Transportation Research Part C 89 (2018) 234–253

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

An autonomous system for maintenance scheduling data-rich complex infrastructure: Fusing the railways' condition, planning and cost☆

Isidro Durazo-Cardenas*, Andrew Starr, Christopher J. Turner[1], Ashutosh Tiwari[2], Leigh Kirkwood, Maurizio Bevilacqua, Antonios Tsourdos, Essam Shehab, Paul Baguley, Yuchun Xu[3], Christos Emmanouilidis

Cranfield University, Cranfield, Bedfordshire MK430AL, United Kingdom

ARTICLE INFO

Keywords:
Data-driven asset management of rail infrastructure
Systems integration
Condition-based maintenance
Intelligent maintenance
Data fusion
Cost engineering
Planning and scheduling
Systems design and implementation

ABSTRACT

National railways are typically large and complex systems. Their network infrastructure usually includes extended track sections, bridges, stations and other supporting assets. In recent years, railways have also become a data-rich environment.

Railway infrastructure assets have a very long life, but inherently degrade. Interventions are necessary but they can cause lateness, damage and hazards. Every day, thousands of discrete maintenance jobs are scheduled according to time and urgency. Service disruption has a direct economic impact. Planning for maintenance can be complex, expensive and uncertain.

Autonomous scheduling of maintenance jobs is essential. The design strategy of a novel integrated system for automatic job scheduling is presented; from concept formulation to the examination of the data to information transitional level interface, and at the decision making level. The underlying architecture configures high-level fusion of technical and business drivers; scheduling optimized intervention plans that factor-in cost impact and added value.

A proof of concept demonstrator was developed to validate the system principle and to test algorithm functionality. It employs a dashboard for visualization of the system response and to present key information. Real track incident and inspection datasets were analyzed to raise degradation alarms that initiate the automatic scheduling of maintenance tasks. Optimum scheduling was realized through data analytics and job sequencing heuristic and genetic algorithms, taking into account specific cost & value inputs from comprehensive task cost modelling. Formal face validation was conducted with railway infrastructure specialists and stakeholders. The demonstrator structure was found fit for purpose with logical component relationships, offering further scope for research and commercial exploitation.

☆ This article belongs to the Virtual Special Issue on "Big Data Railway".

* Corresponding author.
E-mail address: i.s.durazocardenas@cranfield.ac.uk (I. Durazo-Cardenas).
[1] Present address: The University of Surrey, Rik Medlik Building, Guildford, Surrey GU2 7XH, United Kingdom.
[2] Present address: The University of Sheffield, Portobello Street, Sheffield S1 3JD, United Kingdom.
[3] Present address: Aston University, Birmingham B4 7ET, United Kingdom.

https://doi.org/10.1016/j.trc.2018.02.010
Received 14 August 2017; Received in revised form 5 February 2018; Accepted 15 February 2018
0968-090X/ Crown Copyright © 2018 Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/BY/4.0/).

1. Introduction

1.1. The British railway maintenance system

Railway systems are key enablers of economic prosperity. The British railway is considered the fastest growing and safest in Europe (Network Rail-1, 2013). It has been estimated that over 1.6 billion annual passenger and 282,000 freight journeys took place during 2014–2015 (Anon., 2015; Nair, 2015). The British network infrastructure comprises over 20,000 miles of track, 30,000 bridges, 2500 stations, some of which are almost 200 years old, and a multitude of geographically dispersed signaling, electrification and crossing systems. It is the very large number of assets, the long distance interactions and complex interdependencies, as well as its day to day management that clearly make the British railway compliant with the widely accepted classification of a complex transportation system (Magee and de Weck, 2004).

Railway network infrastructure has a very long life, but inherently degrades. Interventions are necessary but they are costly, and can cause lateness, damage and hazards. For example, over 40, 000 infrastructure incidents causing delays were reported in the British railway in 2011–2012; with the top contributors being signaling malfunction, rail crossings, broken rails and detection failures (Network Rail-2, 2013). Effective asset management is crucial to maintain the network reliability and ultimately the users' safety. Network Rail PLC manages and maintains the British railway infrastructure. As many as 17,000 people are engaged in the maintenance activities (Sheeran et al., 2015).

The British railway operates a complex maintenance system by virtue of its own nature; one where multiple failure modes can occur in large numbers, from a range of geographically dispersed assets. Network Rail's comprehensive asset management program is supported by state-of- the- art track inspection trains which are run periodically across the country. These specialist "New measurement trains" (NMT) are equipped with a range of sensors including scanning lasers, high-resolution cameras, ultrasonic systems and linear variable differential transformer probes (LVDT), amongst others. Inspection periodicity varies in accordance with the track speed and importance. During inspection, a number of railway track profile parameters are evaluated against their corresponding thresholds; which are based on Network Rail's standard NR/L2/TRK/001/mod11 (Network Rail-3, 2015). When any of the thresholds is exceeded, a maintenance corrective action notice is issued. Standard prescribed intervention timescales are specified depending on fault severity, and these broadly range from track closure to 28 days to enact repair. Additional measures might include, for example, speed restrictions.

The gradual adoption of remote condition monitoring systems (Network Rail, 2014) for an increasing number of the assets means that more data than ever is being acquired. At present large amounts of data are captured, analyzed and fed into several of the internal systems and databases transferred from the inspection trains to the data centers. Data Collection Services analyze and process the data, performing quality control, post processing and conversion to standardized formats. To support decision making, a linear asset decision support system (LADS) has been recently introduced (ORR, 2014). LADS overlays several geo-tagged information streams, over a graphical representation of the railway. This enables visual identification of the network's "hot spots". However, reasoned supervisory decision making input (e.g. diagnosis and prediction of failure), as well as scheduling of the required tasks for incipient faults, is still required.

1.2. Towards more complex, data-rich automatic railway maintenance systems

Train services are expected to substantially increase over the next 30 years (TSLG, 2012). As a result, railway systems are undergoing a profound modernization. As railway systems progress, the complexity in scheduling maintenance jobs efficiently will increase; and inevitably, a more refined and autonomous decision making capability will be demanded. Fig. 1 illustrates transitional concepts and high-level system characteristics of the UK railway, from a conventional complex system to a complex maintenance system and to a complex data-rich automatic maintenance system. This illustration was constructed from elements in the literature and communications with the stakeholders. The elements in the definition of a complex system (Magee and de Weck, 2004) applied to railway infrastructure primarily refer to the network's complex operation and size, the number of assets and personnel, the long distance interactions and interdependencies. Underneath its operation, a complex stakeholder hierarchy exists, including public funders and regulatory bodies, rolling stock companies (train owners), train operating companies, service providers and the general public (ORR, 2016).

Despite some recent technological developments, railways still rely on resource intensive and less economical time-based preventive maintenance practice to ensure the network availability. A further level of complexity applies when its assets deteriorate and unplanned interventions become necessary. Preventive interventions are carefully planned by qualified personnel to minimize train service disruption. Network disruptions are cost penalized at tabulated rates. Unplanned corrective actions generate a "domino" effect that affects many of the closely related network interactions (Wang, 2008). Their economic and public image impact costs are critical. Clearly railways are becoming a data-rich environment with a substantial number of data streams and operating data-bases. In addition to track inspection data-sets, delay incidents, flooding, CCTV, cost, timetables and users social media data are generated daily (Durazo-Cardenas et al., 2016; Rahman et al., 2015). It is clear that a new structured automated approach is necessary to more efficiently transform current and new generated data to maintenance decisions, to more comprehensively support safer and continuous operation of the network.

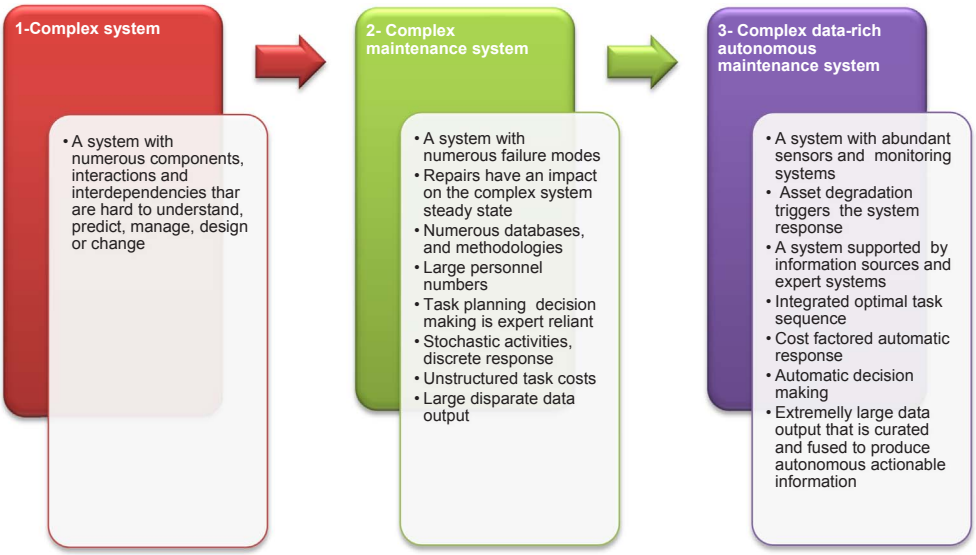


Fig. 1. Railways transition towards a complex data-rich automatic maintenance system, from a complex maintenance system.

1.3. Integration through systems engineering and data-fusion principles

Integrated systems development generally relies on systems engineering methods for specifying processes, control functions, interfaces and associated databases. For example, the widely used waterfall approach provides a structured sequence of design, implementation and test phases; including formal reviews and delivery of documentation (Waltz and Hall, 2009). The initial phase starts with the definition of the system followed by subsystem, preliminary and detailed design phases. Using this approach, high-level system requirements are defined and partitioned into a hierarchy of increasingly smaller subsystems & components. The advantages include:

- ability to build large systems by decomposing them into smaller, manageable and testable units;
- ability to work with multiple builders, designer, vendors, users and sponsors;
- ability to define and manage risks by identifying the source of potential problems;
- formal control and monitoring of the system development process.

Data fusion establishes links between data and information sources, and closes the loop from the minutiae of data collection to strategic decision making. Formal development of the concepts and guidelines for data fusion model architectures were initially developed by the US Department of Defense (DoD). Data-fusion is closely associated with systems engineering principles (Liggins et al., 2009; Steinberg et al., 1999). For example, the widely used Joint Directors of Laboratories (JDL) data fusion model (Hall and Llinas, 1997; Liggins et al., 2009) specifies a hierarchal process comprising the following levels:

i data pre-processing,
ii object refinement,
iii situation refinement,
iv impact assessment
v process refinement.

Other subsequent approaches to the fusion process have been developed; and comprehensive reviews have been presented by a number of authors (Esteban et al., 2005; Khaleghi et al., 2013; Sinha et al., 2008). Data fusion principles have been adopted by a wide range of science and engineering disciplines, including a number of Condition-based maintenance (CBM) systems that employ a combination of data fusion and data-mining (Niu et al., 2010; Raheja et al., 2006). While data-fusion systems development considers particular application attributes and requirements, commonly observed stages include (Hall and McMullen, 2004):

i requirement analysis;
ii sensor selection;
iii architecture selection;
iv algorithm selection;
v software implementation;
vi testing and evaluation;

Additionally, international standard BS ISO 13374 (BS ISO 13374, 2007) provides guidelines for architectural development of integrated condition monitoring systems.

1.4. Related research and limitations

1.4.1. Data driven fault diagnosis in railways

Railways infrastructure faults and inspection methods are described by Popovic et al. (2013) and Jovanović et al. (2014). Early railways infrastructure modelling for condition monitoring based on inspection data is presented by Jovanovic (2004). Recent research publications present big-data challenges and prospects for asset management for railways infrastructure (Thaduri et al., 2015) and Nunez and Attoh-Okine (2014). Nunez et al. (2014) also demonstrates the use of visual analytics to support decision making using frequency analysis. Morant et al. (2016) used failure analysis of previous incidents to support maintenance decision making of signaling systems. Li et al. (2014) analyzed several of the railways data rich databases, exploring several analytical and machine learning methods to build vehicle failure prediction models to optimize the network usage. All of these publications demonstrate the potential to effectively diagnose faults using big-data in modern railways.

