnet, sounds horn to driver. If driver does not respond to horn brake applied after 2-3 seconds. If driver responds no action by system.



Control Centre
Controls points and signals.
Watches train position from
train detection sections.



Figure 1 – AWS application Trains with AWS or full TPWS can share the line.

Where high current DC traction systems are used there is a risk of cross track cables creating 'stray' magnet fields. In this case the AWS detectors on the trains were desensitised and extra strength magnets provided for the trackside AWS functions to minimise the risk of false AWS information from the stray fields.

Another challenge is the increasing use of bi-directional signalling. With the magnets in the centre of the track, trains in either direction will detect the fields. Thus, trains passing over an installation intended for the opposite direction of movement will experience a warning since the permanent magnet will be the last item passed. On tracks where this is an infrequent occurrence a sign may be installed to remind the driver this warning does not apply to them. In locations where such movement is more frequent or would result in the need to regularly acknowledge irrelevant AWS warnings, the permanent magnet is surrounded by an electromagnet coil which suppresses the permanent magnetic field when a reverse direction movement is signalled. Thus, outputs are required from the interlocking system not just to indicate when the signal is showing a green aspect but also when a reverse direction route is set. Occasionally the signalling arrangement may enable one AWS installation to be used for two signals reading in opposite directions. In this case the permanent magnet will have two electromagnets one either end associated with the relevant signal.

Multiple aspect signalling has resulted in most signals being equipped with AWS meaning the need for driver acknowledgement can become a semi-automatic reaction, especially when operating a congested railway on restrictive aspects. Indeed, this caused the Southern Region of British Rail to experiment with a more sophisticated form of Signal Repeating AWS (SRAWS) during the early 1970's but in the end this was not adopted. The problem of instinctive response is exacerbated by the application of the system to severe permanent and all temporary speed restrictions. These requirements arose after

"Another challenge is the increasing use of bi-directional signalling"

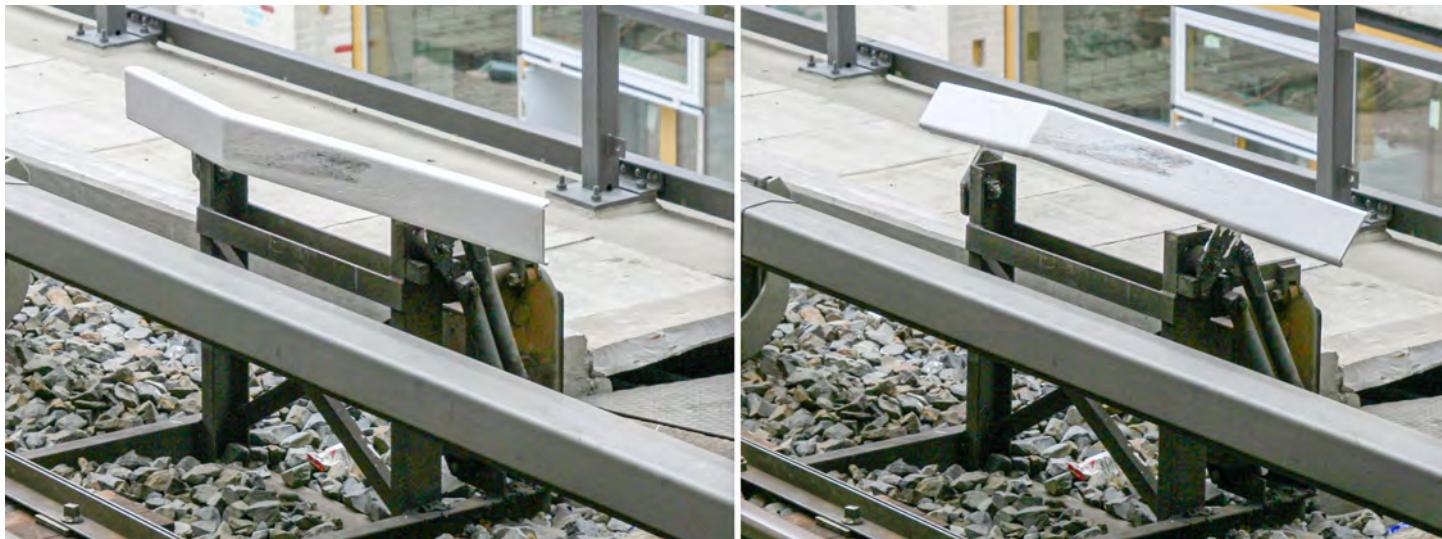
"Train stops are still a mitigation measure rather than a control system"

