Technical Publications

Miniature Train Speed Monitor

Mk2 Technical Report

Peter C. Grossi

Disclaimer

The information contained in this document is for evaluation and guidance purposes only. No responsibility is accepted for loss or damage incurred directly or indirectly as a consequence of using or communicating the information contained herein.

Copyright notice

subject to it being stored, transmitted, printed or reproduced in its entirety, including this copyright notice. No part of it may be copied, transmitted, stored on electronic medium or otherwise reproduced for commercial purposes without the written consent of the author, except according to the provisions of the Copyright, Designs and Patents Act 1988.

Table of Contents

ISSUE RECORD

20/10/22	Initial report.
25/10/22	Addition of first indoor trials.
31/10/22	Increase of sensitivity in the Track Units (revised schematic). Clarification of some explanatory detail relating to Track Units. Site trials report and observations.
6/11/22	Spelling corrections and clarifications

ACKNOWLEDGEMENTS

The author gratefully acknowledges advice and contributions from the following:

Cardiff Model Engineering Society, for the members' advice and access to their rail network at The Heath, Cardiff for design proving.
Miniature Railway Magazine (David Henshaw, Editor-in-Chief) for corrections and advice related to publication of an article based on this report in December 2022.

INTRODUCTION

This is the second development incarnation of an experimental project. The earlier version was successful in proving the suitability of an innovative method of detecting and monitoring trains on a miniature railway network, but some limitations were discovered that complicated the installation. These limitations have been overcome by segregating the analog and digital components.

This report make no assumption about the reader being familiar with the earlier version.

SPECIFICATION

A device was required to enable the speed of trains running on a public-service miniature railway to be reliably measured. It being considered inappropriate, and possibly excessively complicated, to install such a device on all engines operating on the network, while some may be “resident” and others may be occasional or one-off visitors. Compliance with designated speed limits also needs to be reported either by track-side indication or to the signal box.

The device therefore needs to be track or trackside mounted. This requires it to be weather-proof over all reasonable temperatures and all operable weather conditions. It also needs to be robust against other environmental conditions such as leaves, dust and steaming oil, and not easily damaged by small animals or careless boots.

If part or all of the device is to be mounted between the rails it must be of sufficiently low profile as to be entirely below the top of the tracks in case it should foul any part of a train undercarriage. A similar restriction applies if mounted outside the rails unless it is well clear of the entire allowable load gauge width. This is an important consideration for raised tracks where the carriages have running boards hanging some way below track level. Mounting on sleepers may be convenient for stability and simplicity of installation, but this limits the height of any part of the unit to be no more than 25mm or 1” for 5” gauge and 22mm or 0.9” for 3½ “ gauge.

It was not required to report the actual speed, but to report whether a train fails to comply with the required average speed limit over a short designated length of track. It would also be useful to provide an indication that the designated sector is occupied at any time. All indications need to be reported locally by trackside indications and/or through suitable switched signals to a signal box or other devices.

The track may be 3½”, 5”, 7¼” or a dual gauge, and the device must be able to operate with a variety of steam and electrically powered engines.

It is also required to be of low or no maintenance and either permanent or very quick and easy to install as required.

Please note:
While it is not specifically required to do so, it is desirable for the components developed for this application to be easily interconnected in such a way as to allow multiple monitoring points over a whole network. This would allow a greater degree of central monitoring, leading to better overall co-ordination and safety.

BACKGROUND

The problem of measuring train speeds on miniature railways is not new, and has presented a substantial problem across the world where many such installations are to be found.

On model railways, situated in a benign indoor environment, optical sensing is easy and reliable, where speeds can be measured by the time delay between interrupting a pair of light or infra-red beams. These can work very well even in bright sunlight, but in an outside permanent installation this method is very prone to environmental interference. Even in enclosed shelters set back from the track the beam can be obstructed by leaves or a collection of dirt, ice or snow.

Another method tried for outside installations involves isolating a track sector and measuring the conductivity between rails on an electrically isolated sector when connected by metal wheels and axles. This is commonly used for signalling, but it has a variety of problems. Rust and corrosion, as well as oil, wet leaves, snow and ice can affect the reliability of contact, and wet sleepers can provide an alternative conductive path. The isolating gaps needs to be very narrow so that a wheel can pass without undue hammering as they hit the edges, but these gaps may close in hot weather and provide unwanted links between sectors. Also, fishplates are needed to maintain stability across the gap, and these need to be insulating in all weathers and contamination. It cannot be permitted for more than one train (or part of) to be on one sector at the same time. This, while commonly used, is not much liked.

Akin to the above, a load gauge or vibration sensor could be attached to a small track sector. Inductive sensors have been used to detect the slight deformation under weight. This is not an attractive option as it requires a degree of surgery that could destabilise the track itself. And such a sensor may find it difficult to distinguish between an approaching train from one that is present, especially if the isolating gaps become clogged.

Mechanical devices that make contact with the wheel flanges are in use in some installations. These are subject to wear and tear, require cleaning and lubrication, and may not be suitable for the variety of wheels that may be found among the various trains in regular or occasional use. For signalling purposes this has been used, but is not much liked.

Clearly, anything that interferes with the bodies of the engine or carriages, such as a between-tracks sprung lever of some sort would be problematic in coping with a wide variety of shapes and sizes. And many constructor-owners may not take kindly to having their pride and joy assaulted in this way.

Theoretically a Doppler sonar device could be devised that would measure speed directly. At 6mph (8fps), with the speed of sound close to 1000fps and a transmitted frequency of 40kHz the return signal would be changed by 1.6%, or 640Hz. This is detectable for a clean signal, but with a weak return signal and the confusion of reflective structures on a train it seems unlikely to be accurate or reliable for present purposes.

Sonar signal delay measurement could determine the range at any instant. Using the above estimates, the delay between the outgoing and return signals would change by a millisecond every 6ins of movement. Again, this is technically interesting but could be compromised by multiple reflections from the train.

Hand-held radar or laser devices are available for consumer use, and are seen at some installations, but they present problems connecting to installed indicators and signalboxes. And as part of a fixed installation they would suffer from some of the problems mentioned above. There may also be a safety problem with unprotected laser light that could find someone’s eye

Conclusion

From the above the requirement implies a non-contact, non-mechanical device, and this requires some sort of radiation. Having considered the options of light (or infra-red), radar or sonar, the latter has been explored in this project as it appears not to have already been extensively studied, and may be a promising option in terms of potential reliability as well as cost.

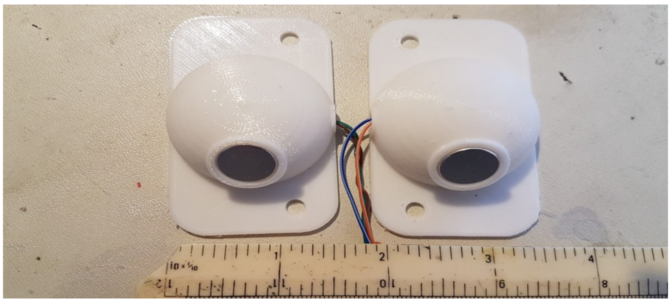
A SOUND APPROACH

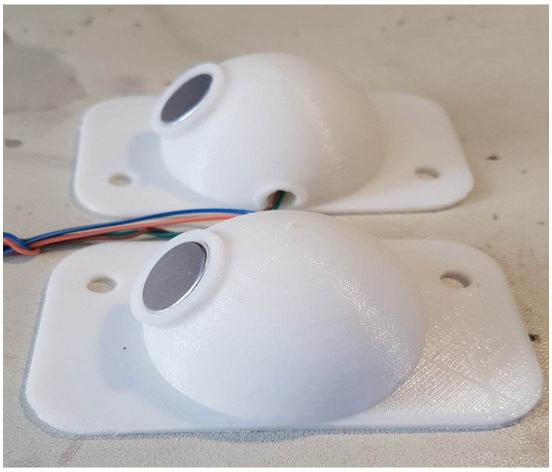
Ultrasound devices are now widely used on motor vehicles to assist reversing, and they are also commonly used in factories for counting items on conveyors. They are not required to measure speeds directly but can provide a reliable indication of the presence of solid objects. They seem to work over a wide range of temperatures, although mounted on motor vehicles they are not subject to quite the same environmental rigours of an outside railway installation. But taking all this into account they seem to offer possibilities.