1.4.2. Planning and scheduling algorithms

Lidén (2015) comprehensively reviewed the challenges of planning infrastructure maintenance jobs; with emphasis on the strategic and operational issues. In terms of algorithmic approaches to optimize planning and scheduling, Bouillaut et al. (2012) provide an approach and a decision support tool for the reliability maintenance of underground rail tracks, analyzing interventions impact. The approach taken uses a Bayesian network for the modelling of maintenance strategies to detect and prevent broken rails. Lidén and Joborn (2017) sought to optimally integrate traffic-free windows with the network maintenance plans using a mixed integer model. Su et al. (2017) proposed a multi-level model for optimizing maintenance interventions; considering degradation modelling, intervention scheduling time and operations analysis. Santos et al. (2015) proposed the use of decision rules model (DRM) for impact mitigation and costs saving during railways heavy duty maintenance operations. Guler (2013) described a decision support system for railway track maintenance and renewal programs, which is comprised of rules developed from interviews with track experts and secondary research sources. Nyström and Söderholm (2010) present an expert knowledge method for the prioritization of railway maintenance actions. Zhao et al. (2009) examined scheduling activities in the form of synchronized rail track component renewal. In this work they utilize a genetic algorithm approach to optimize track renewal activities so minimizing the cost incurred and the track possession time. Similarly, Zhang et al. (2013) also used a genetic algorithm to plan maintenance searching for a minimized cost solution. Guler also (Guler, 2017) proposed the use of genetic algorithms to support decision making in preventive maintenance in the Turkish railway infrastructure. Rail track monitoring data and fundamental cost estimations were factored-in.

1.4.3. Cost models for railways

Cost estimation of railway maintenance has been attempted through various approaches. Patra et al. (2009) give a series of models for estimating the lifecycle cost (LCC) some of the maintenance activities required by the Swedish railway. The equations achieve a LCC by applying the net-present value (NPV) equation to the sum of individual maintenance intervention costs for each year. These equations are therefore presented in such a way that it is possible to use the equation to calculate the cost of a single intervention. The work of García Márquez et al. (2008) investigated the cost/benefit of installing condition monitoring equipment to switch and crossings. These critical infrastructure components are frequently a source of delays. The presented work included more detail on the British railway denial of service costs, which are potentially significant and are included within the cost module developed for this paper. A more complex approach to cost/effort prediction is the Petri net, which has been demonstrated for some specific railway assets; such as bridges (Le and Andrews, 2013) or track tamping (Andrews, 2012). Petri nets need high quality asset datasets and often access to expert opinion. Developing a Petri net and validating it is a significant task, making this approach currently too challenging to implement as part of an autonomous maintenance system. While the work of Patra et al. (2009) and García Márquez et al. (2008) guided our approach to cost estimation, both are assuming that the task of cost estimation will be done by a human. For the system developed within this work the cost estimation needs to be performed by an algorithm within an autonomous system.

1.4.4. Intelligent systems integration approaches

Increasingly the rail industry is looking to autonomous and intelligent systems to address infrastructure decision making maintenance needs. Data fusion principles are widely used for the integration of many decision making systems. In addition to the widely known sensor state-estimation part of the fusion process, subsequent revisions of the JDL data-fusion guidelines proposed the co-existence of a resource-management component (Steinberg et al., 1999). The latter has been generally associated with the planning and assignment of tasks to available resources. However, to date most of the data fusion research has generally focused in algorithm application and refinement for sensor data association, object refinement, estimation and classification; with very few developments of the resource planning fusion component (Scholz and Gossink, 2012).

In the closely related field of prognostic health management (PHM), a limited number of application-specific publications have presented attempts to couple replacement part logistics to the estimation of remaining useful life (RUL) (Hess et al., 2005; Jianhui et al., 2003). As with data-fusion, much of PHM research has focused primarily on diagnosis and RUL estimation algorithms; subsequent actions have been commonly regarded as business management functions (Sikorska et al., 2011). In the present work however, task scheduling, cost engineering and added value are essential inputs for automatic actionable maintenance.

Intelligent railway systems have been also presented. For example Bombardier's ORBITA (Provost, 2010) combines train sensor data for rolling-stock fleet condition based maintenance, but doesn't provide fully autonomous decision making.

Progressive efforts have been conducted across a number of national railway systems. The fundamental differences of the previous research with our work include using a top-down systems engineering approach for autonomy, a more in-depth cost analysis including denial of service cost, and more detailed maintenance task and crew intervention management.

In this paper, we present the design of a high-level architecture for a complex data-rich maintenance system based on data-fusion systems engineering principles. This novel system effects maintenance decisions automatically, from fused technical and business drivers, i.e. faults diagnosis, optimum task sequencing and cost effectiveness. We demonstrate the design principle using real track datasets to simulate the systems response, and to validate fault alarms, scheduling algorithms and cost models.

To our knowledge, a structured approach integrating asset degradation, planning and cost, to automatically allocate a large number of maintenance jobs in a complex data-rich system, has not been presented before.

2. Design of an automatic system for complex infrastructure maintenance scheduling

The integration of an automatic system for railways infrastructure maintenance scheduling clearly necessitates systems engineering/data fusion principles. The vast size and nature of the national railway infrastructure and its complex operation requires the analysis of any system to be broken into manageable fragments. In line with these principles, the system requirements, component inputs/outputs and subsystems interaction were isolated in order to design the integrated architecture and to analyze and implement the necessary algorithms.

2.1. Requirement analysis

In order to enable efficient data to information to decision transitions, the early stage of formulating the fusion architecture i.e. knowing what to ask and where, is of vital importance (Esteban et al., 2005). This process begins by understanding of the system overall aims; for example: what decisions are sought? What constitutes a successful system? Understanding of the decision environment and the anticipated inferences also play a part (Waltz and Hall, 2009).

To successfully capture the new system requirements, effective engagement and communication with the stakeholders was fundamental. This was initially generated through a group-discussion with a number of railway and other industry senior specialists sponsoring this research program. In addition to the specific application requirements, structured discussion dynamics responded also to questions and topics such as:

- What is the level of autonomy and decision levels required?
- What is the value from data? What are the data needs?
- How is the data captured? How much is captured?
- What are the key issues?

Requirements and preferences were captured using a variety of tools such as mind-maps, flow diagrams, and schematics (Turner et al., 2015), and were updated in a series of quarterly structured meetings. Further engagement with lower-level railway specialists, such as project managers, planners, and systems engineers enabled a deeper understanding of the network operation, internal processes, cost and repair practice, and observed railway standards. The derived requirements are summarized in Table 1. As seen, autonomy was one of the chief attributes. Dealing with an extremely large number of incidents can lead to ineffective maintenance planning, as well as incorrect human fault diagnosis (Dhillon, 2014). It is also desired that the new system does not rely on humans to enact optimal maintenance actions in a timely manner. To achieve this level of autonomy, the system must first be able to accurately infer when the assets have degraded and require intervention. Secondly, the system needs to infer and define optimal, cost-effective maintenance task sequencing.

Table 1
Requirement analysis summary for an automatic system for complex infrastructure maintenance scheduling.

Stakeholder requirement	System attribute
Autonomy	Autonomous response to asset degradation alarms and scheduling of maintenance jobs cost-effectively, with minimal or none supervisory input
Cost structure and accuracy	Overall maintenance costs estimation by comprehensive breakdown analysis; including incidents and denial of service charges
Fused output	Clearer situational awareness, maximize data utilization and reduction of storage and data management
Visualization	The system must display key asset, planning, resources and cost information in a logical, analytical and intuitive manner for risk evaluation. Dashboards are preferred due to their easy visualization and intuitive operation
Platform compatibility	Easier integration and alignment with current infrastructure systems, data streams, data-bases, etc.
Rail standards & knowledge observant	The system must acknowledge British railways standard procedures, protocols and incorporate current processes knowledge
Accurate location determination	To address uncertainty issues, precise location is a key enabler of future railway condition-based maintenance
Asset utilization	Accurate usage and loading of railway assets to enhance situational awareness

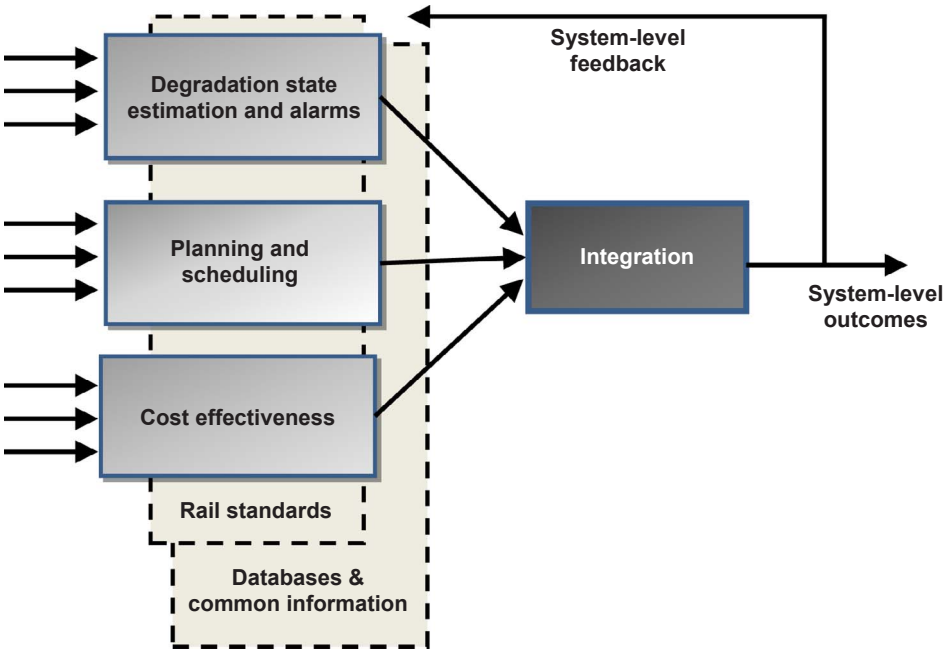


Fig. 2. Generic high-level integration architecture schematic.

2.2. Architecture development

Architectures are formal frameworks used to express the convergence of data and information from different sources. They comprise of a system of components whose structure and integration enable it to perform functions that the individual components could not otherwise accomplish (BS ISO 13374, 2007; Klein, 2004). From the previous requirement analysis, four work-stream components were devised:

1. *Degradation state estimation and alarms.* The objective of this component was to analyze asset degradation and raise timely trigger alarms for initiation of maintenance tasks planning. This required the implementation of an efficient data fusion strategy to collect and aggregate reports from the network sensors and mobile platforms, as well as real-time inspection reports.
2. *Planning and scheduling.* This component's objective was the automation of optimized maintenance tasks sequencing, and to produce actionable network maintenance schedules, considering fundamental operating maintenance parameters such as time, cost and staff availability.
3. *Cost analysis.* This component performed direct estimation of costs, as well as other strategic drivers associated with the maintenance plans scheduled.
4. *High level integratio.* This component was used to formally converge the individual work-stream outputs into a functional fused system output. This component also aggregated component reports for structured visual representation to the graphical user interface.

Following a series of technical discussions with the components research experts, the resulting high-level integrating architecture was developed using data fusion and condition monitoring principles (BS ISO 13374, 2007; Hall and Llinas, 1997) see Fig. 2.