accidents caused by excess speed. The need to continue to reduce the harm caused by accidents together with these human factor weaknesses subsequently resulted in the development of the Train Protection Warning System (TPWS).

A feature of AWS and many similar systems is they give the driver one of two indications at every signal. Thus, any failure of the train borne equipment is quickly revealed to the driver and can be managed through appropriate procedures. In the case of AWS, the use of a permanent magnet ensures that a complete failure of the trackside installation is very remote and similarly the driver will note a wrong indication. It should be understood, AWS is not a fail-safe system with a high SIL rating. It is therefore important that faults are reported promptly and suitable measures taken to operate safely after discovery.

Basic protection systems

The earliest protection systems were train stops placed on the track at the location of a stop signal, as previously described. When the signal is showing a stop aspect an arm is raised typically 50–60mm above rail height. This engages with a trip cock on the train which opens a vent on the brake pipe forcing a brake application. The train will be passing the signal at danger before the brakes are applied in which context this is a SPAD mitigation measure rather than prevention. When the signal is clear the trip arm is lowered below rail level. Train stops were normally provided on metros operating in tunnels because of the significant risk of a collision in such locations. Extension to other locations has followed. To gain full protection requires a long enough overrun to bring the train to a halt from the maximum attainable speed prior to the brake application. For metros this is usually an associated design constraint limiting capacity and creating some operating constraints such as restricting speed through closed stations when the overlap at the next signal is based on the maximum attainable speed from a standing start. For main line railways the concept is often not practicable due to higher speeds and extended braking distances.



"Classic" trainstops on the Berlin S-Bahn, photo taken in 2007.
Photo Kabi/Wikimedia.