Ultrasound devices of the types considered are compact and operate with low voltages, which are important factors. Suitable devices, of which there is a variety, may be about 15mm diameter and about 20mm deep including leads, and may operate at 20V. Some have a gauze top face to permit efficient sound wave projection and sensitive reception, but these would not be weather-proof and would require a cover of some sort which would dull their operation. However there are devices totally enclosed in an aluminium cap. These do not require further protection over the face, but they are less sensitive.

Over a period of some months a number of experiments have been conducted using a borrowed 5 inch carriage bogie. A number of enclosures have been designed, printed in plastic and tried to see how ultrasound reflection may offer a reliable means of detection. Also an attempt was made to sense wheels by projecting sound over the track, but as the devices on both sides of the track had to be entirely below track-top level this could not be made to work reliably.

After some trials, it was found that the most reliable arrangement was to project the sound upwards from between the rails. Not directly upwards but at 45 degrees across a track where it would illuminate the wheels, frame, axles, axle boxes and any other clutter that may be attached to the underside of the train. The receiver was set at a similar angle to catch any reflections. The beam pattern of the devices is such that there could be insignificant reflection from the track itself.





All trains have wheels! So whatever other features and differences they may have, this is one feature that, if it could be used, would constitute a universal approach to a solution. And as the devices are tuned to an ultrasound frequency they should not be susceptible to vibration from the trains themselves.

The result was that using small encapsulated devices with approximately 16 volt excitation of the transmitter, a pulsed return of about 1mV was available from the receiver. This is not much, but as long as other interference could be kept low it could be amplified to produce a clear logical response for further processing.

Conclusion

This part of the investigation took the most time as it required extensive experimentation to establish a feasible and reliable detection system, and to develop suitable between-tracks enclosures. But it clearly indicates that an ultrasound system using cheaply available components is a very feasible possibility and should be pursued further.

THE SYSTEM STRUCTURE

From the foregoing a strategy of using reflected ultrasound signals to determine the presence of a train seems to offer possibilities. But as the devices do not measure speed directly (e.g. by Doppler effect) this requires two sensor units (one at each end of the designated sector). It also requires a control unit to process the signal returns, calculate the elapsed time, and provide the necessary indications.

Unlike simple mechanical switches this solution requires several components and processes:

- The ultrasound devices placed at each end of the sector (pairs for transmitter and receiver).
- Environmentally suitable enclosures for the devices and connections.
- Power source for the transmitter excitation signals at the precise resonant frequency.
- Signal processing and amplification.
- Analysis of relative timing for the required speed and distance.
- Status indications and reporting.

These processes are discussed in more detail below.

The devices

The ultrasound devices used in the developed units are explored in the earlier sector. They are commonly available, inexpensive, environmentally protected and compact. They therefore appear to be eminently suited suited for this application. The ones used are mechanically tuned to a precise frequency of 40kHz, which determines the excitation signal.

The enclosures

The problem of the enclosures is discussed in the earlier sector, and it would appear that a suitable design has been discovered. While this presents a problem obtaining compliant commercial enclosures the design is easily and cheaply produced with 3-D printing, for which production files are available.

Signal generation

Whether they employ piezo-electric components or mechanical vibrators ultrasound devices are specified to operate at a very precise resonant frequency. While this could be provided by simple analogue timing devices these depend on passive components that have wide tolerances and would need some sort of adjustment to tune them sufficiently accurately.

This represents two problems: adjustable components may not be reliable in the environment, even in sealed enclosures, and they are inclined to be temperature-sensitive. It is also a good rule that, when designing electronics to operate in a potentially hostile environment, it is best to avoid using adjustable components, switches and moveable links wherever possible.

Therefore it was decided to use crystal-controlled oscillators with frequency division to get the accurate timing. Suitable crystals are inexpensive and specified to be accurate to a few parts per million over the temperature ranges expected. While the selected devices may operate at 20V peak-to-peak, a transistor bridge circuit can be used to double the available output from a more conveniently available supply (e.g. 12V).

Signal processing

In practise the available return from the receivers when tested was not much more than 1mV. This required amplification and conversion to strong logic-level signals for digital processing. Analogue operational amplifiers (Op Amps) were used for the former, being inexpensive and substantially reducing the alternative component and construction complexity. These were followed by transistor switches to secure strong digital signals.

Analysis

The signal analysis was performed digitally using programmable devices. The alternative, using combinational logic, would greatly increase the component count and severely restrict flexibility for future modifications if requirements change. The devices used were from a range produced by MicroChip. These can be configured to use a crystal for precise timing, and have a very flexible way of using the connections for input and output purposes. They use a simple 8-bit program instruction set which is programmed in Microsoft Assembler (MASM), which is a widely used low-level language. These devices can be programmed in C, but having limited processing power, and for a simple operation, a low-level language was found to be more efficient.

Having no operating system, screen or keyboard interface, or network port, they are very small, very cheap and very well suited for embedding in simple “intelligent” modules, as applications such as this require. But not many people have the means to program such devices, or the programming skills, however if an operator requires a different operational characteristic then variants could be pre-programmed on demand or selected as options.

Status and reporting

The signal analysis needs to allow for several operational possibilities. Throughout this report a “sector” refers to a length of track that is marked at each end with a “start” sensor and an “end” sensor; this may encompass platforms, points or other features. This project assumes that the track is one-way, although the device can be programmed to accept trains in either direction:

- A train enters the sector and leaves it at a normal operational speed (perhaps from a slow walk to about 10mph).
- A train enters the sector and leaves it at an average speed above a designated limit (perhaps 5 – 10mph).
- A train enters the sector and stops before it reaches the end.
- A train reaches sector end but stops before it has completely left the sector.

The second possibility represents a potential hazard for the train and it’s passengers, and the last two possibilities represent a potential hazard for any train that may be following. So it seems advisable to provide some sort of indication, not just for excessive average speed, but for sector-in-use and sector-obstructed situations.

The speed limit on any particular track sector may change and it would be inconvenient to modify the electronics or programming to suit (remembering that field-adjustable components are not a preferred option). So if the train speeds are measured by checking the transit time (leading edge to leading edge) the speed limit can be set simply by placing the sensors at an appropriate distance, and which can be easily changed. Since the programmable device is run from a crystal reference, accurate timings can be determined for the train movements for each circumstance, and digital signals generated accordingly.

For a given programmed sector minimum time allowance the sector length can be easily calculated. For example, a 3-second time allowance for a speed limit of 6 mph would require a sector length of 24 feet. The following table shows representative values for different pre-set minimum time allowances and speed limits. The time allowance is programmed into the system according to user requirement.

It should be understood that if the sector sensors are placed on sleepers the following figures may not be realisable precisely. So for safety if the nearest sleeper is not close to the required distance the end position should be brought back to the previous sleeper. The speed would then be marginally overestimated rather than underestimated.

