Locomotive monitoring system using wireless sensor networks

P. L. Croucamp, S. Rimer
Department of Electrical and Electronic Engineering
University of Johannesburg
Johannesburg, South Africa
Email: suvendic@uj.ac.za

C. Kruger
Advanced Sensor Networks, Meraka
CSIR
Pretoria, South Africa
Email: ckruger1@csir.co.za

Abstract—Theft of cables used for powering a locomotive not only stops the train from functioning but also paralyzes the signalling and monitoring system. This means that information on certain locomotive’s cannot be passed onto other locomotives which may use the same track at different times hence increasing the possibility of accidents which can be costly. We propose to design and implement a standalone wireless sensor system that can be used to monitor a locomotive and which runs on a separate platform than that used to actually operate the locomotive.

Keywords—railway signalling; wireless sensor network; Arduino; train collision detection

I. INTRODUCTION

Cable theft has become a problem not only in South Africa but globally as well. Recently, there have been many accidents caused by cable theft which paralyzes the locomotive monitoring and signalling systems [1]. These accidents do not only cause serious injuries to passengers but can claim lives as well.

On 31 January 2013, a passenger train accident which occurred near Pretoria injured over 200 people and although no casualties were reported, there were serious injuries [2]. The expenses left by the accident amounted to approximately R42 million [3]. Half of this amount was needed to compensate the families of the injured passengers and the remaining half was to be used to repair the infrastructural and technological damage caused by the accident.

The rail industry however is solely dependent on these cables to operate the locomotives which require large currents to operate, currents as large as the order of 1500 A or even more. However, for the locomotive signalling and monitoring systems, the use of cables is not necessary meaning that this system can therefore be replaced.

When cables used for powering the locomotive are stolen, this not only stops the train from functioning but also paralyzes the signalling and monitoring system. This therefore means that information on this locomotive cannot be passed onto other locomotives which may use the same track at different times hence amounting to accidents that can be costly.

By solving this problem lives can be saved and millions of Rands can also be saved and used in other profitable projects. The public can also have peace of mind knowing that their lives are in safe hands due to reliable monitoring and signalling systems that are highly efficient. What is required is a standalone system that will be used to monitor the locomotive and which cannot be paralyzed by cable theft. The system has to be safe in the form that it cannot be “stolen”. It must not be subject to any criminal activities and has to withstand the harsh conditions that can prevail in the rail industry.

The objective of this paper is to describe the design and implementation of a Wireless Sensor Network (WSN) that can be used to monitor locomotives on a track. The system runs on a separate platform than that used to actually operate the locomotive.

Generally, we can define a Wireless Sensor Network (WSN) as a distributed group of sensors which can monitor physical and environmental conditions and then pass information to a central location. In most cases, for control purposes the networks that are used are bi-directional. A WSN application can consist of numerous sensors that monitor quantities such as temperature, vibration, acceleration, direction and many others. These networks can be used for applications including military, industrial control processes, health monitoring and as in this project train/locomotive monitoring and many more [4], [5].

II. BRIEF SUMMARY OF SIGNALLING SYSTEMS

We start our theoretical background with railway signalling in general and then the railway signalling system employed in the South African Rail Industry. We define railway signalling as a system that is used to control trains and prevent collisions. Trains do not stop quickly and do not stop within sighting distance of the driver. The most common form of rail traffic control is to run trains using a fixed timetable as to which train uses a track and at what time. This reduces mobility and is a slow system generally. It is essential though since train tracks are not as flexible as air or a tar surface used by airplanes or cars respectively. Therefore trains may only use a section of the track at a given time during which another train is not permitted to use that very same section [6]-[8].

However, trains can run on opposite directions and meeting points are therefore inevitable. At such points, one train must stop in order to allow the other train to move. In most cases, these track sections are based on a certain timetable as to which train uses that section first or reaches the meeting point first.
The timetable has several disadvantages and these are given below:

- There is no confirmation that the track ahead is clear.
- The system does not allow for electrical or train engine failure though the system is setup so that there is sufficient time for the failed train crew members to setup warning signs.
- The system is not flexible. Addition or rescheduling cannot take place without notification thus wasting time.
- A fourth problem is a corollary of the third: the system is inefficient. To provide flexibility, the timetable must give trains a broad allocation of time to allow for delays, so the line is in the possession of each train for longer than is otherwise necessary.

These problems can impact on the economy in a sense that minerals or any other goods being transported may arrive later than scheduled or passengers using trains may not arrive for work etc.