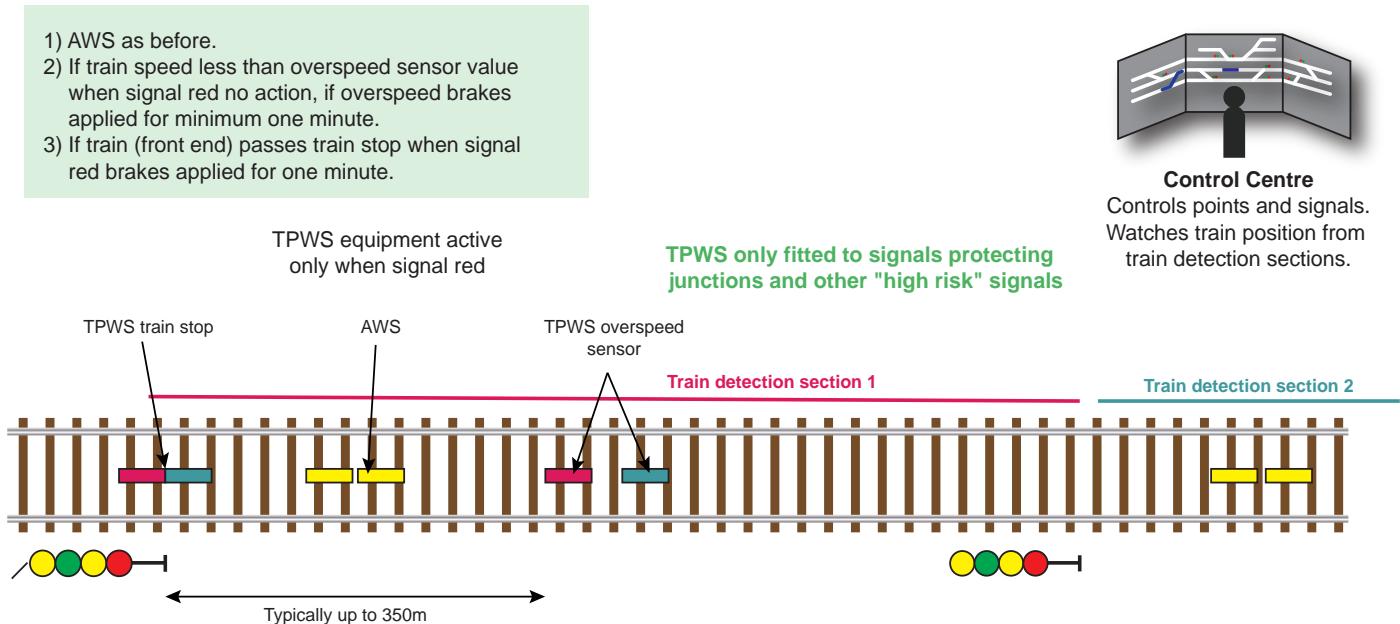


Figure 2 – Typical TPWS installation.

~~Train stops have moved on significantly from this technology, but they are still a mitigation measure rather than a control system. An early advance on mechanical contact train stop systems was "Indusi", first installed in Germany and surrounding countries in the 1930s and revived in the 1950s. Several enhancements have subsequently occurred leading to the system known as PZB90. Three tones are generated on the driving vehicle and interact with passive tuned coils mounted as required trackside. The tuned coils cause the relevant frequency on the motive unit to resonate resulting in an impedance change and thus activating the required action. The 2000Hz tone is a train stop forcing a brake application and is will typically resonate on passing a red signal. The 1000Hz tone is typically associated with a caution signal and requires a driver acknowledgement followed by a speed reduction within a set time whilst a 500Hz tone requires a lower maximum speed and is often placed shortly before a red signal.~~

~~Indusi tuned coils are installed on the sleeper ends so they are naturally only detected by a vehicle travelling in one direction and do not need suppression for opposite direction movements. TPWS in the UK has related functionality in that tones, in this case six of them around 65kHz, are transmitted by track mounted loops this time centrally mounted in the track. Three frequencies, two arming and one trigger, are used for the normal direction of movement and the other three for trains in the opposite direction on bi-directional track. These provide the functions of a 'speed trap' for a train travelling faster than deemed safe and a train stop (Figure 2). Should either the speed trap or train stop function be activated a full brake application is initiated and maintained for one minute, bringing the train to a stand. The location at which the train stops is a function of the train brake performance and the initial speed prior to the brake application. There is no certainty this will be in a safe location.~~

"TPWS is only provided at locations that are deemed high-risk"

~~TPWS is considered to have a high probability of stopping a passenger train within the overlap provided the approach speed is not greater than 75mph (120km/h) and adhesion can deliver the braking effort of the train. Where approach speeds are greater than 75mph and risk is assessed to be high additional overspeed loops may be provided earlier on approach to the signal (Figure 3 overleaf). The system substantially mitigates the risk with a SPAD but does not eliminate it.~~

~~The speed trap function of TPWS is determined jointly by the train and the infrastructure. The train has a preset on board timer. At the trackside the arming and trigger loops are separated by a predetermined distance corresponding to a defined speed based on the standard timer setting. Current implementations assume two standard timer settings, one for passenger trains and a longer setting for freight trains. Thus, the speed trap will activate if a train passes above a predefined speed based on its 'type'. This gives rise to not infrequent cases where a marginal overspeed results in a heavy brake application that was probably 'unnecessary' in terms of the train being able to stop where required. This can cause drivers to dislike the system which is a concern as it may lead to misuse.~~