	Sector distance feet		
Speed limit mph	Time required 2 secs	Time required 3 secs	Time required 4 secs
4	10.67	16	21.33
5	13.33	20	26.67
6	16	24	32
8	21.33	32	42.67
10	26.67	40	53.33
12	32	48	64

In determining whether a train has stopped in the sector another timer is used, with a fixed delay allowance in seconds. For example if a train averages less than one tenth of the permitted speed the warning could be issued through an indicator or switched signal. This would remain set until it is automatically cancelled when the end sensor is eventually triggered. This is programmed into the system according to user requirement.

In order to clear the sector-in-use condition the system needs to know when the last of the train has passed the end sensor. As the carriages pass over the sensor it will be intermittently triggered as they pass, but there will be gaps when nothing is sensed, for example between carriages. But if the gaps are timed, and a limited allowance is given for each of them, then a reasonable guess can be made when nothing is sensed for a suitably safe length of time. This also is programmed into the system according to user requirement.

But how would the system know if a train stops after the end sensor is triggered but before it has completely left the sector? If the train stops with some item reflecting the ultrasound then the end sensor can determine that the train is still there and the previous comment applies. But if the train stops with a gap between carriages over the sensor the situation is not clear. In the end, the only reliable way of determining whether the track is clear is by looking at it, but a third sensor further down the track at a distance at least the length of the longest train could be used. This is broadly how mainline systems traditionally work, with three detection points controlling each of the sector signals. This is not a requirement in this specification but could be considered for further development.

Future development

If, in service, this system proves to be reliable within it's own limitations, a more fully integrated system could be determined, which can overcome this limitation and which can be co-ordinated with trackside signals. The units could also be connected with a digital link to the signal box (which could be a laptop computer), where the above timings and other parameters could be monitored and changed at will during operation.

CONSTRUCTION Mk2

This is a development of the earlier Mk1 and operates on exactly the same principles. Changes were required in the analog design to enable easier installation and to desensitise it against radiated interference.

The device was split into three parts. Each pair of ultrasound devices is serviced by a close-in analog unit, the Track Unit. These can be mounted very closely to the ultrasound devices between the tracks and enable very short leads between the devices and the amplifier/detector circuits. These can easily be kept separate to prevent crosstalk. The two Track Units generate the device excitation locally to reduce crosstalk issues with the interconnecting cables.

The third part is the digital unit, or TrackSide Unit. This sends a 40kHz reference signal to the Track Units, together with a 10V DC supply. It receives the detected signals in digital form so there is no longer a problem with large excitation voltages running alongside millivolt sensor data.

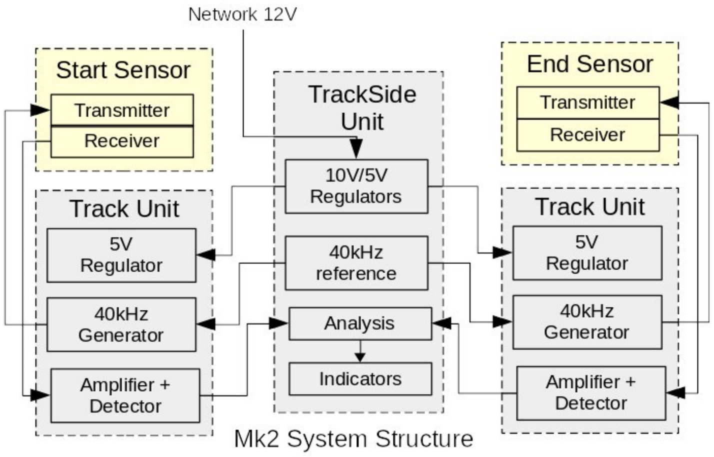
The TrackSide Unit may be placed at any convenient location near the track. But as it may be several tens of metres or more from either or both of the Track Units, measures were needed to ensure robust communication. In earlier projects the author experimented with using standard CAT5 cable to communicate between digital devices over long distances. For this purpose balanced line drivers were successfully used, of type ST485. These can be used as receivers or transmitters driving into twisted pairs. The two output signals copy the input signal and it’s inverse, effectively providing a balanced 10V peak-to-peak signal from a 5V supply. It was found that it could reliably send and receive mixed signals down the four pairs of a standard CAT5 cable over a total distance of at least 1km.

To complement the local LED indicators MOSFET devices are provided as earthing switches, so that other trackside devices or relays could be driven remotely without needing to be concerned with their operating voltages (DC only).

Structure

The TrackSide Unit requires a single 12V supply from the network, but it does not use the network supply directly. This supply is expected to be electrically noisy due to other connected devices which may include relays and motors. So the supply is first regulated to 10V with high frequency suppression, and is then regulated down to 5V for the programmable analyser.

The total current consumption from the 12V supply is 5mA plus about 8mA for each of the local indicator LEDs when lit (of which four be lit at any time).



Mk2 System Structure

The 10V supply is sent to the Track Units where it is used directly to generate the 40kHz excitation independently for each of the transmitters. It is regulated in each of the Track Units to provide 5V locally for the amplifiers and detectors. The excitation signal is less critical of noise, and benefits more from the higher power output, but the active components need to be well protected so they are hidden behind two regulator stages.

Each sensor comprises an ultrasound transmitter and a receiver. With the component types used here (SN36969) both are the same and can perform either function. Each transmitter is driven from the 40kHz generator in the Track Unit, and each receiver is connected independently to the amplifier in the Track Unit. The ultrasound devices are housed separately because it has been found that if they are housed in the same unit the sound short-cuts through the structure and compromises the reflected signal clarity. The transmitter/receiver pairs are placed between the tracks and looking upwards at an angle towards the wheels and frame. In a dual gauge layout the common rail is used.

Not shown in the structure diagram are the balanced-line signal driver/receivers for the 40kHz reference and detected digital signals. These enable the three units to be placed at large distances from each other. Also not shown are the connections from the TrackSide Unit and the site infrastructure for signalling as these depend entirely on the particular site requirements.

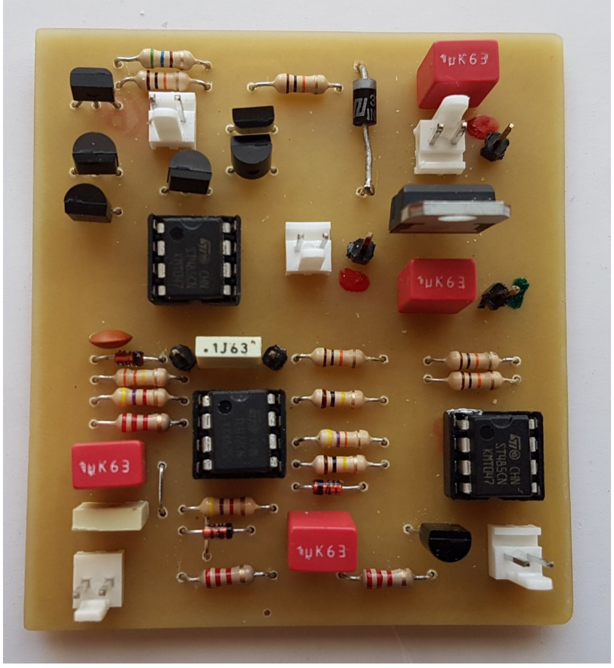
System applications

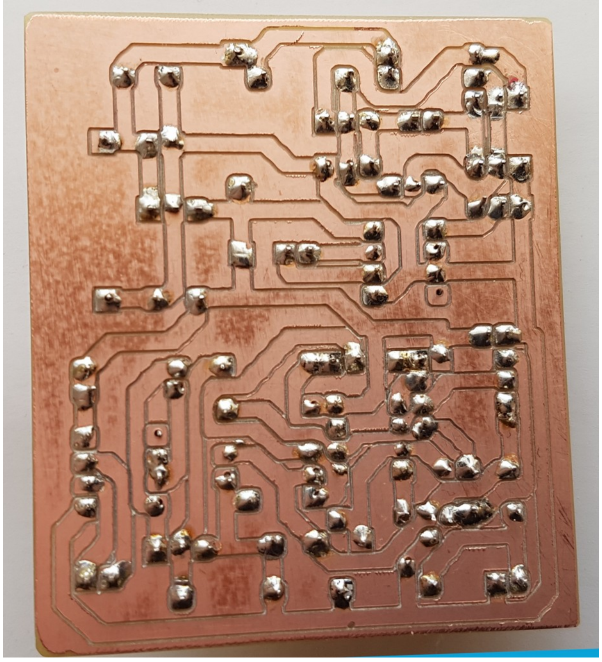
The Track Units have a very simple electrical interface, and perform a simple function which can be applied to many different purposes. Therefore someone who is confident with electronic design could use any number of these devices in whatever configuration they require in conjunction with their own design of TrackSide Unit (or virtual signalbox) to collect, analyse and report track data across their network.