A. Train Detection

The most common way to determine which track section is occupied is by using a track circuit. The rails at either end of each section are electrically isolated from the next section, and an electrical current is fed to both running rails at one end. A relay at the other end is connected to both rails.

When the section is unoccupied, the relay coil completes an electrical circuit, and is energized. However, when a train enters the section, it short-circuits the current in the rails, and the relay is de-energized. In modern train systems, color light signals are used instead of flags. The color convention is similar to that of the car traffic system.

1. Green : Proceed at line speed and expect to find next signal displaying green or yellow.
2. Yellow : Prepare to find next signal displaying red.

Generally, there are two signalling conventions used globally. These include route signalling and speed signalling.

Route Signalling: With this signalling method, the driver is informed which route to take. The driver uses his route knowledge together with speed restrictions. This system requires the driver to know the different routes and in most cases the driver is only permitted to drive on a track section of a route he is familiar with. This therefore restricts emergency routes as the driver may not know the route.

Speed Signalling: Here, the driver is not informed which route the train will take, but the signal aspect informs him at what speed he may proceed. Speed signalling requires a far greater range of signal aspects than route signalling. However, it is less dependent on the driver’s route knowledge.

Historically the signalling system used at Transnet (South Africa’s railway operator) had two purposes, namely:

- Facilitating the safe movement of trains by preventing derailments and collisions (primary purpose).
- Increasing the capacity of the railway network by safely allowing a higher density of trains on the network.

These functions were carried out by the operating personnel with the use of hand signals before a new signalling method was implemented. All of the train detection and interlocking functions were vested in people. The initial development of technology-based signalling systems was therefore based on the processes which were in place just before the start of the implementation of signalling. The primary purpose was then redefined as:

**To facilitate the safe movement of trains by ensuring that they comply with their issued authorities of movement (both speed and extent of authority)**

The signalling system and the functions that it had to fulfill are given in Table 1. Also given in the table are the ways in which these functions were implemented using conventional signalling and the proposed ways of implementation in the new train authorization systems.

<table>
<thead>
<tr>
<th>TABLE I. REQUIREMENTS AND CHARACTERISTICS OF SIGNALLING SYSTEM [8]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Required Functions</strong></td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td><strong>Conventional Signalling</strong></td>
</tr>
<tr>
<td>Position of Train</td>
</tr>
<tr>
<td>Train Complete</td>
</tr>
<tr>
<td>Required Route</td>
</tr>
<tr>
<td>Section Defined and Clear</td>
</tr>
<tr>
<td>Formulated Authority</td>
</tr>
<tr>
<td>Communication of Authority</td>
</tr>
<tr>
<td>Ensure Compliance with Authority</td>
</tr>
<tr>
<td>Centralized Train Authorization Management centre</td>
</tr>
</tbody>
</table>

We can summarize by saying that the main difference between the conventional signalling system and the new “transmission based” or “communication based” train authorization system proposed above is that the communication of authorities and the updating of train positions is conveyed via a radio network and not via ground based equipment and cables.
A second significant difference is that the determination of the position of a train and its completeness is done on board the train in an acceptably safe manner.

This is not exactly the same principle that is being used in this paper’s design but similarly, knowing the position, speed and the acceleration of the train as well as other quantities such as momentum, it is possible to locate the train and also to make sure that there are no collisions on certain track sections.

III. SYSTEM DESIGN

As we will be using analog sensors, most sensors, and specifically the sensors that we will be using for this project give out a certain voltage. For example, with the proximity sensor, the output voltage varies with the distance measured, so if the voltage can be measure so can the distance.

However, most PC’s are digital in nature. They therefore only differentiate between high and low. Therefore to counter this problem, most computers have an Analog to Digital Converter (ADC). It will thus convert a voltage to a number so that it can be processed. To do this, a microcontroller has to be used. Arduino is a single-board microcontroller designed to make the process of using electronics in multidisciplinary projects more accessible [9]. The hardware consists of a simple open source hardware board designed around an 8-bit Atmel AVR microcontroller, though a new model has been designed around a 32-bit Atmel ARM. The software consists of a standard programming language compiler and a boot loader that executes on the microcontroller. The Arduino integrated development environment (IDE) is a cross-platform application written in Java, and is derived from the IDE for the Processing programming language and the Wiring projects [9].