~~TPWS is an intermittent train protection system because it is only provided at locations that are deemed high-risk. Those are defined as signals protecting a convergent junction and selected other signals (where the SPAD risk is assessed as high), buffer stops and severe speed reductions where the approach speed is or exceeds 60mph (approximately 100km/h).~~

~~The TPWS system is active to stop a train so a key requirement is confirming the trackside equipment is functioning when required. This is done by monitoring the current fed to the transmitting loops. Should this fall outside the intended parameters an indication is given to the supervising signaller and the signal in rear held~~

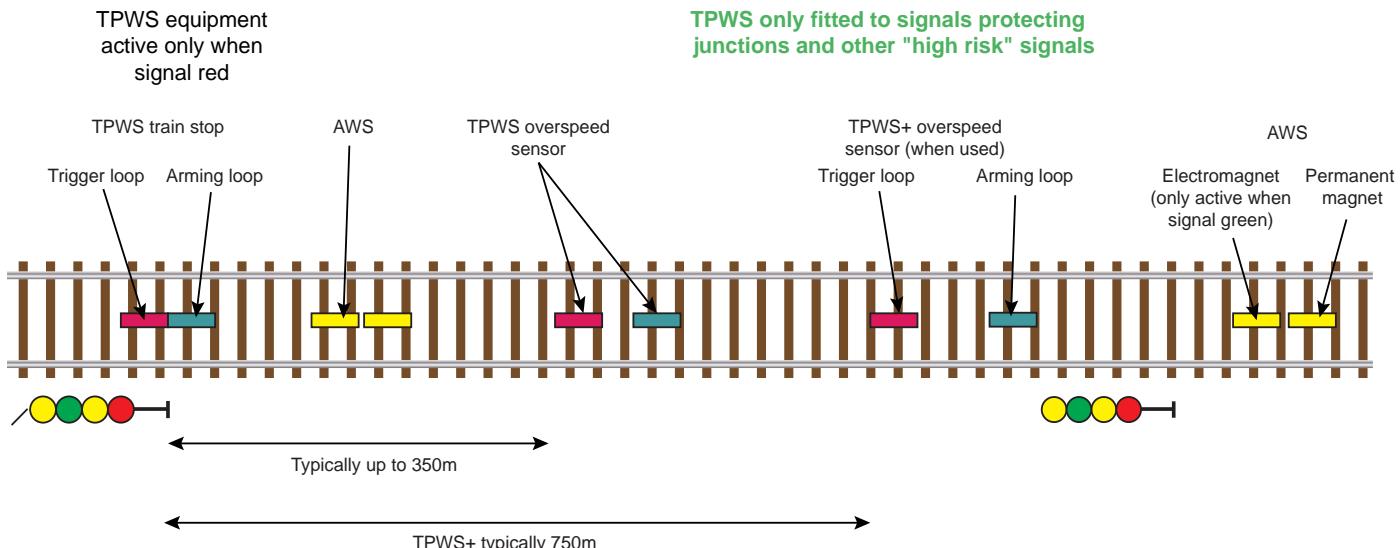


Figure 3 – Typical TPWS installation including AWS.

~~at danger. Similarly, the on-board equipment is self-diagnostic to confirm it is functional at the start of a journey and in some cases intermittently thereafter. As with AWS this system was not designed in accordance with SIL 4 requirements, so safety needs to be controlled by suitable procedures in the event of a fault.~~

~~Neither PZB90 nor TPWS can prevent all accidents as they are intermittent systems. Both systems are provided with facilities to allow the driver to reset the system after a brake intervention and continue with the journey. Such reset is required to be conducted under controlled circumstances involving communication with the signaller or a controller. In both cases there is evidence of drivers sometimes resetting without following the correct procedure, with the potential for a serious accident.~~

Full ATP systems with intermittent update

"Full ATP systems continuously monitor the movement of the train"

Full ATP systems continuously monitor the movement of the train and intervene with a brake application should the speed go outside the safe envelope. An indication of a suitable speed is given to the driver, this speed usually having a small margin beneath the safe speed envelope. These systems are normally designed in accordance with SIL 4 requirements.