The following describes the Track Units in detail as well as the TrackSide Unit that was developed with reference to a particular installation, where it was only required to report excess train speeds and track occupation for a sector within a network.

Track Units

Electronics





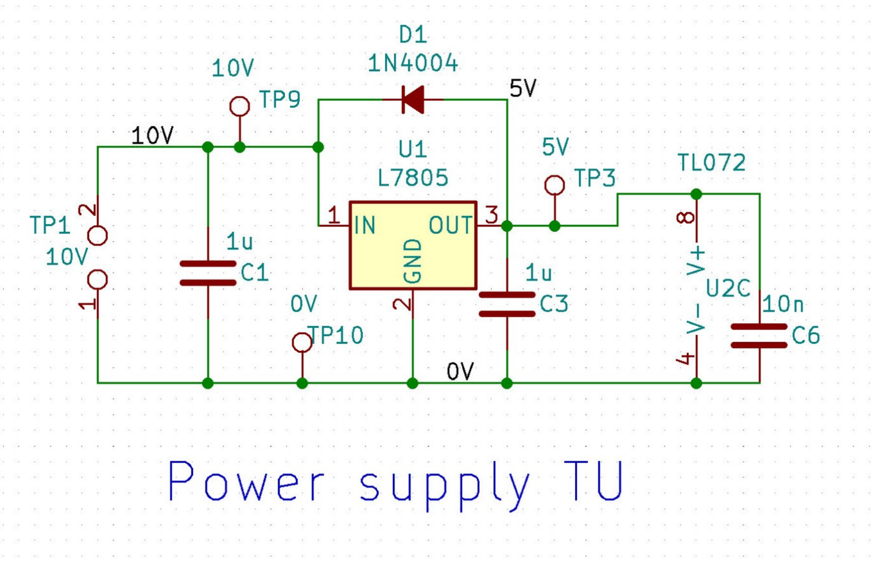
The two Track Units are identical so only one is shown. They are each made in the same way as the TrackSide Unit.

The top right corner of the component picture shows the power regulators, and the top left corner shows the H-bridge components feeding the 40kHz to the transmitter. The small IC nearby is the line receiver for the reference 40kHz.

Bottom left is the amplifier IC and associated network, with the incoming analog signal in the corner. At bottom right is the detector transistor and the line driver for the detected signal to the TrackSide Unit.

The circuits

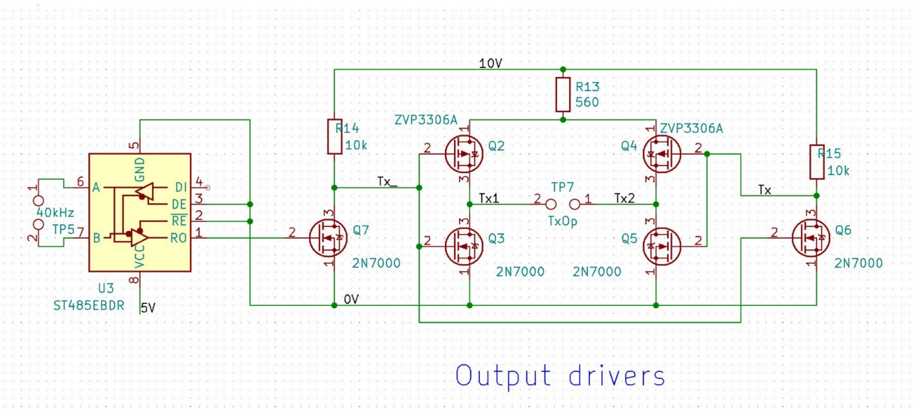
The circuit schematic is broken into three sections. A local power conditioner, a power driver for exciting the ultrasound transmitter, and an amplifier/detector to process the input from the ultrasound receiver



The Track Unit requires a 10V supply from the network. This is obtained from the TrackSide Unit, so only one network supply is required overall.

While this may be noisy it is considered suitable to be used directly for the device excitation, but for the sensitive analog components and digital interfaces it is necessary to reduce it to 5V.

The input is is suppressed by a capacitor to sink high frequency noise before being regulated. The regulator device is a fixed-voltage variety as no variation or adjustment should be considered necessary. The diode provides protection of the regulator against a sudden collapse of the 10V supply which could damage it by applying a reverse voltage from the internal power line. The output is damped by C6 which reduces propagation of any noise generated internally.



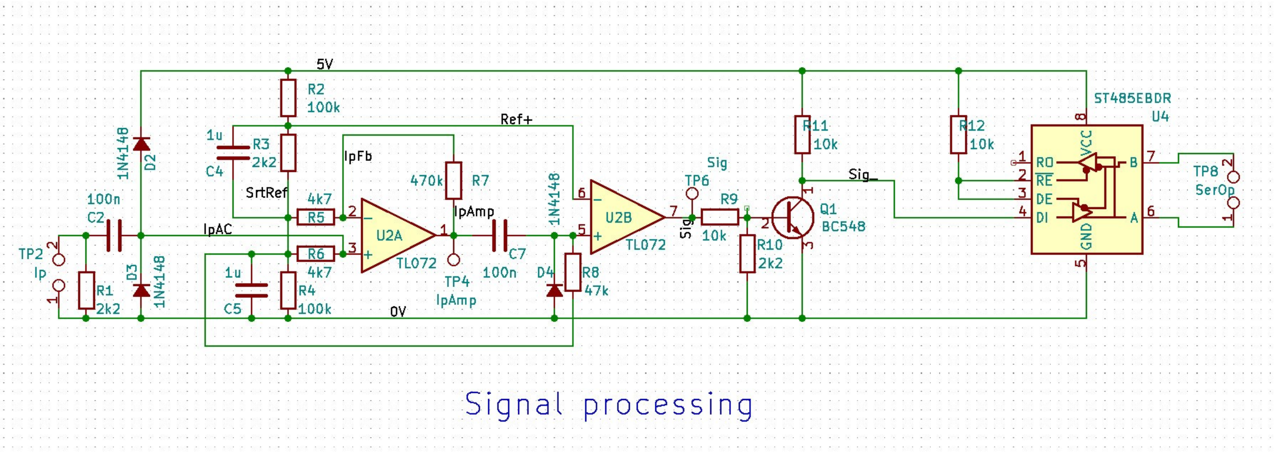
To get the greatest possible signal return from reflections of the transmitted ultrasound, the greatest possible voltage is required for the transmitter. To this end the outputs are driven from the 10V supply. They are not driven directly from the 12V network supply as it may be noisy and contain large transients. Also, a bridge circuit is used to give a bipolar output, so the transmitters are given nearly 20V peak-to-peak.

R13 is used to protect the components and the power regulators in case of a short-circuit in the installation wiring. Each of the outputs is delivered by a twisted pair cable to reduce susceptibility to electrical interference. This acts as a balanced line, which is intended to be entirely isolated and not grounded at any point.