The anticipated outputs from the first three components, degradation state estimation and alarms, planning & scheduling and cost analysis become the inputs of the overarching integration component. The integration system output delivers an optimized impact, availability, cost and capacity response that is based on the health state of declared fault entities. A feedback loop continually refines this process. The architecture also employs the underlying components "*common information*" and "*databases*". Common information refers to the input sources that the various components share of during the fusion process. In the railways context, asset location and the operational schedule sources are examples of common information sources. Databases store information and knowledge process inputs, such as railway standards criteria, rules, digital maps, degradation models, maintenance processes and tasks.

The high-level architecture shown in Fig. 2 was presented to the project stakeholders for concept evaluation. They considered the architecture appropriate for the delivery of current requirements and in agreement with their longer term operational technical strategy. Therefore the development of the research subsequent stages was approved.

3. Proof of concept demonstrator development

This section initially provides details of the demonstrator scope and assumptions, the lower architectural levels principles, the

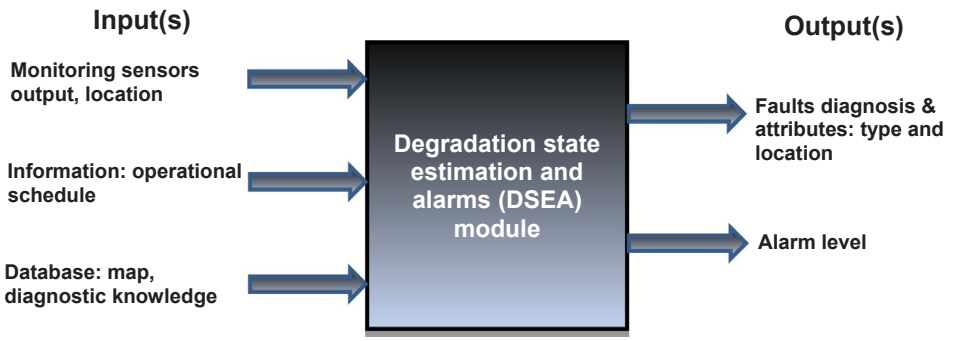


Fig. 3. Degradation state estimation and alarms and analysis and diagnosis module I/O analysis.

dataset characteristics, the system components interactions, as well as program specific algorithms tested.

3.1. System demonstrator scope and assumptions

The scope of the system was to demonstrate the integration of condition monitoring, planning and scheduling algorithms and cost analysis to automatically plan a large number of maintenance tasks. To achieve this, the research developed a proof of concept demonstrator; which was built using the architecture shown in Fig. 2 as a "blueprint". In agreement with the project stakeholders, a scaled-down system was prepared, dealing with the degradation monitoring, planning and scheduling and cost analysis of 10 railway track faults occurring over a 5-week period. This served to validate the system concept, its logic and to conduct elemental algorithm functionality tests. The demonstrator simulates how a real track incident and inspection dataset can be transformed into a discrete number of asset degradation severity alarms. These alarms trigger a scheduler algorithm which outputs an optimized intervention plan that factors-in a number of cost variables. Gradual asset degradation was assumed for all assets. Maintenance shift and flexible crews were both assumed to be available to respond to the incident alarms scenarios. Out of hours response incurred a supplementary cost. Maintenance depots and crews were assumed to be located within an 80 mile range of the intervention site.

A full scale national rail demonstrator is out of the scope this proof of principle research. Although a larger scale demonstrator can be attempted in subsequent iterations, this will likely be a commercial driven development.

3.2. Lower-level integration architecture development

The immediate lower levels of the architecture presented in Fig. 2 were derived using black box analysis (Green, 2014; Sánchez, 2007) to deduce each of the 3 architecture module input and outputs, and to determine specific functional, communication and performance requirements, see Figs. 3–5. At this point, the analysis was conducted without reference to the internal algorithmic structure of the individual components.

3.2.1. Degradation state estimation and alarms (DSEA) module: I/O analysis

Asset monitoring data, asset location as well as information sources are the fundamental inputs to the *Degradation state estimation and alarms* (DSEA) module. The fusion and analysis of these determines the health state of assets and their position on the network. The operational schedule input provide situational context. Data-base inputs provide location information such as maps, and also degradation knowledge for diagnostic inferences during the lower fusion processes. On inference of an actionable fault, the module generates an output alarm that initiates the scheduling process.

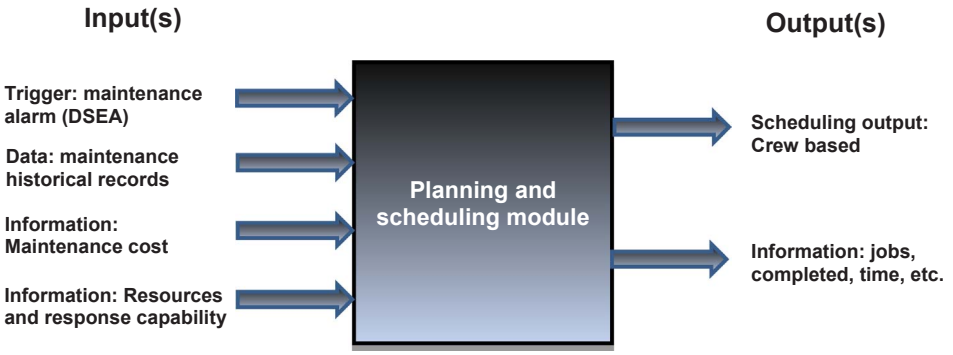


Fig. 4. Planning and scheduling module I/O analysis.

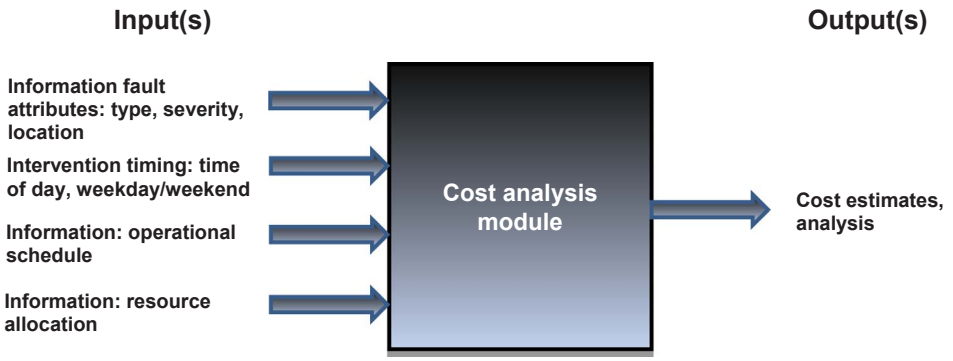


Fig. 5. Cost analysis module I/O analysis.

3.2.2. Planning and scheduling module: I/O analysis

The alarm raised by the DSEA module formally initiates the planning and scheduling module. In order to optimize planning, historical maintenance tasks and their sequencing must be first deconstructed and analyzed. The planning and scheduling module evaluates and mines maintenance repository input records. The cost of the maintenance activity is an important input parameter and is also factored-in. The available resources and response capability are fundamental inputs to deliver commensurate response plans. After optimally reconfiguring the maintenance data into business processes, the planning and scheduling module delivers a Gantt chart of all scheduled maintenance activities. Network maintenance information and progress parameters are also outputs delivered.

3.2.3. Cost analysis module: I/O analysis

The purpose of this module is to use identified cost drivers to model and estimate the overall maintenance costs and value. Its primary output is the estimation of the costs of the identified actionable faults by the DSEA module. The inputs required are the fault attributes: type, severity and location; which determine parts, direct labor and transportation costs. The planned intervention time input determines the costs incurred by service disruption and labor overtime; with the operational schedule providing operation situational context. Cost directly influences the planning of maintenance and any trade-offs.

3.3. System module interactions

We have used Unified Modelling Language (UML) to characterize and describe the system and modules interactions; and as backbone for the demonstrator integration programming. UML offers means to communicate complex information effectively using visual modelling (Holt, 2004). UML is widely used in railway related systems research; some example applications include signaling (Jabri et al., 2010) and reliability engineering (Bernardi et al., 2013). Standard practice is detailed in international standard procedure BS ISO/IEC 19501. A number of diagrams are specified: use case, class, statechart, activity, sequence, collaboration and component.

Use case diagram interactions are typically non-sequential. UML's use case diagrams utilize "actors" to initiate tasks and to describe the interaction between an actor and the system. The research evaluated a number of use case scenarios and UML diagrams (Turner et al., 2017) illustrating the modules interactions and behavior. The demonstrator top-level use case for an automatic planning and scheduling system is presented in Fig. 6. The system's modules are depicted in the system by 4 actors: integration, DSEA, planning and scheduling, cost and value analysis. The interactions in the use case can be broadly described:

- The DSEA actor updates the "identifies fault scenario" and "degradation monitoring" use cases from asset monitoring data. The "identifies fault scenario" use case is used and updated by the "integration" in the user Human Computer Interface (HCI) to display fault information, such as degradation trends, fault location, and severity.
- The "Triggers fault" use case by the DSEA actor initiates the systems response to the infrastructure degradation. The DSEA actor employs "identifies fault scenario" and "monitors degradation" use cases.
- The integration actor issues the "Generate plan & scheduling request" use case, using the "Triggers fault" use case.
- Both the Planning and Scheduling and the Cost Analysis actors use the "Generate plan & scheduling request" use case to action the "Plans maintenance and schedules tasks" and the "Estimates costs & value" use-cases, respectively.
- The "Plan maintenance and schedule tasks" use case uses the "Estimates costs & value" use case, at which point a Gantt chart is issued.
- The integration actor uses "Display degradation alarms" "Display impact-cost matrix" and "Display usage & resources information" use cases, for display in the HCI.

3.4. Demonstrator dataset description

The demonstrator focused on track asset incidents. Due to its overall importance, a reliable railway track is a priority for

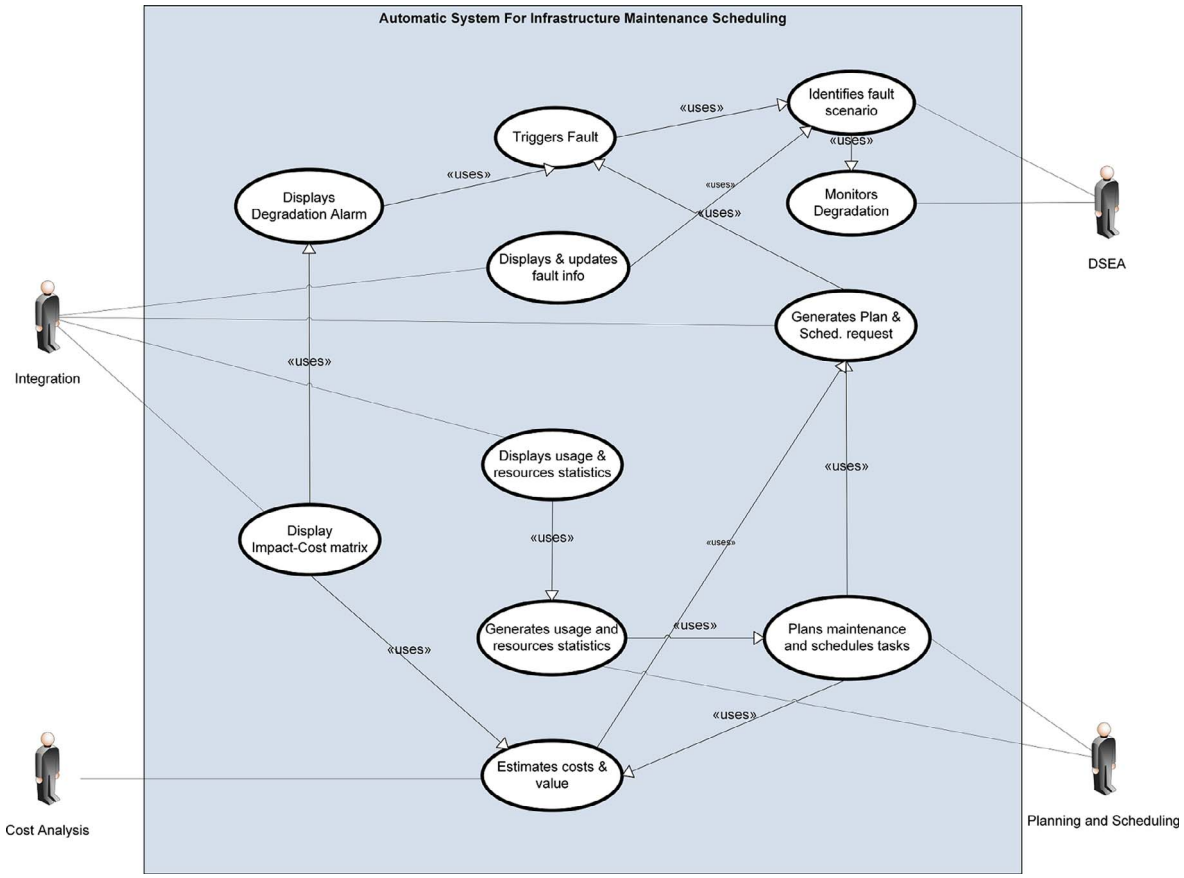


Fig. 6. UML use case illustration for automatic system for infrastructure maintenance scheduling.

stakeholders; and is also one the foundations of the railway's future technology strategy (TSLG, 2012).