To enable the train to continuously calculate and monitor the safe speed profile information is needed about several parameters:

- Infrastructure profile.
- Distance to the end of the movement authority (i.e. stop signal) (EoA).
- Current speed of the train.
- Distance travelled.
- Braking performance of the train.
- Maximum train speed.

The infrastructure profile must reflect the permitted speed, including any temporary speed limits, covering at least the distance within the current movement authority. This includes speed limits over junctions which may vary depending on the route allocated. It also needs to include the gradient profile over the same distance as changes to gradient can have a significant effect on stopping distance.

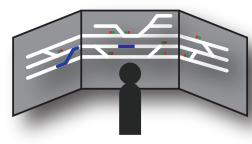
The location of the EoA either needs to be known or at least it is beyond the range that currently needs to be considered (e.g. when driving on green signals or the equivalent). To calculate the ongoing speed profile, it is obviously necessary to know the current speed and how far the train has travelled to enable the remaining distance to any speed change to be determined.

The braking performance of the train is a key parameter in performing the safe speed envelope calculation.

The data about the infrastructure including route set and signal status is usually fed to the train intermittently as data telegrams by balises or beacons, most of which will be associated with a signal (Figure 4). Some systems use inductive loops laid in the track to pass the relevant telegrams. Infill may be provided where the intermittent transmission of data has a significant impact on operational capability.

The need to know speed and distance travelled results in some fairly complex configurations of equipment on board the train. The wheel to rail interface is notoriously challenging with relatively low adhesion. This makes simple odometry inadequate for the purposes of ATP especially as the objective is a system with high safety integrity. Slip (during acceleration) and slide (during braking) are relatively likely, both creating erroneous information if reliance is placed on wheel rotation. Some systems use an unpowered and under braked (or unbraked) axle as the sensor input.

- 1) Beacon tells train 'colour of signal', permitted speeds distance to next beacon, gradient etc.
- 2) Train computer calculates maximum safe speed, advises driver by display in cab.
- 3) Driver obeys lineside signals and signs.
- 4) If driver exceeds safe speed brakes applied.



Control Centre
Controls points and signals.
Watches train position from
train detection sections.

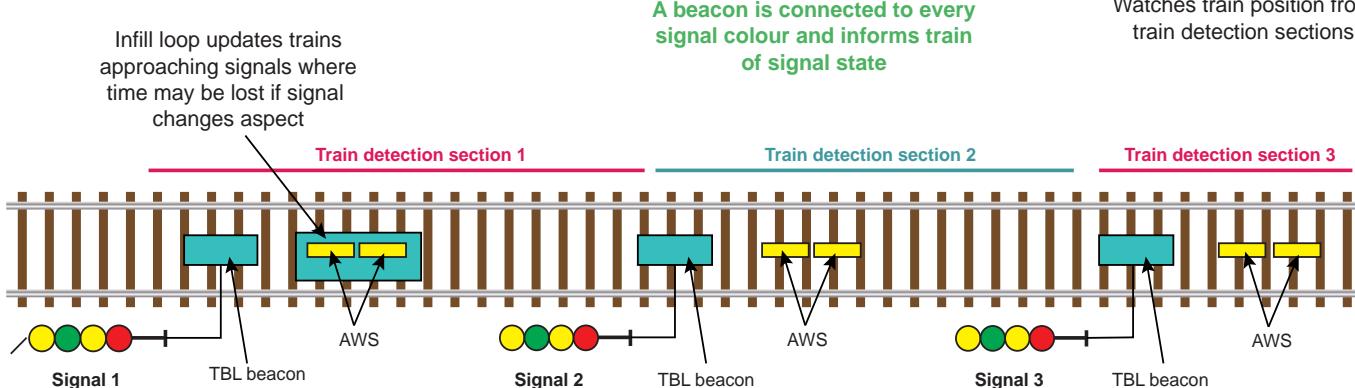


Figure 4 – ATP with intermittent update (as used by Network Rail between London and Bristol).