The reference 40kHz input is provided by a balanced digital line from the TrackSide Unit and received by Q3. But the device does not provide complementary outputs so transistor Q6 provides an inverse control signal to one side of the bridge. This avoids the necessity of providing a second input from the TrackSide Unit.

The input signal is received directly from the ultrasound receiver and is of the order of 1mV to be detected.

The signal processing consists of three stages. The first stage is protected by an input isolating capacitor and a resistor R1 to provide a consistent load to the receiver. It is also protected by a pair of diodes D2 and D3 to ensure that the input signal does not venture more than about half a volt outside the power lines, in case there should be a large transient picked up by the cable.



The first two stages use the two operational amplifiers (Op Amps) provided in a single 8-pin package. These devices have an extremely high gain and are intended to be used with a feedback circuit to provide whatever amount of amplification is required. The way to understand these devices is to realise that the output on pin 1 is kept at whatever voltage is required to maintain zero volts between the inputs on pins 2 and 3. Here, pins 2 and 3 are both connected to the same point in the potential divider R2, R3, R4, and which is kept at a steady DC level by C5 at around half the line voltage, so the device is kept in the middle of it’s operating range. So if an input raises the voltage on pin 3, the op amp raises pin 1, which is fed back through R7 until pin 2 has the same voltage.

The amplification is therefore governed by the ratio of R7 to R5, which keeps it at around 100. So the pulse output into C7 is at about 100mV for an input of just over 1mV. C7 then isolates the DC level into the second op amp.

The inputs to U2B are biassed by the voltage across R3 of the potential divider mentioned earlier, which is negative by about 55mV . This is maintained at a steady DC level by C4.

As the negative input is held slightly higher than the positive input this causes the output on pin 7 to be normally low. But on receipt of a pulse of more than 55mV the output switches to a high level. Diode D4 ensures the pulse from C7 does not drive the input below the negative rail if a large transient should be received, and R8 in conjunction with C7 provides a time constant of 4.7ms which is long enough to ensure the 40kHz pulses are not suppressed.

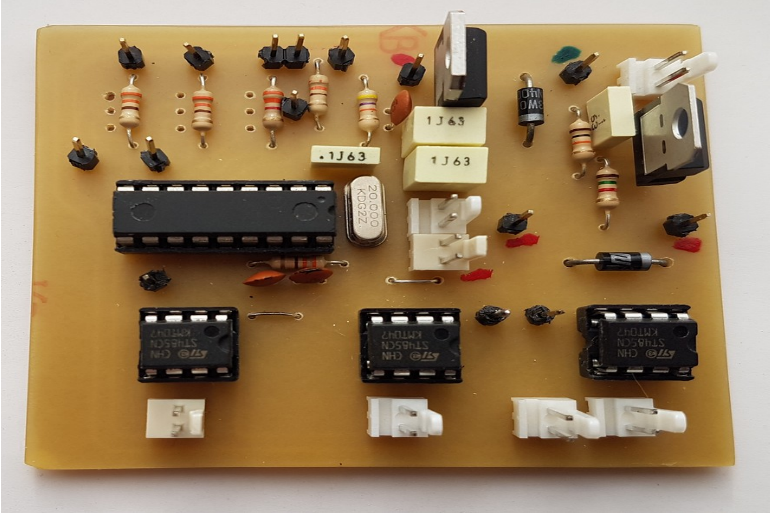
Transistor Q1 provides a switch to ensure a strong logical output. R9 and R10 prevent the base being over-driven while ensuring adequate sensitivity. The output is provided to the TrackSide Unit through the balanced line driver U4.

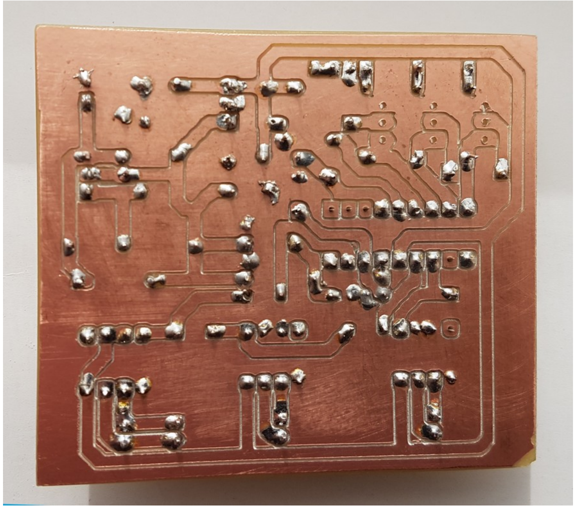
The output should be connected through a twisted pair cable, such as used in Ethernet installations, and with the specified line driver type it should be reliable over distances of at least 1.5 km. The 40kHz reference input should be provided by similar means, and the 10V power can be carried in the same way. Therefore this device can be reliably connected using three pairs within a standard Ethernet cable (e.g. CAT5), which is easily and comparatively cheaply available.

When analysing the digital output from the Track Units it must be appreciated that when a reflective surface passes the ultrasound devices a continuous stream of pulses will be received at 40kHz until the reflection stops. This can be useful, as the validity of the detection can be determined by disregarding isolated pulses that may be received spuriously from interference. But this has implications if the device is to be used for counting axles or bogies.

TrackSide Unit

Electronics





The TrackSide Unit was made from single-sided copper laminate on a resin substrate, and engraved to delineate the tracks. The equipment and materials required for a photo-chemical process as used in economic mass production were not available, but the design files were prepared in the commonly used Gerber format, and could be provided for alternative production methods if required.

The active components were mounted in sockets for ease of development, and to enable the programmable component to be easily removed for firmware modifications. In-circuit programming is possible with these devices but the required facilities were not available.

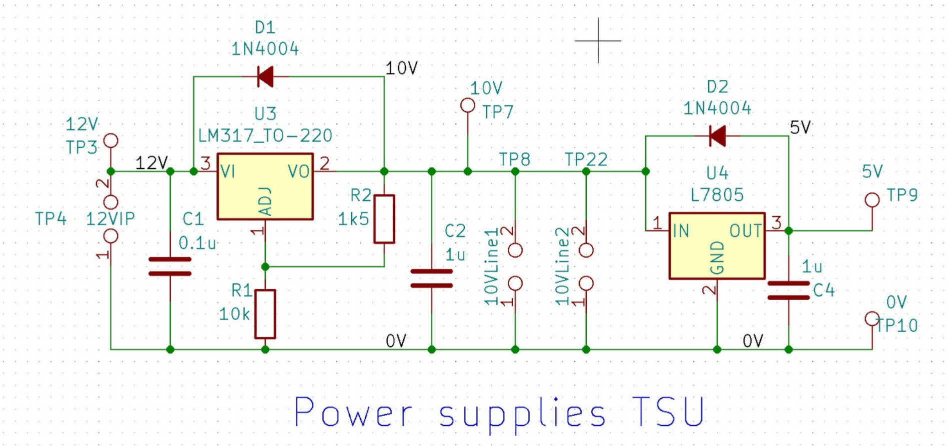
The top right corner of the component picture shows the power regulators, and the top left corner shows the connections to the four LEDs (mounted on the enclosure): green for operating; red for train-in-sector (which blinks if the train stops); an additional red for stain stopped; and yellow for excessive speed. Provision is there to fit MOSFETs for external switching.

The large (18-pin) component is the programmable device to generate a stable 40kHz and to analyse the inputs. The pin headers next to it provide the 10V regulated output for the Track Units. Along the bottom are the line driver ICs for the digital links to the Track Units. Bottom left is the connection for the start detector Track Unit; bottom centre is for the end detector Track Unit, and bottom right are connectors for the 40kHz reference signals for each of the Track Units.