Monthly infrastructure incident and inspection representative datasets were obtained. In agreement with Network Rail, the dataset used for analysis covered the route between Waterloo and Southampton stations. This is considered a primary line with continuous train traffic. The incident dataset comprised of twenty-nine descriptive fields, covering the financial year 2013–2014. The system principle was demonstrated using the data acquired over a 5-week period. During this time, 1991 track related incidents were reported. This time window was commensurate with the prescribed repair times of standard 2015 NR/L2/TRK/001/mod11 (Network Rail-3, 2015), which recommends a maximum of 28 days repair of severe faults.

3.4.1. Data pre-processing and preparation

Datasets as provided were not ready for the demonstrator input. A procedure was implemented to examine and prevent formatting errors, including checks for empty values, date logs and duplicates. Out of the twenty-nine descriptive fields in the original datasets, the ten most relevant were extracted for the demonstrator development. Table 2 shows the data fields used in the demonstrator. The data fields that were not used included for example: responsible manager, responsible organization, delivery unit

Table 2
Incident dataset fields used in demonstrator.

Field	Description
Job Number	Identification number
Due Date	Date issued
Importance	Job priority
Incident type	Text description
Location	Relative, distance from maintenance depot
Time estimate	Estimated time of job
Fault Group	Fault association
Alarm	ON/OFF Binary
Description	Fault description attribute
Cost	Cost incurred per incident

name and section end code, amongst others.

Although the dataset comprises a large number of class incidents, the system principle was demonstrated for the top five frequent group items:

- IS – Track defects (Other).
- IR – Broken/cracked/twisted/buckled/flawed rail.
- IZ – Other infrastructure.
- PB – Condition of the track.
- IT – Bumps reported - cause not known.

Typically, incidents description attributed failure to dips in the track, rail cracks, track circuit failures, and bumps in the track. Examples of other incidents not considered were attributed to trespass, vegetation, driver, etc. Faults severity was ranked 1 to 5; with 5 being the most critical fault requiring priority maintenance intervention, 4 is a warning alert, 3–1 indicate healthy or low priority interventions. This classification is in accordance to track geometry standards (Network Rail-3, 2015). A 6th level prescribing immediate line closure was not implemented in the demonstrator.

3.4.2. Forward-facing video and network maps feeds data

The incident dataset used was complemented with corresponding contextual visuals including driver's view video and maps of the affected network route map. For these, the proof of principle demonstrator used video feed repository drivers' training videos. Network route timetable data was obtained from the National Rail DARWIN information engine and the maps were fed from http://traintimes.org.uk.

3.4.3. Modular message passing

Integration served to link the executables from each module using C#. Message passing between the modules is illustrated in Fig. 7; as data log and scheduled log messages objects. These 2 objects are abstract representations of a scheduled job log containing maintenance incidents class attributes. They are passed in the form of a file with the methods enacted in the software modules.

3.5. Modules functionality, algorithms and implementation

Having established each of the system modules' I/Os, their interactions and the specific modules' functionality the modules algorithmic approach was developed.

3.5.1. HCI and visualization

The HCI of the demonstrator employs a dashboard. Dashboards provide graphical displays of interactive measurement-driven plots and gauges that depict trends, identify outliers, and drill-down capabilities. An analytics driven dashboard clearly concurs with the risk evaluation visualization that stakeholders require. The demonstrator dashboard observes established design principles (Selby, 2009):

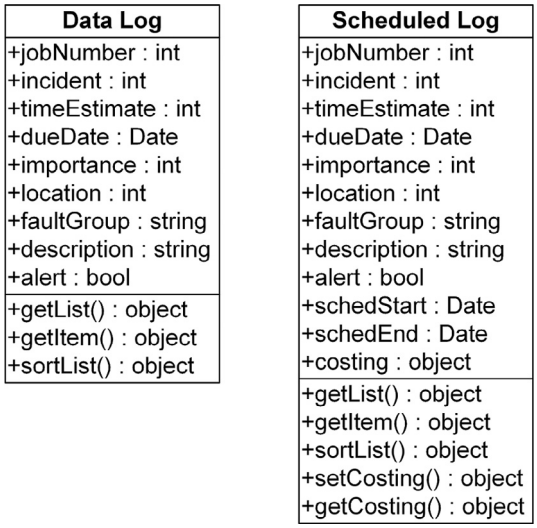


Fig. 7. Modular message passing characteristics. Left: data log for DSEA, Integration and planning and scheduling modules. Right data log for integration, planning and scheduling and cost analysis modules.



Fig. 8. Demonstrator HCI-dashboard for an automatic system for railways data-rich complex infrastructure maintenance scheduling integrating assets condition, planning and cost.

- support for different metric sets and different key performance indicators;
- utilization of different displays for different types of information, i.e. asset and information fault, maintenance planning updates and cost information;
- visualization of data trends;
- updates overall status using red-yellow-green indicators.

In addition, following the stakeholders request, the dashboard also integrated contextual train driver's view video and route network maps. The initial prototype used a Microsoft™ (MS) Excel constructed HCI dashboard because of its portability, widely available software license and stakeholder inter-changeability. MS Excel is also widely understood by engineers and railway practitioners, so any alterations suggested could be readily implemented. Fig. 8 shows an image of the dashboard used for the demonstrator.

3.5.2. Degradation state estimation and alarms (DSEA) module

The Degradation State Estimation and Alarms (DSEA) module raised degradation alarms that initiated the scheduler algorithm. It employed mid and low level data fusion processes to infer and present railway assets health, as well as their precise location in the network, which are this module's fundamental inter-modular message passing components. A combination of multiple measurements from fixed and mobile unsynchronized sources were used (Bevilacqua et al., 2015). Representative rail inspection datasets examined were substantially rich, reporting up to 13 measurements of rail-track quality such as twist or gauge every 0.22 yards. Additionally, railway operational schedules and maps were used for situational context. The fusion and visual analytics processes examined in our research included a range of estimation and statistical algorithms e.g. fuzzy logic (Dote and Ovaska, 2001) and Kalman filters (Boehringer, 2003), for collecting data, making health and location inferences and to report the information, depending on the types of data to be handled.

Simulation of degrading rail track measurement data was compared against rail standard thresholds to generate health state level alarms (Bacete, 2016). Current railway infrastructure asset health diagnosis rigorously adheres to rail standards that sanction pass/no-pass criteria for the data collected by the inspection trains and prescribe nominal intervention timescales. In early discussions with the project stakeholders we were advised that the current train location level of uncertainty could potentially impact the future asset management strategy. Therefore considerable efforts were put towards analyzing location uncertainties, performance of resolving algorithms and new measurement approaches for inertial measurement units (IMU), global navigation satellite system (GNSS), visual odometry and balises. For example, exploratory work conducted showed that the GNSS positional resolution of inspection trains could be improved from 1 m (nominal) to 30 cm by complimentary visual odometry; while also enhancing true positioning through GNSS "dark" signal areas (Bevilacqua et al., 2016).

A structured approach was followed to combine location and sensor data, see Fig. 9.

More comprehensive details of the work undertaken for this module, including data-fusion, location and degradation data analytics can be found in the following authors' publications (Bacete, 2016; Bevilacqua et al., 2015; Durazo-Cardenas et al., 2015; Loizillon, 2016). An example of the work conducted is shown in Fig. 10. This figure illustrates the evolution of the twist track

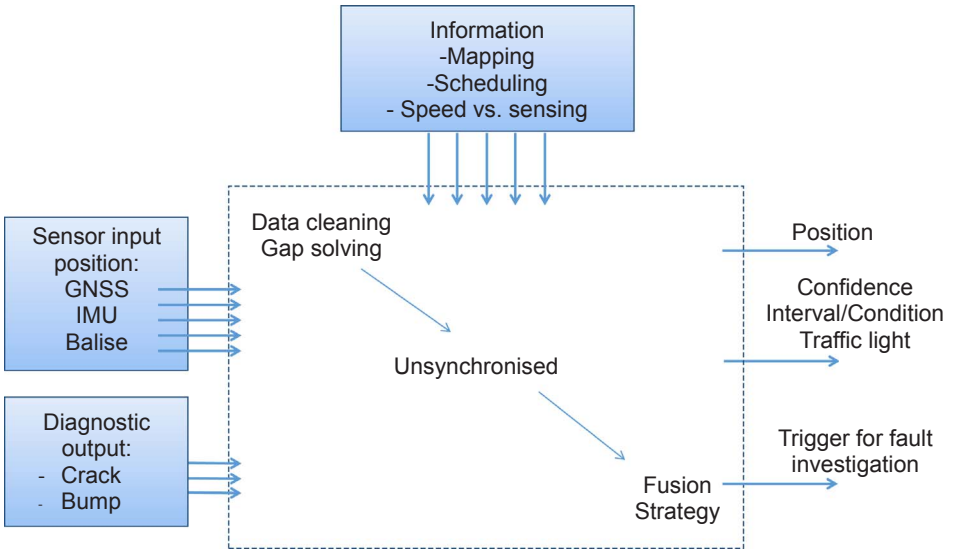


Fig. 9. Data Fusion strategy for integration of unsynchronized degradation and location sensors.

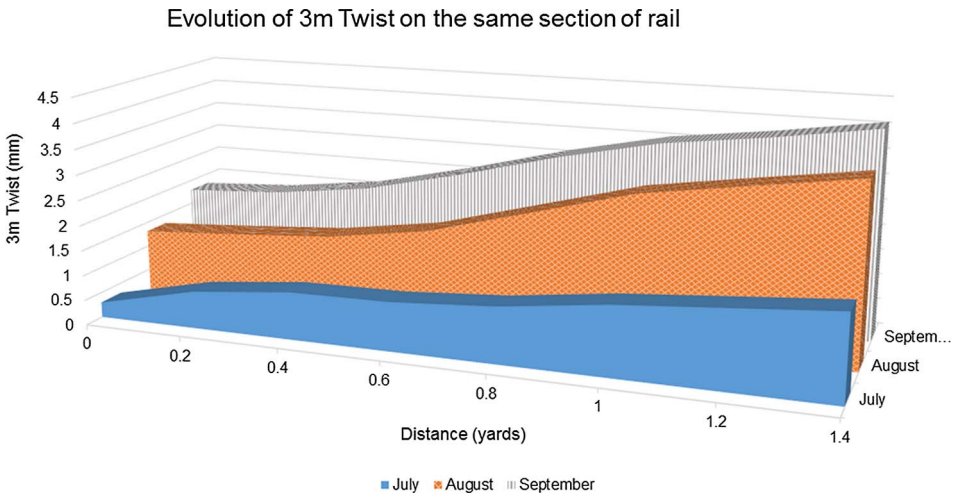


Fig. 10. Evolution of 3 m twist rail track quality parameter for a track section.

parameter over three consecutive months on the same track section obtained by analyzing inspection data. Through new visualization approaches, the progressive deterioration of track parameters can be more clearly observed.