A. Mechanical Analysis (Dynamics)

Since we are dealing with locomotives that can carry hundreds of human beings, or tons of minerals, it is necessary to define the different physical elements that can be incorporated in the project such as the relationship between distance, speed, velocity, mass, time and forces in general such as frictional forces, gravitational forces, drag forces and numerous others.

The dynamics of the system can be described using equations of motion which generally describe the behavior of a physical system with respect to time. They allow for the conservation of momentum which is the product of mass and velocity.

A simpler way of understanding momentum physically is the fact that a system which consists of a heavy mass takes a longer time to speed up and therefore takes a longer time to slow down as well. As this project is a locomotive monitoring system, it is necessary to take momentum into account as we will be dealing with how much time and effort the locomotive has to use in order for it to start or stop. We will now define all the variables that have to be considered and also derive equations that will be used in the project.

We can define the velocity as:

\[ v = \frac{r_f - r_i}{t} \]  

We now have acceleration as:

\[ a = \frac{v_f - v_i}{t} \]  

We then have:

\[ v_f = at + v_i \]  

\[ r_f = r_i + v_i t + \frac{at^2}{2} \]  

\[ r_f = r_i + \left(\frac{v_f + v_i}{2}\right)t \]  

\[ v_f^2 = v_i^2 + 2a(r_f - r_i) \]  

\[ r_f = r_i + v_i t - \frac{at^2}{2} \]  

Given these equations we can easily solve for all the parameters. However, since we are dealing with dynamics and not kinematics, the forces and the energies of the physical systems must be taken into account. Like velocity, linear momentum is a vector and possesses a direction as well as a magnitude.

\[ p = mv \]  

Linear momentum is a conserved quantity, meaning that if a closed system isn’t affected by any external forces then the total linear momentum cannot be changed. This therefore means that the total momentum is constant.

This is implied by Newton’s laws of motion. We therefore now derive the law of conservation of momentum using Newton’s second and third laws of motion.

\[ F_1 = \frac{dp_1}{dt} \text{ and } F_2 = \frac{dp_2}{dt} \]  

We then have:

\[ \frac{dp_1}{dt} = -\frac{dp_2}{dt} \text{ or } \frac{d}{dt} (p_1 + p_2) = 0 \]  

If the different velocities before and after interaction are defined as follows,

\[ v_{f2} : \text{Final velocity for quantity 2.} \]  

\[ v_{f1} : \text{Final velocity for quantity 1.} \]  

\[ v_{i2} : \text{Initial velocity for quantity 2.} \]  

\[ v_{i1} : \text{Initial velocity for quantity 1.} \]  

\[ m_2 : \text{Mass for quantity 2.} \]  

\[ m_1 : \text{Mass for quantity 1.} \]  

Finally we have:

\[ m_1 v_{i1} + m_2 v_{i2} = m_1 v_{f1} + m_2 v_{f2} \]  

This law will hold regardless of how complicated the force between the particles is.
B. Sensing

Initial Sensing: For simplicity, we will assume that the tracks being used will only accommodate for one direction only. In other words, it is not possible for the track to be used bi-directionally at any time (i.e. the track is only unidirectional). Therefore, we do not necessarily need any device to detect the direction at which the train is travelling in. We are only interested in the distance from the traffic lights. Therefore, we only need a sensor that will let us know how far the train is from the traffic lights as shown in Fig. 1.

Signalling: Once the distance to the traffic lights has been calculated or obtained, the train with the shortest distance to the traffic lights can now be signaled to proceed with a green light or flashing lights. The train driver that will use the track next will know that his/her train is the next in line to go. For this paper, we only considered the case where only two trains are to be monitored as shown in Fig. 2.

Final Sensing: The final part is to determine whether or not the track is clear. We need a sensor which will report that the track is clear or in use as shown in Fig. 3. Once the track is clear, then the next train can be signaled to go.

C. Signalling

Signalling can be achieved through LED’s. The general convention will be that if train 1 has a smaller distance between it and the traffic lights (where the distance sensor will be mounted), then the light indicated will be green. If train 2 has the smaller distance then the light indicated will be green. There will be 4 LED’s with two different colors, green and red as shown in Fig. 4.

![Fig. 4. : Signalling of the system](image)

If distance 1 is less than distance 2, then LED 2 and LED 4 should be switched on. Then if distance 2 is less than distance 1, then LED 1 and LED 3 should be switched on.