The circuits

The circuit schematic is broken into sections.

The TrackSide Unit requires a 12V or greater supply from the network (up to 36V), which is suppressed by a capacitor to sink high frequency noise.

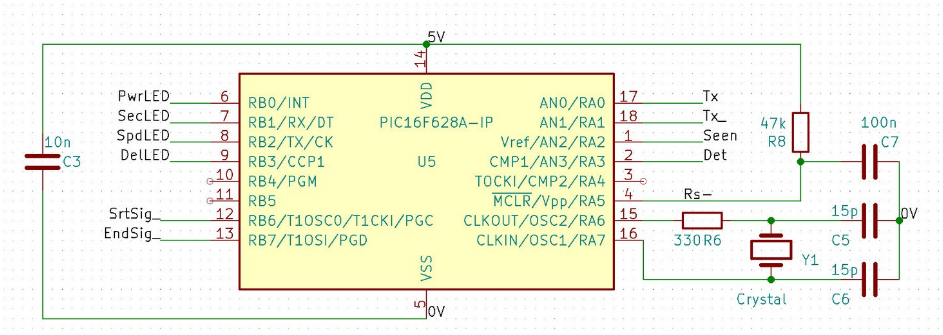


The first regulator maintains a constant voltage between the output and the Adj pin, so it produces an output governed by the ratio of R1 and R2. 5% tolerance resistors are used here, as elsewhere, because the actual voltage is not critical. D1 provides protection of the regulator against a sudden collapse of the 12V supply which could damage it by applying a reverse voltage from the internal power line. The output is damped by C2 which reduces propagation of any noise generated internally.

The 10V supply is provided for the Track Units and is not used within the TrackSide Unit except to derive the 5V line required by the digital components.

A versatile regulator is used to provide 10V for the analog components instead of a fixed-voltage type because if the supply is prone to drooping a lower regulated level may be necessary, and this can be adjusted by simply changing one resistor.

The second regulator is of a type that is intended only to produce the required 5V supply for the digital components. The output is damped by C4, but the processor has a small additional capacitor to suppress high frequency internally generated noise. Again, a diode is used to protect the regulator against sudden supply failure.



Processor, Clock and Slow Start

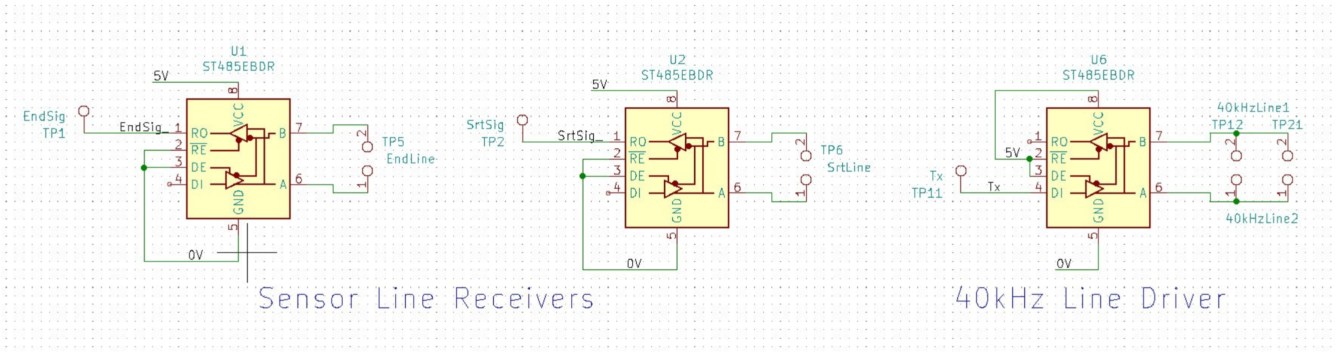
The processor is an 18-pin device from MicroChip. It incorporates an 8-bit RISC (reduced instruction set computer), and has sufficient other connections and working memory to cater for all the required needs (with a few spares for further development). It is configured to use a 20MHz crystal X1 connected to pins 15 and 16. This provides an instruction clock which runs the processor at 5MIPs.

The processor also provides up to three internal timers of which only one is used. This provides a stable timing reference for the 40kHz excitation and all the required speed and track event timing calculations. It is produced by a binary division that actually results in 39kHz. This is close to the resonant frequency of the ultrasound devices, although it does incur a few dB of losses. For simplicity this is referred to as the 40kHz signal throughout this report.

The 40kHz signal is output to pin 17 and it’s inverse on pin 18. These are provided as reference signals for the Track Units to generate the excitation power to the transmitters.

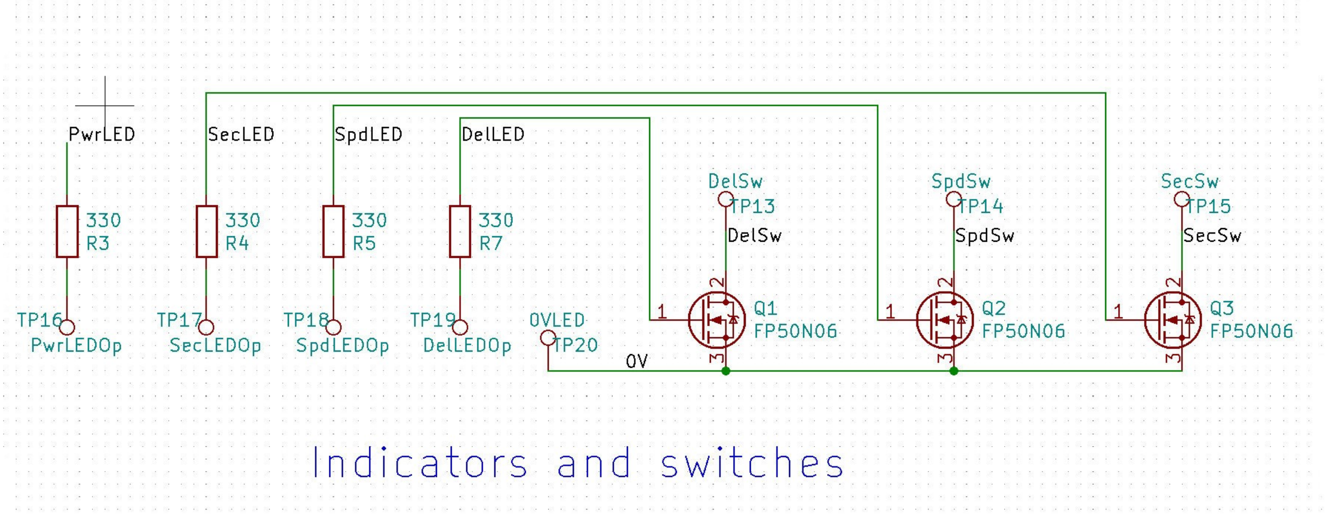
Pins 12 and 13 are configured as logical inputs from the two Track Units shown in the structure explained above. These inputs are read by the program to determine what is going on with the trains. When a part of the train (e.g. a wheel) passes the sensors there is a short period when the ultrasound is reflected. This normally results in a burst of pulses which input to these pins. The program counts these pulses to provide confidence that the signal is genuine and not a transient from some other electrical equipment.

Pins 6, 7, 8 and 9 provide outputs for LED indicators. The green LED on pin 6 is turned on by the processor after performing setup checks. The red LED on pin 7 indicates that a train has been detected entering the sector and has not yet left. If the train takes too long this LED is flashed to draw attention to the track being blocked. The yellow LED on pin 8 is turned on if the train reaches the end sensor too quickly, and remains on until the train leaves the sector. The fourth output on pin 9 is designated for additional indication that a train has stopped in sector, but it can be programmed for other purposes if required.