"Odometry for ATP is often a mix of systems to counteract the weakness of any one of them"

Another risk with wheel rotation is the slow but steady change in size as the wheel wears. Doppler radar is sometimes used, but is not without challenges especially in damp or snowy conditions and of course like wheel rotation it is subject to a steady build-up of position error unless corrected. Use of a Global Navigation Satellite System (GNSS) offers another means. However, it rarely covers all conditions, examples being tunnels and urban or other 'canyons' where lost or reflected signals can affect the calculated position. Thus, odometry for ATP is often a mix of systems to counteract the weakness of any one of them and to support the determination of the margin of error on the actual estimated position which will tend to increase over time and distance travelled. Balises or beacons on the infrastructure are frequently used as a reference point to correct accumulated error. Not surprisingly odometry problems are potentially the biggest cause of train equipment failure for some ATP systems.

A further result of the challenge with odometry is the need for a "release" speed with many ATP systems. Because the train cannot know exactly where it is the ATP system will assume worst case error and therefore bring the train to a halt a short distance prior to the current EoA. This is a significant problem if the next update function is located close to the EoA and in the case of the EoA being at the end of a station platform when, for station purposes, it is important to get close to the platform end. To counteract these challenges, it is normal for the ATP system to return full control to the driver on final approach to the EoA provided the train is then operated below a low "release" speed. This of course means there needs to be a short safe overrun beyond the EoA should the train pass beyond the EoA, when a "trip" state will be entered to bring the train to a stand.

Braking capability is the final parameter the system needs. This is generally relatively simple for fixed formation multiple unit trains, although a given axle may under-perform due to a maintenance

issue. It is a far more complex challenge for variable formation locomotive hauled trains, especially freight trains. The usual arrangement is to request the driver to enter or select brake data as part of train preparation. One challenge in defining the brake performance is whether the system needs one or both the service brake and emergency brake parameters, and ensuring the rolling stock and signal engineer are discussing the same function! The rolling stock engineer will often regard the emergency brake as the brake of last resort (i.e. a vented air pipe) which may result in a lower deceleration rate because electric braking and wheel slide prevention systems may be disabled. If the emergency brake rate is lower than the service brake rate it will tend to dominate the operating profile because it will require the train to be slowed earlier than the service brake. Thus, it is important to be clear between parties exactly what the parameters mean and how they will be used.

The inputs from the above data sources are typically fed to a multi-channel processing system prior to being presented to the driver as a permitted speed indication and when necessary to the brake activation interface.

Full ATP systems with continuous update

These systems have the same basic functionality and challenges as described for intermittent update systems, but the data about infrastructure condition is able to be updated at any time. Such systems allow in cab signalling as the driver can be informed of an update to signalling conditions via the cab display.

Early systems achieved continuous update by transmitting codes to the train via track circuits. In practice the number of codes capable of transmission this way is severely restricted. So usually the train was given a maximum running speed and a target speed to achieve at the end of the section. Information such as speed restrictions



Opposite ends of the coded track train protection speed spectrum:

Left, the LGV-Est line in France, equipped with TVM430.

Right, London Underground's Central Line.

Photos Shutterstock/olrat, Shutterstock/Tatty.



or gradient would have to be used as part of the design process when configuring the system for the railway. This method of providing continuous update is limited to lines using rolling stock with very similar brake performance, or at least has to be based on the worst brake performance train permitted. It was the backbone of the original London Underground Victoria Line automatic train operation system from 1965 until recently and is the basis of the signalling on many of the French LGV lines signalled using TVM 430. A significant challenge with such systems is the need to ensure the train can detect the track circuit code prior to shorting the track circuit current as part of the train detection role. Thus, assuming the detector is at the front of the train, it is necessary for all track circuit feeds to be ahead the train which means that reversal of feed and detector functions is necessary for any lengths of track used reversibly.