For demonstration purposes, the system used a condensed dataset that included a prescribed number of asset degradation incidents and their location. Simulation was used to supplement intermediate degradation measurements.

3.5.3. Planning and scheduling module

3.5.3.1. Business process representation. The initial task of this module was the identification of the infrastructure maintenance processes that would more clearly benefit from optimization. It also conducted more detailed explorations in order to determine the more efficient scheduling of such tasks in an overall maintenance schedule. For demonstration purposes, a single rail maintenance process (tamping) was modelled along with a train washing process. The process representation activity helped to inform the types of maintenance activities impacting the railway infrastructure and their potential to influence planning and scheduling decisions. This work then evolved into the focus on modelling rail maintenance tasks as matching to groups of rail maintenance workers.

3.5.3.2. Planning and scheduling of maintenance jobs. The scheduling approach modelled the rail maintenance problem in terms of ten sets of work crews (with each set composed of a group of multi skilled rail workers). Five of the work crews were set to be available for jobs during the hours of 7 am – 7 pm. The remaining five were allotted high availability status meaning that they would be available for call out 24hrs a day. The data mining approach taken in this work utilized two algorithmic approaches, one based on a heuristic and a second that utilized a single objective Genetic Algorithm (GA). All software was written in C# (utilizing Microsoft

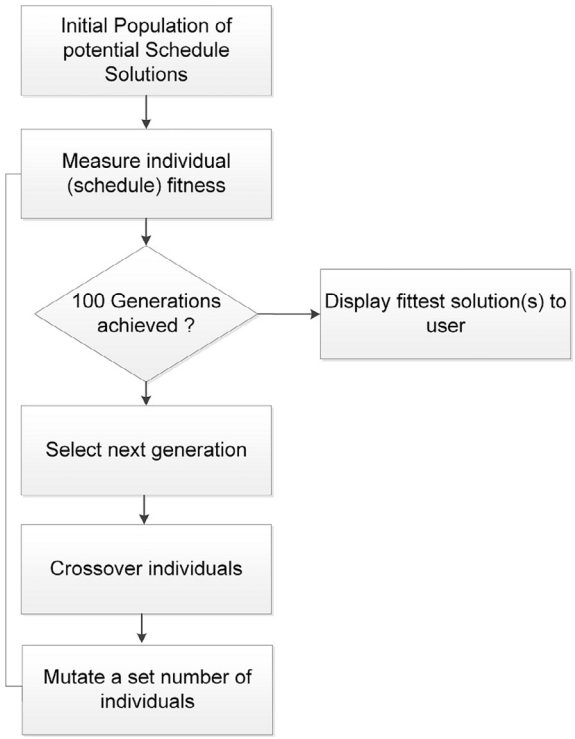


Fig. 11. Scheduling generic algorithm flowchart.

Visual Studio) and included a windows desktop (forms) user interface to enable user interaction with the system demonstrator.

With the heuristic approach the following rules were applied:

- Maintenance jobs for scheduling are divided into 5 groups based on priority with level 5 jobs marked as the highest priority (presenting highest fault risk severity, as defined in Section 3.4.1);
- Provide ability for user to dynamically raise and lower fault group priorities and mix job types;
- Schedule closely located jobs first (with regard to the physical location of the work crews maintenance depot);
- Schedule jobs according to fault type (5 different fault type groups were specified with the ability to prioritize jobs falling into specific fault group types).

The demonstrator interface allows users to experiment with different parameters and provides the ability to raise the priority of jobs based on their priority level, fault type and individual characteristics. While the heuristic approach provides flexibility to the user to experiment with different scheduling parameters it does not allow a full exploration of the possible solution space. Therefore an additional data mining technique utilizing an evolutionary soft computing approach was proposed. A GA is utilized to find the optimized schedule with a given set of jobs. While not guaranteeing to find the perfect solution soft computing approaches such as GA provide the opportunity to efficiently explore the possible search space to identify the 'fittest' scheduling option satisfying a given objective or objectives while meeting a given set of constraints. In this approach a single objective is set, that of minimization of cost. Fig. 11 shows the main steps of the single objective GA employed. The GA utilizes a population of solutions with each individual being a complete potential schedule. The individuals (or chromosomes) of the GA are composed of the job tasks in order sequence of completion. The initial population is generated from a random ordering of jobs composing a maintenance schedule (each of these jobs is referred to as a gene). In addition:

- each gene contains 'pointer' to the detail of the job;
- gene (job) sequences within chromosomes can be swapped (Crossover);
- individual genes (jobs) can be swapped (Mutation);
- each generated solution constitutes schedule structure for fitness assessment;
- the cost generated for each solution (chromosome).

In terms of crossover sequences of jobs can be swapped between 10 of randomly selected individuals (with the swap point chosen at random) for each generation. In terms of mutation an individual job can be changes in the sequence of 5 randomly selected individuals for each generation. The following constraints are also recognized within the GA:

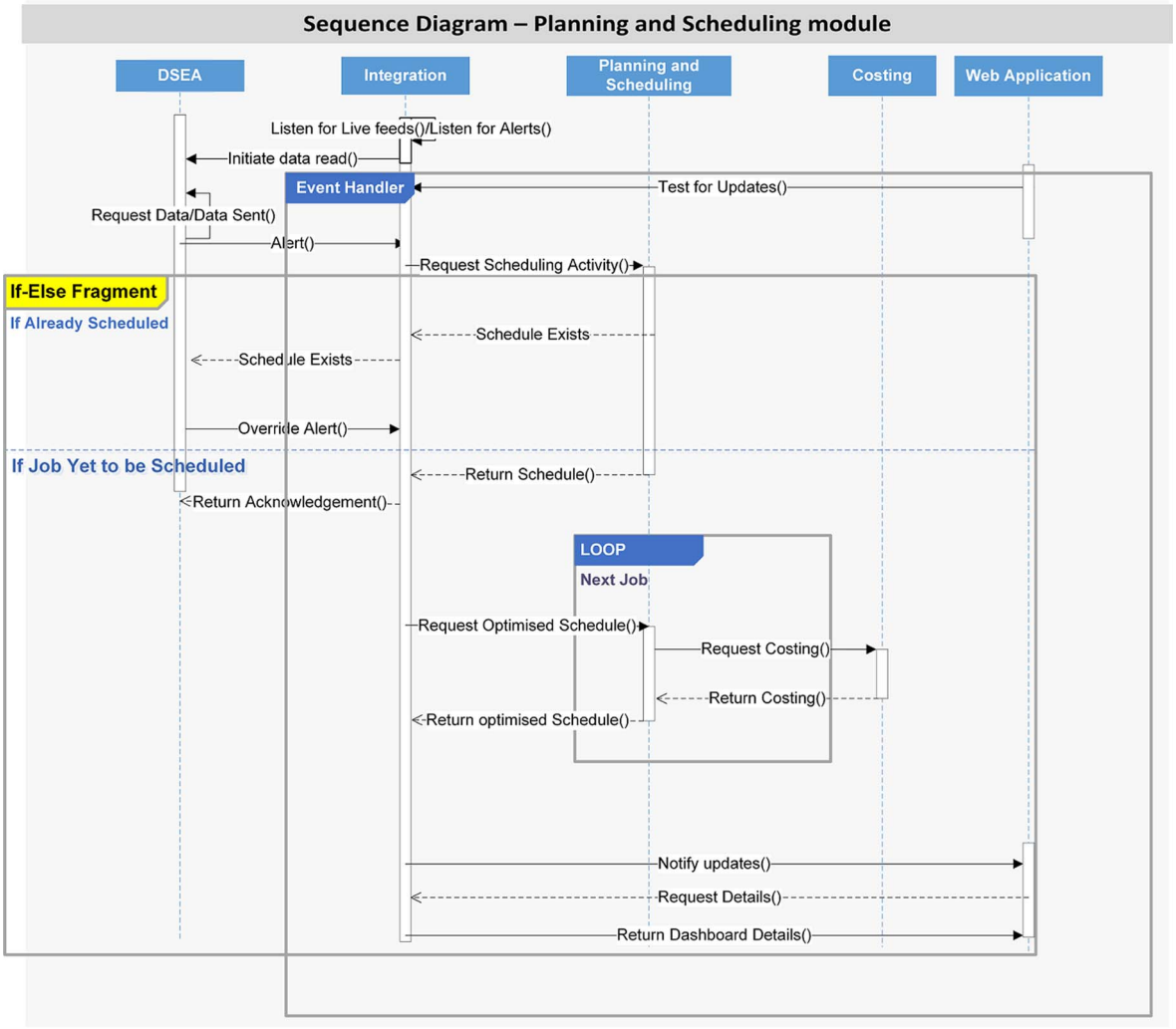


Fig. 12. Sequence diagram for the planning and scheduling module.

- if a job cannot be scheduled at its start date a later slot is tried;
- if selected the user can specify that jobs located near to the depot are undertaken first;
- high priority jobs will be scheduled first.

The fitness function for the GA has the single objective of minimizing cost of the schedule and the following job scheduling rules are encoded:

- every schedule is costed (the overall cost is calculated);
- jobs scheduled to 24hr availability work crews cost more to complete;
- late scheduled jobs are penalized (higher cost);
- highest priority jobs have an additional cost penalty;
- location of job can be taken into account (if the option is selected by the user).

In terms of the job costing each job has a base cost which may be increased to account for travel time and the mix of job types allotted to the same work crew on a given day

Fig. 12 illustrates planning and scheduling module process sequence and the interactions of the demonstrator modules that resulted in the generation of work plans.

3.5.3.3. Scheduler algorithm development. From conception, the planning and scheduling module aimed to take advantage of modern data stores such as Hadoop. When scheduling scenarios are created they are also saved by default by the system so they may be returned to the user for comparison with other generated scheduling solutions.

Initially we intended to develop a multi-objective approach to scheduling involving the optimized trade-off between time and cost. Experiments to this end involved the utilization of a Genetic Algorithm (GA) to provide a Pareto front of solutions to the user showing optimized scheduling scenarios for different objective trade-off points. Standard multi-objective GA algorithm implementations such as NSGAII (Deb et al., 2002) and PAES (Knowles and Corne, 2000) were evaluated as part of this experimentation phase. In addition a GA approach to business process optimization was investigated for its suitability to planning and scheduling. The outcome of the aforementioned experimentation was the conclusion that it was in fact possible to fold the two objectives of time and cost into a single objective of cost. This led to a further investigation of single objective approaches to scheduling. After further research it was possible to identify a single objective GA approach that was suited to the problem posed in the autonomous scheduler demonstrator. The genetic algorithm optimizes the schedule and prioritizes those jobs at the highest alarm level. For every schedule the overall cost was calculated. The cost model and fitness function take into account out-of-hours penalties, denial of service, and location of the job.

3.5.4. Cost analysis module

The cost module development required the identification of a suitable cost breakdown structure. Initial analysis considered its structure to consist of four key cost elements: fault detection, maintenance planning, maintenance activity and “denial of service” cost; with a considerable number of identified potential drivers for these elements (Amorim-Melo et al., 2014; Carlander et al., 2016). Full implementation into the automatic demonstrator, however, proved challenging because current maintenance cost records are not structured for this purpose. Furthermore, the different cost drivers also have different degrees of significance for the overall maintenance cost. Through discussion with industry experts/stakeholders it was known that denial of service charges would be one of the dominant cost drivers and, in some situations, could impact the maintenance cost even more than the cost of doing the maintenance activity itself. The cost analysis module therefore had to prioritize the cost of the activity and the cost of denial of service. Previous research (García Márquez et al., 2008) also discussed denial of service, but instead they focused on establishing the average disruption at a yearly level, rather than per incident level the model developed used here.