IV. Detailed Design

A distance sensor can sometimes be referred to as a proximity sensor. It emits an electromagnetic field or a beam of electromagnetic radiation and then looks for changes in the field or return signal. A voltage is then induced and this voltage can be used to interpret the distance. In general, the larger the voltage, the shorter the distance from the sensor to the object being sensed. The proximity sensor will be used to find the distance between the trains and the traffic lights (LEDs in this case). The proximity sensor must be mounted on the same platform as the LEDs to ensure that there is accuracy. Once the two proximity sensors measure the distance independently, they will use transceiver 1 and 2 to send data to transceiver 3. This means that node 1 (with transceiver 1) and node 2 (with transceiver 2) will be programmed to transmit to node 3 (with transceiver 3). At this point in time, node 3 will compare the data being sent to it. This will be accomplished through a microcontroller.

Depending on which train has the shortest distance, the LEDs will be switched on accordingly (LED 1 and 3 if train 2 is closest to the LEDs and LED 2 and 4 if train 1 is closest to the LEDs). Once the train has been signaled to move, it will use the track section where the vibration sensor will be placed.

When the train approaches, the vibration sensor will detect vibration and keep the system as is. When there is no longer vibration, then it will send a signal to node 3 using transceiver 4 which will be part of node 4 to allow the LEDs to switch oppositely i.e. if LED 1 and 3 were on, then they will have to be switched off and LED 2 and 4 will have to be switched on as shown in Fig. 5.

Vibration sensors measure the magnitude of vibration. For this project, vibration sensors will determine when and where the vibration occurs for a position and referencing system. Their method of operation is that there is a voltage that is induced when the sensor is shaken.
A vibration sensor was chosen to determine whether the track was clear or not. If the track was clear, then there wouldn’t be any vibration but if the track was in use then there had to be vibration. The vibration sensor gave out a voltage which was induced by any vibration.

However, since trains are large bodies, they can cause vibration even when they are far. Generally, the specific vibration sensor used for this project gives out an increasing output voltage due to an increase in vibration. This means that in the real world a different threshold voltage (higher than in the project) has to be determined based on the vibration caused by a real train.

The vibration sensor microcontroller was programmed using the Arduino programmer. Initially, with no vibration, the voltage given out is 0V, when there is vibration the voltage given out is not 0V, meaning that the train is using the track. Once the train is now not using that specific part of the track, then the voltage is then 0V again and that is the state that is needed to determine that the track is clear. This information is then passed on to the signalling node.

V. EXPERIMENTAL RESULTS

We will now analyze the proximity sensor and its behavior. When objects were moved in front of the proximity sensor, the distance changed and so did the voltage. When the distance decreased, the voltage that was induced increased. As the XBee module transmitted data to the laptop, this data showed both voltage and the distance. These values were recorded while varying the distance to and from the proximity sensor. The relationship is shown in Fig. 6 and Fig. 7.

Furthermore, the microcontroller was programmed so that as the objects moved closer and closer, the LED flashed faster and faster. The vibration sensor was connected to the Arduino board. The microcontroller was programmed such that when there was vibration, there was a voltage output and the LED would flash and when there was no vibration then there would be no voltage output and the LED would stay off.

A. Analysis of Results

The basic requirement was fulfilled and the system was capable of monitoring traffic control of trains. The proximity sensor was able to calculate the distance from the target to the traffic lights. Its function was to light an LED if the minimum distance was reached. The vibration sensor’s function was to light on the LED when there was vibration. In terms of the different graphs, it is clearly visible that when the distance is smaller, the voltage is larger. However, the relationship is not perfectly linear.

Interpolation therefore had to be done in order to ensure that the readings were consistent. A table is used to interpolate the actual distance by finding the nearest entry in the table and adjusting it based on the ratio of the measured value to the next table entry. The readings obtained only contained a 1% error.

Proximity sensors such as the one used in this project are used to measure the distance generally provide an analog output that can be measured using analogRead. They can have greater accuracy to their ultrasonic counterparts. Generally, the vibration sensor produces a voltage in response to physical stress. The more it is stressed, the higher the voltage.

VI. CONCLUSION

A system using wireless sensor nodes to monitor train distance and occupation of railway tracks was prototyped. The system indicates that the use of wireless sensors can be a viable system to use for railway signalling systems. The advantages of using wireless sensor nodes is that they are not dependent on copper cable (copper cable is prone to theft), and they are flexible in their deployment and can be placed across multiples areas of a railway track.
REFERENCES