Another early method of providing continuous update was via a loop usually between the rails which radiates data to the train. The German LZB system is perhaps the best known of this form and is also the parent system for several others of this type (e.g. SELTrac). In this particular case the data rate is low by modern standards (1200 baud up link to the train; 600 baud down link). Thus, the data volumes handled between track and train are relatively low (circa 85 bits per telegram). As a result much of the calculation is performed in the LZB control centre and limited target speed data is passed to the train. The loops in the four foot are crossed every 100m which both reduces interference and provides a distance monitoring function to the train by reversing the phase of the radiated signal.

The system is in relatively widespread use through Germany and Spain, especially on high speed lines, with other applications in Austria and elsewhere.

The most recent systems in this category are urban CBTC train control systems and ETCS Level 2 and above, both of which use radio systems to transmit data between track and train.

The major advantage of mobile data packet systems is the much larger and reliable data bandwidth, meaning more of the processing can be performed on the train and be relevant to the specific train, resulting in better optimisation. They also both open the way for the train to inform the infrastructure of its position and its operating performance.

CBTC

Many current resignalling projects for metro railways are based on Communications Based Train Control (CBTC). As the title suggests this is a highly integrated form of train control covering all the important functions of train movement including route setting, train location, movement authority often based on moving block, train supervision and train driving. In the latest systems CBTC is often used on a railway working under Automatic Train Operation (ATO) or a higher grade of automation with the objective of providing the maximum achievable capacity on the routes fitted.

CBTC systems incorporate continuous ATP protection using radio communications between the train and the trackside installation. The actual form of the radio link, which is also used for other aspects of the CBTC system, varies between both supplier and application and may include Wi-Fi and 4G/LTE, although the majority of new systems are likely to use 4G/LTE. The detailed methods of implementation are proprietary to the manufacturer.

Most systems of this type have been implemented on individual lines of a metro or on people mover systems (e.g. at airports). This has the advantage of a limited range of rolling stock making the train parameters consistent and thus maximising the capacity gain from moving block.

Two disadvantages are the need, especially on a high capacity metro, of having a highly reliable system requiring layers of redundancy, and the proprietary nature of the systems from each manufacturer which for commercial reasons restricts the use of such systems on a large network of lines.

"CBTC systems incorporate continuous ATP protection using radio communications"

"The advent of through international trains such as Eurostar required the trains to be fitted with multiple train protection and communication systems"

China's CTCS system shares many concepts with ETCS but has developed very quickly in order to keep up with infrastructure developments across the country.

Photo Shutterstock/kikujungboy.

ETCS

Until the around the 1990s most EU countries had their own train protection systems as well as signalling and communications arrangements. This resulted in relatively closed markets for signalling products and required different systems to be available on any driving vehicle that might cross from one country to another. The advent of through international trains such as Eurostar required the trains to be fitted with multiple train protection and communication systems and for the drivers to be trained in the differing signalling arrangements. The opening up of markets prompted the EU to propose a common standard for rail traffic management and especially signalling, train protection and radio communications across the EU states which culminated in the definition and development of ERTMS including ETCS and GSM-R for track to train communication.

ETCS provides three primary levels of implementation together with others to suit specific needs.

Level 1 is essentially an intermittent update ATP system associated with existing lineside signals with the same facilities and restrictions in terms of 'infill'. The primary track to train communication media is the balise with infill loops an option.

Level 2 adds a Radio Block Centre (RBC) and provides continuous update of movement authority and all associated infrastructure data by radio. For ETCS the currently specified radio link is via a continuously open GSM-R call. However, this is likely to change as train radio moves to packet switched (4G/LTE or 5G). In the interim GSM-R may be supplemented with General Packet Radio Service (GPRS), a 2/3G packet oriented mobile data standard, to support ATP in busy areas. It therefore has the functions of a continuous update ATP system. Lineside signals can be removed but Level 2 retains trackside train detection and interlocking functions. Such a system can provide the basis for Automatic Train Operation (ATO) and has been implemented on the Thameslink core through central London.