The “maintenance planning cost” and “fault detection cost” drivers were not considered within the revised cost breakdown structure. The reduced cost breakdown structure did not reduce the complexity because both ‘Denial of Service Cost’ and ‘Maintenance Activity Cost’ have discontinuous time dependencies. For example, labor rates are highly time dependent with effects from overtime and/or weekends influencing which rate is applied.

3.5.4.1. Maintenance activity cost. From the literature available, the work of Patra (Patra et al., 2009) guided the estimation of the cost of activities. From their work, it was expected that linear equations that used length of the track section worked on as a cost driver would be largely suitable for estimating activity costs. The activity equations presented by Patra et al. (2009) include a non-linear component within the denominator; which is an attempt to use Net Present Value calculation methods to future proof the estimates. This would not be suitable if the rates used were periodically updated, as was assumed within the cost analysis module of this work. Another example from the literature of a similar approach is (García Márquez et al., 2008). We can be confident of the relevance and validity of our approach when building upon the approaches used by these authors. For the demonstrator a linear model was built for each error code, which estimates material costs associated with the activity. The scheduling and planning module passes estimates of task duration, therefore the labor cost and each hour of the task could be calculated separately. This allowed for any overtime calculations to be made should a task overrun typical working agreements. This gave the activity cost some time dependent behavior. Greater time dependency of cost was introduced when denial of service costs were estimated.

3.5.4.2. Denial of service cost. Network Rail’s denial of service costs are primarily linked to schedule 8 fines levied by regulators (ORR-1, 2012). Our analysis of denial of service costs focused upon three drivers for the scale of these fines: time and day, error code (job type) and location of fault upon the network (route criticality). For Network Rail the most relevant location information is the Strategic Route Section (SRS) and the Route Criticality Banding (Ove Arup and Partners Ltd, 2013). Route criticality banding divides the strategic route sections into 5 bands. This simplifies deciding the strategies to deploy across the network for asset management. This system clearly indicates the expected scale of denial of service fines by route. The band definitions are (Ove Arup and Partners Ltd, 2013):

- Band 1: SRS with costs per incident more than two times the mean;
- Band 2: SRS with costs per incident between the mean and two times the mean;
- Band 3: SRS with costs per incident between the mean and half the mean;
- Band 4: SRS with costs per incident between half the mean and one quarter the mean;
- Band 5: SRS with costs per incident less than one quarter the mean.

The location (and therefore the criticality band number) helps to reduce the uncertainty associated with denial of service cost estimation. Separate denial of service cost estimating relationships based upon the banding structure have been developed for each error code used in the demonstrator. These cost estimating relationships are modified by time of day and weekday/weekend variables to estimate the cost of denial of service. The overall cost module executional sequence is illustrated in Fig. 13. The sequence is initiated from a cost estimation request from the Planning and Scheduling module. As seen, the cost module performs the estimation of the activity cost based on labor and materials costs; then the denial of service estimate element is taken into account. Maintenance and cost databases are used to support the overall single figure of merit cost estimation returned to the Planning and Scheduling module.

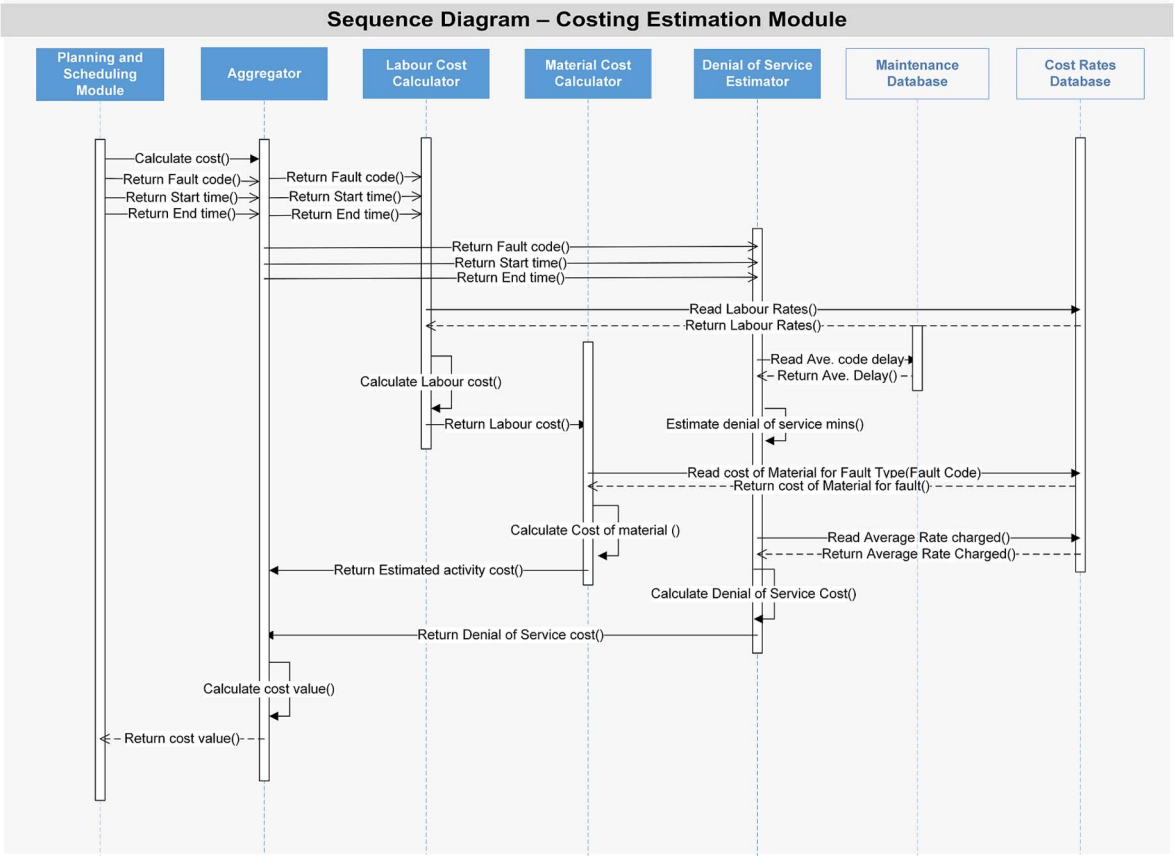


Fig. 13. UML sequence diagram representing the sequential operation of the cost module in the system demonstrator.

3.5.4.3. Cost analysis module limitations. Due to the integration requirements of the system demonstrator, the cost module development had a number of constraints; including:

- reduced scope of the data available for building the estimate;
- suitable for integration within the system as a whole;
- compatibility with the genetic algorithm based Planning & Scheduling module;
- satisfy stakeholder requirements.

Naturally, the scaled-down scope of the demonstrator reduced the number of possible cost drivers that otherwise could be taken into account in the analysis. However, in this paper our intention is to demonstrate cost integration principles using industry's most relevant drivers. Within the greater system demonstrator, the cost module also has to exchange information to other modules, most frequently with the Planning & Scheduling module; which uses a genetic algorithm. To avoid higher computational overheads in the functional demonstrator, uncertainty simulation e.g. Monte Carlo was not performed. However, in-depth cost analysis in the demonstrator is currently undergoing and has been reported separately by the authors (Kirkwood et al., 2015, 2014). Uncertainty modelling will certainly play a part in future automatic planning and scheduling engines.

4. Summary and discussion

Systems engineering and data fusion principles were used to demonstrate an autonomous data-rich infrastructure maintenance scheduling system that integrates railways asset condition, planning and cost.

The stakeholders' system requirements, desired attributes and deliverables were formally captured and analyzed. From these, a cross-disciplinary high-level architecture was developed. The new architecture comprises four fundamental components/modules: integration, degradation state estimation and alarms (DSEA), planning and scheduling and cost analysis. Technical discussions with railway engineers confirmed the emphasis that the railway industry places on tacit knowledge and rules employed for fault diagnosis and decision making.

We employed a black box approach to derive each of the modules underlying inputs and outputs required for the system functionality. This led to the selection of algorithms and tools necessary to achieve the modules outputs. Unified modelling language was

used to illustrate the system modules interactions and formulate message passing. This helped dealing with the system design complexities and enabled smoother code development.

The HCI was the platform for a demonstrator that simulates how curated rail-track incident and inspection datasets are transformed into a discrete number of asset degradation severity alarms. These alarms automatically triggered a scheduler algorithm which outputs an optimized intervention plan that factored-in a number of cost variables. The dashboard used:

- Assets monitoring condition “traffic light” alarm display. Incident severity was ranked from 1 to 5; with 5 being the most critical fault (red), 4 issues a warning alert (amber), levels 3–1 reported degradation levels within rail standards acceptable tolerances (green);
- Asset degradation trends;
- Affected services maps;
- Drivers view of affected route/services;
- Scheduled maintenance operations Gantt and resource usage statistics;
- Planned maintenance cost break down;

Track degradation monitoring has been improved by extracting key information from large inspection datasets, to characterize and plot degradation progression; in a more intuitive and informative manner, as shown in Fig. 10. This figure illustrates the evolution of the twist track parameter, of the same segment of track on the route to Southampton station. This degradation visualization was constructed using data analytics from inspection data acquired over three consecutive months during 2015. As seen twist progressively deteriorates. This trend can be used to signal interventions and anticipate further degradation and to assess the effectiveness of past interventions.

The planning and scheduling task scheduler used heuristics and genetic algorithms to enable the allocation of maintenance jobs to group of crews, producing cost effective optimized plans to deal with incidents in a clear and timely manner. This is illustrated in Fig. 14; which shows one of the Gantt charts demonstrated. In this instance, the Gantt shows the crews daily work load in response to a track circuit failure. The autonomous scheduler is also capable of summarizing and displaying key performance indicators.

The cost analysis module uses a simple cost-breakdown structure to estimate cost of maintenance; but because of the complexity of the effect of time of the maintenance task delivery, multiple values are possible for any given task. Building upon the work of Patra (Patra et al., 2009) means that the approach to maintenance cost analysis has historically been useful. Previous work by the authors (Kirkwood et al., 2015) on denial of service has shown the complexity of that behavior and that using a single probability distribution function can struggle to describe denial of service. The demonstrator captures much of this complexity by accounting for time of day, weekday/weekend and location (route criticality) as determining factors.