Two line receivers, U1 and U2, are used to accept detector signals from the Track Units for detections at start and end of sector, and a single line driver is used to provide 40kHz reference to both the Track Units. The device is rated with a large fanout for shared balanced lines.

The line transmitter/receiver devices are controlled by signals on pins 2 and 3. In this application those pins are permanently connected as each device is only ever used in one mode. A logical high configures it as a transmitter; a low as a receiver. Pin 1 is used for an output to the line, and pin 4 is used to receive an input from the line, so in this application only one of them is used for each device.



The local indicator LEDs are driven through resistors R3, 4, 5 and 7, and an adjacent 0V header provides a return path. The LEDs themselves are mounted in the enclosure rather than on the PCB itself.

The external devices are switched by MOSFETs Q1,2 and 3, and no return path is provided as they are intended to be driven from an external DC supply which need not be within the range used by this circuit, but will need to have a common 0V or return connection. For high-voltage or AC devices an external relay is necessary for isolation.

SOFTWARE (Microcode)

The software can be considered in three parts: the startup routine prepares all the working variables and configuration of the processor, it then starts the interrupt process and hands over to the baseload loop which continually monitors events and keeps the device alive.

The software was developed in Assembler (MASM) using the IDE (Integrated Development Environment) provided with the programming device from MicroChip (a PICkit v3 was used). While other source code languages such as C can be used the limiting power of the processor benefits from a more direct implementation.

Startup

On power-up the program allocates memory addresses for the required selection of program variables which are later used to record what is going on and how long things are taking. There are also user-specified fixed values required later as references for the allowed intervals. While these cannot (in this implementation) be changed dynamically they are declared in a way that enables them to be easily changed. These specify in half-second units:

- The minimum time permitted for a train to traverse the sector. This is set at 3 seconds.
- The maximum reasonable time for a train to traverse the sector. This is set at 22.5 seconds.
- The “dead” time used to determine whether a train has completely left the sector. This is set at 15 seconds.

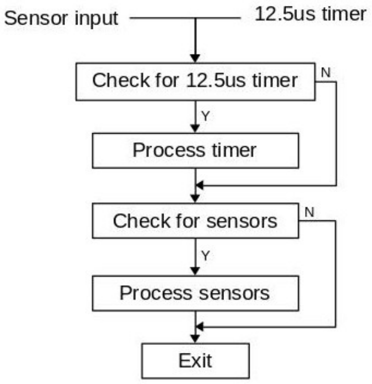
It then goes through an intricate hardware setup sequence which configures the various pins as inputs or outputs. It sets up a master timer based on the 5MHz instruction clock to provide program interrupts at 12.8 microsecond intervals (78kHz), and configures interrupts to occur whenever either of the sensor inputs changes. When all this is finished the green “operational” LED is lit.

Interrupt routine

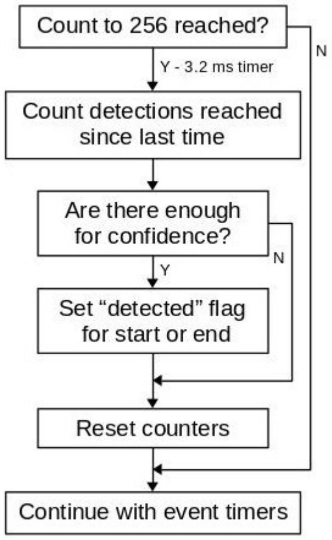
The interrupt program is run under either of two conditions. It is called at regular 12.8 microsecond intervals by the program interrupt timer, and it is also called if there is a change in either of the sensor inputs.

In either case a flag is set in memory that is checked by the program to determine what is happening. So if the timer flag is set the timer routine is run, and if either flag is set the sensor inputs are checked and they are processed if required.

The flags remain set until cleared by the program, and while set they inhibit further interrupts. So if the program is held up it cannot be itself interrupted, which would cause confusion. Therefore the flags are cleared by program when the routine exits.



Mk1 Interrupt Routine



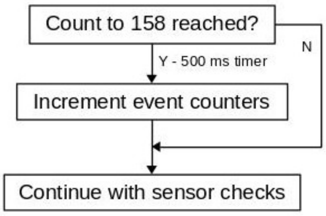
Mk1 Interrupt Timer Process 1

The timer process is in two parts. The first part further divides the interrupt clock into 3.2ms intervals and counts how many detections have been received for each of the two sensors in that time. When there is a genuine detection at the sort of speeds involved there should be a short burst of detections at 26us intervals, so there should be at least a few counted in the time used. This count is compared with a confidence factor to eliminate spurious signals from electrical interference.

If confidence is achieved for a sensor then a flag is set for it. These flags are checked continually in the baseload loop and appropriate action is taken as described below.

The second part of the timer process divides the clock into periods of half a second and counts half seconds in several counters that are used to determine train speeds and delays.

One counts half-seconds since the first detection was received by the start sensor after the sector had been clear; this is checked in the baseload loop to determine how long the sector has been occupied. It is also used to determine whether the train seems to have stopped if it reaches a designated value before the end sensor is triggered, or reaches the end sensor too early and is speeding.

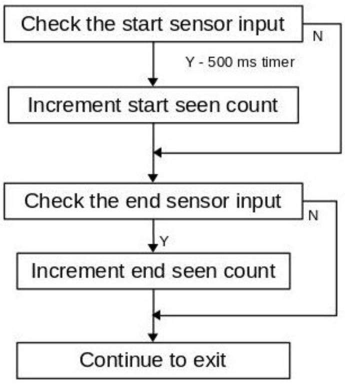


Mk1 Interrupt Timer Process 2

Another counts half-seconds since the first detection was received by the end sensor after the sector became occupied. And another counts half-seconds since the most recent detection by the end sensor; this is used to estimate whether the train has fully cleared the sector.

The final process in the interrupt routine is to check the sensor inputs and count how many times they detect a train. These checks are run every 12.8 microseconds and whenever an input triggers an interrupt.

The numbers collected here are analysed and restarted every 3.2ms by part 1 of the interrupt timer process to determine whether there are enough “hits” at either entry or exit to be confident of a genuine detection.



Mk1 Interrupt Process Sensors

Baseload

The baseload runs a continuous loop checking for changes reported by the interrupt program and to monitor the counters in case anything is happening too quickly or too slowly.