Level 3, which is not yet fully developed, will enable train detection to be based on data received by the RBC from the trains in the area. It will therefore facilitate the reduction or removal of lineside train detection equipment and facilitate the introduction of short virtual blocks or perhaps moving block. A significant challenge at present is from variable formation trains (e.g. freight and locomotive hauled passenger trains) and proving they are complete at all times.

Other ETCS levels also exist, including those working with a national ATP system through a Specific Transmission Module (Level STM), and Level 0 when an equipped train is operating over unfitted infrastructure and is thus subject only to train ceiling speed supervision.

In addition to these primary levels there also exist options for Limited Supervision which effectively provides train protection only at critical locations and thus corresponds to intermittent train protection. Finally, there are proposals for ETCS Hybrid Level 3 which provides virtual blocks as well as conventional train detection, delivering capacity benefits for trains that are Level 3 enabled, without having to bar the route to Level 2 or unfitted trains. An article about ETCS Level 3 and Hybrid Level 3 was published in the April 2017 issue of IRSE News.

Positive Train Control (PTC)

In North America there is currently a programme to implement Positive Train Control (PTC) on all intercity passenger lines, and on freight lines carrying more than 5 million tons per annum of inhalable toxic substances. PTC is a form of ATP in that it requires the prevention of collision due to passing the end of authority or passing through misaligned points and speed supervision to prevent derailment at speed restrictions whether permanent or temporary. There is no one mandated solution but there is the requirement of interoperability between systems as many vehicles travel over several networks. It also requires provision of arrangements to protect track workers. Whilst several systems were proposed most installations have settled on one of three systems and all will operate with each other.





In a similar way that CTCS was created to meet the needs of China, PTC was developed to meet the specific needs of North America's railroads. Union Pacific has been one of the adopters of this technology.

*Photo Shutterstock/
AngelDiBilio.*

The PTC system was mandated by law signed in 2008. The original law required the system(s) to be in place by 2015 but costs and challenges in implementation mean this date has been pushed back to the end of 2020. The range of operating conditions from metro and commuter railways to long heavy haul freight lines means variety of solutions is too great to discuss here and could well be the subject of a separate article.

About the author ...

David joined the S & T department of the Southern Region of British Rail in 1968 as a student engineer, with his first appointment in the telecommunication design office. He supported the testing and commissioning of passenger information systems, including the automatic operation of platform indicators at London Bridge and Waterloo East and setting a career profile linked with development of new systems.

In the 1980s David led a team involved with train describer development and a new system to distribute train location information via the national teleprinter network, followed by the development of the Integrated Electronic Control Centre (IECC) including automatic route setting. He was also involved in improving electromagnetic compatibility between train detection and new rolling stock.

Moving to Railtrack in 1994 he was involved with the development and implementation of TPWS and the development of ERTMS in Europe, his work including the implementation strategy for the UK. Before retirement David managed the safety assurance process for the Cambrian ERTMS pilot. David is one of the contributing editors of IRSE News.

Conclusion

In this article I have attempted to summarise the scope of train warning and protection systems and to highlight some of the major issues the implementation of ATP faces as well as giving an insight into some of the systems that exist.

"Back to Basics" and the IRSE Exam

We hope our 'Back to Basics' series is interesting to all members, bringing topics that underpin our profession into focus in a simply understood, yet comprehensive way, making continuing professional development straightforward.

We particularly hope the articles are interesting to those of you who are new to the industry and are working to build up your knowledge. For those considering taking the IRSE exam, these articles should be particularly relevant for your studies.

As an example, why not think about how you would answer this question from the 2019 Module 5 of the exam, based on what you've learnt from the article?

"Describe a system which provides train protection for a stop signal or limit of movement authority. Using a diagram, clearly describe each of the following:

- (i) *How the speed of a train approaching a stop signal or a limit of movement authority is monitored and controlled;*
- (ii) *How a train is prevented from ignoring a signal displaying a stop aspect or a limit of movement authority; and*
- (iii) *The necessary interfaces to the signalling system including the interlocking and control system."*