To the authors knowledge this is the first time a cost engineering approach has been modified to comply with the challenges of integration within an autonomous system. Therefore, while the cost model presented does not have any particularly innovative

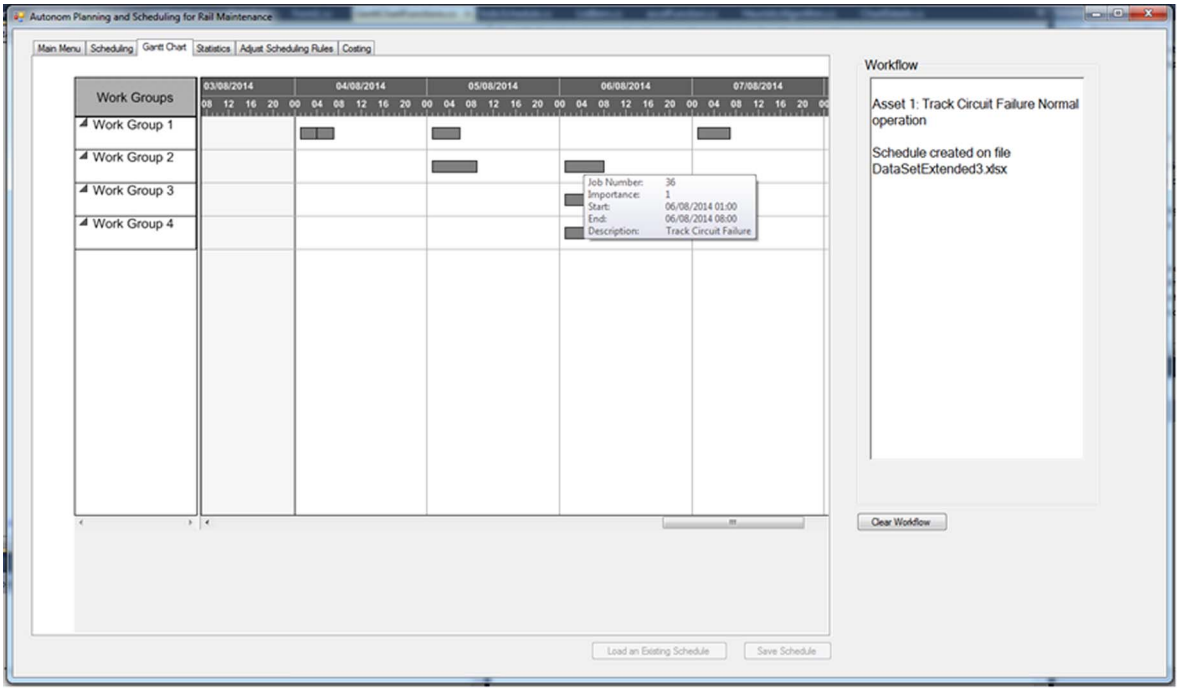


Fig. 14. Autonomous jobs by crew schedule illustration.

methods, the resulting application makes it very novel. While many cost models exist for various maintenance activities, most focus on the cost of the activity and ignore denial of service costs associated with asset downtime. Such an approach is incomplete and will result in poor decisions being made. The approach to developing the cost models used here was deliberately multi-industry: should other industries seek to apply the approach to their cost estimation challenges, the focus on maintenance activity and denial of service should make the implementation less challenging.

Despite modern developments, the British rail infrastructure is still heavily reliant on informed human decision-making, involving the interpretation and contextualization of multiple-sources of information. The British rail so far has benefited by the development of asset management tools like infrastructure inspection trains and decision tools such as LADS (ORR, 2014). LADS enables rail specialists to contextualize a large number of the available information sources. LADS however, does not autonomously output maintenance plan decisions that fuse asset degradation and cost engineering inputs.

With the continual growth of rail services demand and larger monitoring data outputs, the capability to automatically respond to an even larger number of maintenance interventions timely and cost effectively, becomes crucial. This research's approach has used controlled simulation scenarios from real data sets to demonstrate feasible cost-effective intervention planning in a complex data-rich environment; where a multitude of simultaneous interventions are required.

Our research contribution emanates from a systems integration of asset monitoring, state of the art planning and scheduling algorithms and thorough cost modelling. Automatic scheduling of maintenance tasks clearly advances practical systems fusion contribution, from the sensor fusion stage towards resource management; which is often overlooked, as noted in the literature (Scholz and Gossink, 2012).

Our system demonstrator is a successful proof of principle for a British main rail line. We have used curated datasets for demonstration purposes, but future endeavors should attempt to employ railways data streams. Clearly a larger scale strategy will be required to address the British rail needs. Commercial tools and developers will play a critical role to realize the benefits presented.

The approach proposed will contribute to greatly abating unplanned infrastructure maintenance expenditure. This was recently estimated at £120 million per year (TSLG, 2012). Accurate cost savings assessment is challenging because of the vast number of factors involved; but even a conservative 10% reduction will represent substantial savings and smoother railway operations. The system response is not only expected to reduce interventions, but to more cost-effectively plan those required.

4.1. Validation

Railway infrastructure experts were regularly consulted throughout of the research to validate specific module components. However, at the end of the research, formal face validity of the integrated proof of principle demonstrator and its components was conducted. This involved a number of demonstration sessions with stakeholders and railway infrastructure specialists. Over twenty-five individuals were consulted. The system architecture, fault alarm and visualization, planning and cost components were evaluated to ensure satisfaction, alignment with rail standards and best practice. Face validation is considered a key measure of system design success (DeLone and Mclean, 2003; DeLone and McLean, 1992; Yang, 2012).

Both, direct interaction with individuals and group settings approaches were used to collect the views of railway experts and stakeholders. Questionnaires and unstructured interviews were used to acquire the data. Responses were usually elicited after a software demonstration. In the case of questionnaires, the participants usually responded to 10–20 questions. Specific demonstrator attributes were evaluated using closing ended questions and a 1–5 rank system, where 1 was least valuable or poor, and 5 was most valuable or excellent. Questionnaires also captured feedback and desired attributes using a small number of specific open-ended questions.

In general, the conceptual model demonstrator and job planning sequencing was well-accepted by railway experts and it was also seen in agreement with British railways longer term strategy. The visual output provided by the demonstrator dashboard is very powerful, illustrating simulated real-time fault monitoring trends and alerts monitoring, cost-weighed optimal processes planning autonomously. Its structure was found fit for purpose with logic component relationships. Closed ended questions rank typically averaged 3.9; which means the system was well accepted. Suggestions have been noted for the future evolution of the Gantt chart screen so that it can highlight priority jobs to users and allow for data drill down so more information on jobs may be displayed to a user. It is also envisaged that with connection to staff roster systems the scheduling may be able to take account of the real time composition of work crews. Higher accuracy in fault location and overall system network awareness is expected to optimize intervention planning. It is also the case that it is difficult to identify scheduling clashes with the existing application though further enhancements towards providing a fully autonomous version of the demonstrator may address this need. Naturally, a formal benchmark evaluation was not possible due to the lack of existing benchmark response systems addressing similar scenarios.

5. Conclusions

Despite modern developments, the British rail network maintenance intervention planning still employs informed human decision-making, involving the interpretation and contextualization from multiple-sources of information. As demand for new train services increase and the network infrastructure modernizes, planning for an ever increasing number of interventions will be required. This demands systems that not only support the decision making process, but truthfully and cost-effectively schedule the necessary interventions autonomously.

An integrated approach that fuses asset monitoring, planning and scheduling and cost has demonstrated a feasible option to achieve the automatic planning and reduced maintenance costs demands. Our system architecture design and underlying modules are

generic enough to be applicable to a range of scenarios, including complex systems with abundant sensors and monitoring systems, for example in oil rig drilling, nuclear decommissioning and marine systems. It can raise asset degradation alarms that trigger the system response, supported by information sources and expert systems, resulting in cost-factored automatic maintenance tasks and resource sequence.

Our system design has been well received by the British rail industry. It clearly is in line with their future strategy. We have set the foundations for future autonomous infrastructure maintenance planning execution. In further technology readiness iterations, a larger scale strategy and approach will be clearly required.

Acknowledgment

The authors gratefully acknowledge the support of the EPSRC through the Autonomous and Intelligent Systems (AIS) partnership grant EP/J011630 “Integrated through-life support for high value systems". The authors also grateful for the support provided by the AIS partners: principally Network Rail, and BAE Systems, DSTL, European Space Agency, NNL, Scisys, Schlumberger, Sellafield Ltd, and the UK Space Agency. Special thanks to Kevin Blacktop from Network rail for facilitating timely access to railway experts and information exchange throughout the research.

References

Amorim-Melo, P., Shehab, E., Kirkwood, L., Baguley, P., 2014. Cost drivers of integrated maintenance in high-value systems. Procedia CIRP 22, 152–156. http://dx.doi.org/10.1016/j.procir.2014.08.011.

Andrews, J., 2012. A modelling approach to railway track asset management. Proc. Inst. Mech. Eng. Part F J. Rail Rapid. Transit. 227, 56–73. http://dx.doi.org/10.1177/0954409712452235.

Anon., 2015. Freight Rail Usage 2015-16 Q1 Statistical Release. London.

Bacete, M., 2016. Visual Analytics to Support Railway Infrastructure Services. Cranfield University. < http://cclibweb-3.central.cranfield.ac.uk/handle/1826.1/12555 > .

Bernardi, S., Flammini, F., Marrone, S., Mazzocca, N., Merseguer, J., Nardone, R., Vittorini, V., 2013. Enabling the usage of UML in the verification of railway systems: the DAM-rail approach. Reliab. Eng. Syst. Saf. 120, 112–126. http://dx.doi.org/10.1016/j.ress.2013.06.032.

Bevilacqua, M., Tsourdos, A., Starr, A., 2016. Egomotion estimation for monocular camera visual odometer. In: 2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings. IEEE, pp. 1–6. http://dx.doi.org/10.1109/I2MTC.2016.7520579.

Bevilacqua, M., Tsourdos, A., Starr, A., Durazo-Cardenas, I., 2015. Data Fusion Strategy for Precise Vehicle Location for Intelligent Self-Aware Maintenance Systems, in: Proceedings – International Conference on Intelligent Systems, Modelling and Simulation, ISMS. 10.1109/ISMS.2015.37.

Boehringer, F., 2003. Train location based on fusion satellite and train-borne sensor data. In: Zhao, Y. (Ed.), International Society for Optics and Photonics, p. 76. doi:10.1117/12.487062.

Bouillaut, L., Francois, O., Dubois, S., 2012. Optimal metro-rail maintenance strategy using multi-nets modeling. Int. J. Perform. Eng. 8, 77–90.

BS ISO 13374, 2007. Condition Monitoring of Machine Systems-data Processing, Communication and Presentation. BSi British Standards.

Carlander, L., Kirkwood, L., Shehab, E., Baguley, P., Durazo-Cardenas, I., 2016. Integration of Cost-risk Assessment of Denial of Service within an Intelligent Maintenance System. In: Procedia CIRP. doi:10.1016/j.procir.2016.03.229.

Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans. Evol. Comput. 6, 182–197. http://dx.doi.org/10.1109/4235.996017.

DeLone, W.H., Mclean, E.R., 2003. The DeLone and McLean Model of information systems success: a ten-year update. J. Manag. Inf. Syst./Spring 19, 9–30. http://dx.doi.org/10.1073/pnas.0914199107.

DeLone, W.H., McLean, E.R., 1992. Information systems success: the quest for the dependent variable. Inf. Syst. Res. 3, 60–95. http://dx.doi.org/10.1287/isre.3.1.60.

Dhillon, B.S., 2014. Human error in maintenance: an investigative study for the factories of the future. IOP Conf. Ser. Mater. Sci. Eng. 65, 120–131. http://dx.doi.org/10.1088/1757-899X/65/1/012031.

Dote, Y., Ovaska, S.J., 2001. Industrial applications of soft computing: a review. Proc. IEEE 89, 1243–1265. http://dx.doi.org/10.1109/5.949483.

Durazo-Cardenas, I., Amaraegbeni, C., Starr, A., 2016. Fusion of Railway Network Data Streams for Asset Usage in Intelligent Maintenance Systems. In: EUROMAINTENANCE 2016. EFNMS, Athens, Greece, pp. 59–64.

Durazo-Cardenas, I., Starr, A., Turner, C.J., Kirkwood, L., Maurizio, B., Baguley, P., Essam, S., 2015. An Intelligent Maintenance System for the Rail: Integrating Condition, Planning and Cost. In: Railway Engineering-13th International Conference & Exhibition. Edinburg, Scotland.

Esteban, J., Starr, A., Willetts, R., Hannah, P., Bryanston-Cross, P., 2005. A Review of data fusion models and architectures: towards engineering guidelines. Neural Comput. Appl. 14, 273–281. http://dx.doi.org/10.1007/s00521-004-0463-7.

García Márquez, F.P., Lewis, R.W., Tobias, A.M., Roberts, C., 2008. Life cycle costs for railway condition monitoring. Transp. Res. Part E Logist. Transp. Rev. 44, 1175–1187. http://dx.doi.org/10.1016/j.tre.2007.12.003.

Green, M., 2014. The Application of Black Box Theory to System Development. In: Systems Engineering in DC Proceedings, SEDC, 2014.

Guler, H., 2017. Optimisation of railway track maintenance and renewal works by genetic algorithms. J. Croat. Assoc. Civ. Eng. 68, 979–993. http://dx.doi.org/10.14256/JCE.1458.2015.