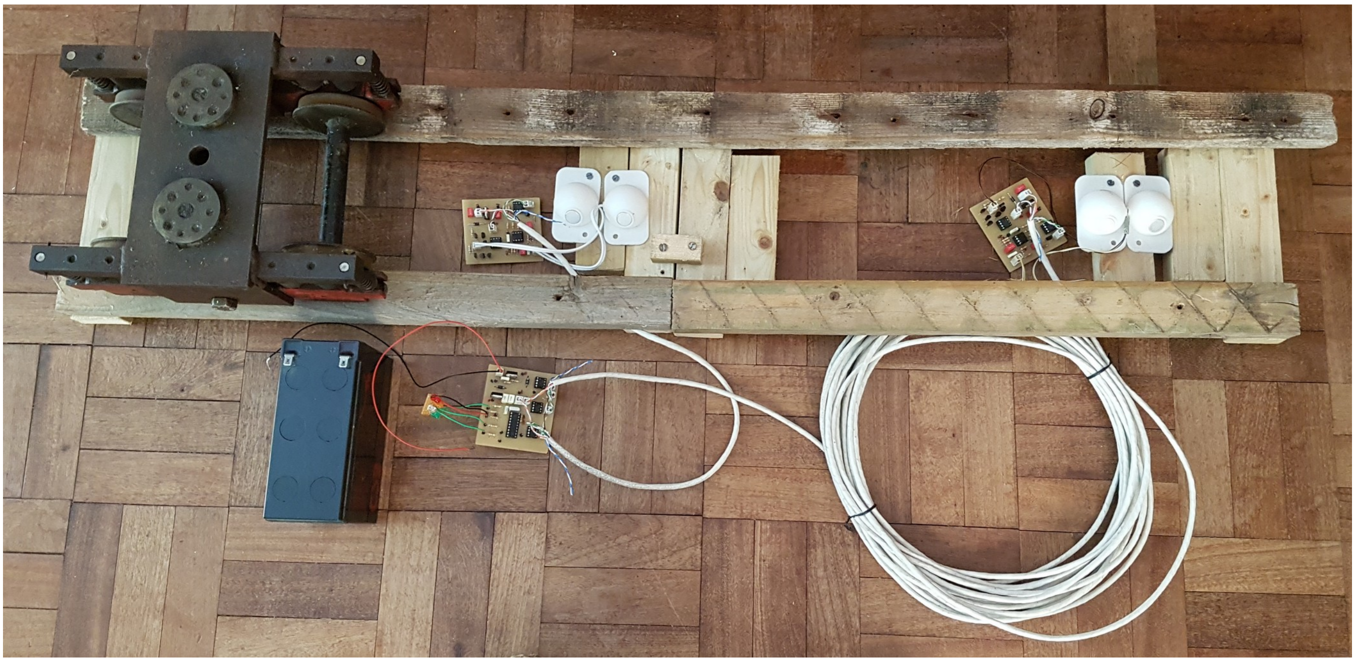
The baseload program is structured as a continuous loop, continually performing checks according to operational status:

Sector clear. The most recent train has cleared the sector. This is determined by the train having been seen exiting by the end sensor and there has been a delay since any part of the train has been seen after that. When the sector is clear the Sector Occupied light and the Speeding light (if lit) are turned off.
Train entered. A train has been detected by the start sensor but has not yet been seen by the end sensor. The Sector Occupied light is lit and the Sector Occupied counter is reset at the first detection so the time in sector can later be calculated.
Train exiting. A train has bas been seen to enter the sector and has been seen also by the end sensor. The Sector Occupied light remains lit because there will be some length of train remaining in sector. On first detection the time in sector is calculated with reference to the Sector Occupied counter. This is compared with the allowed minimum time.
Train exceeding speed limit. When the train is first seen exiting, the time in sector is checked against an allowed minimum time and the Speeding light is lit if it is too soon. This remains lit until the train has completely exited the sector.
Train delayed. If the time in sector reaches a designated value without the train being seen to be exiting it is deemed to be delayed for some reason. The Sector Occupied light flashes as a warning and remains so until the train is eventually seen to be exiting. This light then remains continuously lit in the normal way until the train has entirely left the sector.
Train exited. This is another name for Sector Clear, and takes us back to the beginning. It is deemed to occur when a train has been seen exiting and no detections have been seen by the end sensor for a designated period of time. ***It may be falsely determined if a train stops for a number of seconds over the end sensor such that the sensor does not detect anything (e.g. between carriages)***.

INITIAL (INDOOR) TRIALS

Setup

In order to verify that the system works according to functional specification a test track was set up using scrap wood, and a borrowed 5” carriage bogie.



The ultrasound sensors can be seen as the two pairs of white domes between the tracks. The Track Units are on short wires connected to the ultrasound devices, and the TrackSide Unit is next to the track, on the end of longer wires. To verify the communication reliability a coil of standard quad twisted pair cable of length 50 feet is shown connecting to one of the Track Units.

The so-called TrackSide Unit, which can be connected using much longer CAT5-type cables, need not actually be near the side of the track, and could be placed in a signalbox. In an expanded network of sensors this would be more convenient and allow a more creative evaluation of the distributed sensor signals.

The electronic units have not been enclosed, pending site trials to determine practicable designs.

The whole is run from a small lead-acid type 12V battery, shown next to the bogie.

Results

The system worked reliably and was not found to be unduly sensitive in the placement of the devices in relation to the track. However it was found that the distance of the devices from the track had an effect on where in the bogie it was first detected. This is not significant for the essential purpose of detecting a train for the purpose of determining it’s speed, but may have implications regarding the use of this arrangement for counting axles or bogies. This is an issue that will need to be considered for future developments.

INITIAL (SITE) TRIALS

Setup

The system was trialled on a track of the Heath Park Miniature Railway in Cardiff.

In order to ensure a consistent position of the ultrasound devices a mounting plate was prepared that is placed over the devices and interlocks with their bases. This securely locates them to ensure consistent performance.

The Track Units were placed between sleepers on a dual-gauge track (5” and 7¼”) at a distance of 20 feet. This, with the time allowance of 3 seconds programmed into the TrackSide Unit, corresponded with a speed of 5 mph, as required by the network. The same cable used for the indoor trial (length 40 feet) was used, and kept away from the track.

The photograph shows one of the sensor units mounted on a temporary sleeper, and the electronics were placed in waterproof boxes positioned next to the track, with the short signal wire pairs well separated. The weather-proofing was well tested through the day because it rained continuously for the entire time!

The TrackSide Unit was placed in a temporary box with the LED indicators exposed (not photographed, but other than improvised weather protection was the same one as was tested indoors). No changes or adjustments were made to any of the units or cables.

The site could not conveniently provide their own 24V supply, so the test was conducted using a small 12V lead-acid battery.

A diesel locomotive with several carriages was used to test the unit.

Results

The devices consistently detected the train and indicated correctly when the allowed speed was exceeded. It also accurately responded to the end of the train as detections ceased.

As a simple test of vulnerability, a large and very wet leaf was draped over one pair of devices, but they continued to work accurately.

Overall the exercise was a success, although it was not convenient to test the system with a large variety of different rolling stock. It was found that the Track Units met the functional requirements and could easily be installed on existing track. However some adjustment to sleepers would be needed to fit the units between 3½” tracks.

It was felt that the system could be easily integrated with an existing relay-based signalling system. The provision for connecting the Track Units over distances of several hundred metres allows the TrackSide Unit to be placed in the signalbox, which is very desirable for installations that have a signalbox.

The system demonstrated a capability for use as-is for monitoring train speeds (subject to availability of a 12+V supply and a suitable trackside indicator), and showed potential for being developed into a more extensive monitoring system, either independently or in conjunction with an established relay-based system.

In conversation with the site operators a number of suggestions were made regarding future developments:

Axle-counting is a very desirable further development, so the departure of a train from a sector can be more reliably determined, and this should be investigated promptly.

If an axle-counter were sensitive to the direction of movement it would cater for shunting operations, which can take place in both directions, even on a one-way track. It was therefore decided to consider how two closely-positioned sets of devices could be used to count axles or bogies in either direction, and whether the Track Units in their present form could achieve this purpose. Further trials are to take place when this has been investigated.

Subject to the TrackUnits being suitable for axle or bogie counting and bi-directional operation with acceptable reliability the designs could form the basis for substantial operational improvements on similar sites.

It was considered important that the TrackSide Unit design is programmable. This allows the functionality to be diversified according to site requirements without significant electronic changes. If several inter-related track sectors were to be equipped, then the TrackSide Unit would need to become something more like a standard gauge Evaluator, in electronic form.

Some thought should be given to developing a visual design aid so that track operators could manage a network of sensors and their interlocks without having to reprogram the TrackSide Unit, in whatever form it becomes. Such a system would require something a lot more sophisticated than a simple programmable chip. A possibility would be to replace the TrackSide Unit with a Raspberry Pi. This would need to be connected through a panel of line drivers/receivers of the type used here, but it could then be controlled with a suitable program and a graphic interface.

No other system was known to offer the same potential for versatility, low maintenance and reliability in regard to operational and environmental conditions.