Guler, H., 2013. Decision support system for railway track maintenance and renewal management. J. Comput. Civ. Eng. 27, 292–306. http://dx.doi.org/10.1061/(ASCE)CP.1943-5487.0000221.

Hall, D.L., Llinas, J., 1997. An introduction to multisensor data fusion. Proc. IEEE 85, 6–23. http://dx.doi.org/10.1109/5.554205.

Hall, D.L., McMullen, S.A.H., 2004. Mathematical Techniques in Sensor Data Fusion. Artech House inc, Norwood, MA.

Hess, A., Calvello, G., Frith, P., 2005. Challenges, issues, and lessons learned chasing the “Big P”: real predictive prognostics part 1. IEEE Aerosp. Conf. Proc. http://dx.doi.org/10.1109/AERO.2005.1559666.

Holt, J., 2004. UML for Systems Engineering: Watching the Wheels, 2nd ed., pp. 373.

Jabri, S., El Koursi, E.M., Bourdeaud'huy, T., Lemaire, E., 2010. European railway traffic management system validation using UML/Petri nets modelling strategy. Eur. Transp. Res. Rev. 2, 113–128. http://dx.doi.org/10.1007/s12544-010-0030-5.

Jianhui, L., Namburu, M., Pattipati, K., Liu Qiao, Kawamoto, M., Chigusa, S., 2003. Model-based prognostic techniques [maintenance applications]. In: Proceedings AUTOTESTCON 2003. IEEE Systems Readiness Technology Conference. IEEE, Piscataway, NJ, pp. 330–340. doi:10.1109/AUTEST.2003.1243596.

Jovanovic, S., 2004. Railway track quality assessment and related decision making. In: Conf. Proc. – IEEE Int. Conf. Syst. Man Cybern., vol. 6, pp. 5038–5043. doi:10.1109/ICSMC.2004.1400992.

Jovanović, S., Božović, D., Tomičić-Torlaković, M., 2014. Railway infrastructure maintenance planning based on condition measurements and analysis. J. Croat. Assoc. Civ. Eng. 66, 347–358. http://dx.doi.org/10.14256/JCE.959.2013.

Khaleghi, B., Khamis, A., Karray, F.O., Razavi, S.N., 2013. Multisensor data fusion: a review of the state-of-the-art. Inf. Fusion 14, 28–44 doi:10.1016/j.inffus.2011.08.001.

Kirkwood, L., Shehab, E., Baguley, P., Amorim-Melo, P., Durazo-Cardenas, I., 2014. Challenges in cost analysis of innovative maintenance of distributed high-value assets. In: Procedia CIRP. doi:10.1016/j.procir.2014.08.006.

Kirkwood, L., Shehab, E., Baguley, P., Starr, A., 2015. Uncertainty of net present value calculations and the impact on applying integrated maintenance approaches to the UK rail industry. Procedia CIRP 245–249. http://dx.doi.org/10.1016/j.procir.2015.07.056.

Klein, L.A., 2004. Sensor and Data Fusion: A Tool for Information Assessment and Decision Making. SPIE Press.
Knowles, J.D., Corne, D.W., 2000. Approximating the nondominated front using the Pareto Archived Evolution Strategy. Evol. Comput. 8, 149–172. http://dx.doi.org/10.1162/106365600568167.
Le, B., Andrews, J., 2013. Modelling railway bridge asset management. Proc. Inst. Mech. Eng. Part F J. Rail Rapid. Transit. 227, 644–656. http://dx.doi.org/10.1177/0954409713490924.
Li, H., Parikh, D., He, Q., Qian, B., Li, Z., Fang, D., Hampapur, A., 2014. Improving rail network velocity: a machine learning approach to predictive maintenance. Transp. Res. Part C Emerg. Technol. 45, 17–26. http://dx.doi.org/10.1016/j.trc.2014.04.013.
Lidén, T., 2015. ScienceDirect Railway infrastructure maintenance – a survey of planning problems and conducted research. Transp. Res. Procedia 10, 574–583. http://dx.doi.org/10.1016/j.trpro.2015.09.011.
Lidén, T., Joborn, M., 2017. An optimization model for integrated planning of railway traffic and network maintenance. Transp. Res. Part C Emerg. Technol. 74, 327–347. http://dx.doi.org/10.1016/J.TRC.2016.11.016.
Liggins, M.E., Hall, D.L., Llinas, J., 2009. Handbook of Multisensor Data Fusion – Theory and Practice. CRC Press.
Loizillon, P., 2016. Prognostics Conceptual Modelling Framework to Support Railway Infrastructure Services. Cranfield University. < http://cclibweb-3.central.cranfield.ac.uk/handle/1826.1/12577 > .
Magee, C.L., de Weck, O.L., 2004. Complex system classification. Incose 18. http://dx.doi.org/10.1002/j.2334-5837.2004.tb00510.x.
Morant, A., Larsson-Kråik, P.-O., Kumar, U., 2016. Data-driven model for maintenance decision support: a case study of railway signalling systems. Proc. Inst. Mech Eng. Part F J. Rail Rapid. Transit. 230, 220–234. http://dx.doi.org/10.1177/0954409714533680.
Nair, N., 2015. Passenger Rail Usage 2014-15 Quarter 4 Statistical Release. London.
Network Rail-1, 2013. Our vision – Network Rail [WWW Document]. URL < http://archive.nr.co.uk/aspx/122.aspx > (accessed 8.3.17).
Network Rail-2, 2013. Annual return – Network Rail 2013.
Network Rail-3, 2015. Track geometry – Inspections and minimum actions' NR/L2/TRK/001/mod11.
Network Rail, U., 2014. Asset Management Strategy October 2014.
Niu, G., Yang, B.-S., Pecht, M., 2010. Development of an optimized condition-based maintenance system by data fusion and reliability-centered maintenance. Reliab. Eng. Syst. Saf. 95, 786–796. http://dx.doi.org/10.1016/j.ress.2010.02.016.
Nunez, A., Hendriks, J., Li, Z., De Schutter, B., Dollevoet, R., 2014. Facilitating maintenance decisions on the Dutch railways using big data: The ABA case study. In: 2014 IEEE International Conference on Big Data (Big Data). IEEE, pp. 48–53. http://dx.doi.org/10.1109/BigData.2014.7004431.
Nunez, S.G., Attoh-Okine, N., 2014. Metaheuristics in big data: an approach to railway engineering. In: 2014 IEEE International Conference on Big Data (Big Data). IEEE, pp. 42–47. http://dx.doi.org/10.1109/BigData.2014.7004430.
Nyström, B., Söderholm, P., 2010. Selection of maintenance actions using the analytic hierarchy process (AHP): decision-making in railway infrastructure. Struct. Infrastruct. Eng. 6, 467–479. http://dx.doi.org/10.1080/15732470801990209.
ORR, 2016. An overview of the rail industry in Great Britain [WWW Document]. URL < http://www.orr.gov.uk/_data/assets/pdf_file/0008/16379/an-overview-of-the-rail-industry.pdf > (accessed 8.3.17).
ORR, 2014. Quarters 1-2 of Year 1 of CP5 1 April to 11 October 2014.
ORR-1, 2012. Periodic Review 2013 Consultation on Schedules 4 and 8 Possessions and Performance Regimes – November 2012.
Ove Arup and Partners Ltd, 2013. Part A Reporter Mandate AO/030: PR13 Maintenance & Renewals Review: Policy and WLCC Model. London.
Patra, A.P., Söderholm, P., Kumar, U., 2009. Uncertainty estimation in railway track life-cycle cost: a case study from Swedish National Rail Administration. Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit 223, 285–293. http://dx.doi.org/10.1243/09544097JRRT235.
Popovic, R., Lazarevic, L., Vukandinovic, V., Tepic, G., 2013. Rail inspection of RCF defects. Metalurgija 52, 537–540.
Provost, M.J., 2010. Bombardier *Orbita:* railway asset management for the 21st Century. Saf. Reliab. 30, 46–56. http://dx.doi.org/10.1080/09617353.2010.11690903.
Raheja, D., Llinas, J., Nagi, R., Romanowski, C., 2006. Data fusion/data mining-based architecture for condition-based maintenance. Int. J. Prod. Res. 44, 2869–2887. http://dx.doi.org/10.1080/00207540600654509.
Rahman, S.S., Easton, J.M., Roberts, C., 2015. Mining Open and Crowdsourced Data to Improve Situational Awareness for Railway. In: Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015 – ASONAM '15. ACM Press, New York, New York, USA, pp. 1240–1243. doi:10.1145/2808797.2809369.
Sánchez, P.J., 2007. Fundamentals of simulation modeling. In: Proceedings – Winter Simulation Conference. pp. 54–62. doi:10.1109/WSC.2007.4419588.
Santos, R., Teixeira, P.F., Antunes, A.P., 2015. Planning and scheduling efficient heavy rail track maintenance through a Decision Rules Model. Res. Transp. Econ. 54, 20–32. http://dx.doi.org/10.1016/j.retrec.2015.10.022.
Scholz, J.B., Gossink, D.E., 2012. A resource management blueprint for fusion and command and control. In: 15th International Conference on Information Fusion, FUSION 2012. IEEE, Singapore, pp. 1005–1011.
Selby, R.W., 2009. Analytics-driven dashboards enable leading indicators for requirements and designs of large-scale systems. IEEE Softw. 26, 41–49. http://dx.doi.org/10.1109/MS.2009.4.
Sheeran, R., Jenner, A., Kay, M., 2015. A Short Guide to Network Rail Network Operations Infrastructure Projects Long-term planning for the future. London.
Sikorska, J.Z., Hodkiewicz, M., Ma, L., 2011. Prognostic Modelling Options for Remaining Useful Life Estimation by Industry, vol. 25, pp. 1803–1836. doi:10.1016/j.ymssp.2010.11.018.
Sinha, A., Chen, H., Danu, D.G., Kirubarajan, T., Farooq, M., 2008. Estimation and decision fusion: a survey. Neurocomputing 71, 2650–2656. http://dx.doi.org/10.1016/j.neucom.2007.06.016.
Steinberg, A.N., Bowman, C.L., White, F.E., 1999. Revisions to the JDL data fusion model. In: Dasarathy, B.V. (Ed.), Proceedings of SPIE. International Society for Optics and Photonics, pp. 430–441. doi:10.1117/12.341367.
Su, Z., Jamshidi, A., Núñez, A., Baldi, S., De Schutter, B., 2017. Multi-level condition-based maintenance planning for railway infrastructures – a scenario-based chance-constrained approach. Transp. Res. Part C Emerg. Technol. 84, 92–123. http://dx.doi.org/10.1016/J.TRC.2017.08.018.
Thaduri, A., Galar, D., Kumar, U., 2015. Railway assets: A potential domain for big data analytics. Procedia Computer Science. 457–467. http://dx.doi.org/10.1016/j.procs.2015.07.323.
TSLG, 2012. THE RAIL TECHNICAL STRATEGY 2012.
Turner, C., Thoppur Ravi, P., Tiwari, A., Starr, A., Blacktop, K., 2017. A software architecture for autonomous maintenance scheduling: scenarios for UK and European rail. Int. J. Transp. Dev. Integr. 1, 371–381. http://dx.doi.org/10.2495/TDI-V1-N3-371-381.
Turner, C.J., Tiwari, A., Starr, A., Durazo-Cardenas, I., Blacktop, K., 2015. An intelligent framework and prototype for autonomous maintenance planning in the rail industry. In: VEHITS 2015 – Proceedings of the 1st International Conference on Vehicle Technology and Intelligent Transport Systems.
Waltz, E., Hall, D., 2009. Requirements derivation for data fusion systems. In: Handbook of Multisensor Data Fusion – Theory and Practice. CRC Press, pp. 849.
Wang, W., 2008. Delay Time Modelling. In: Springer Series in Reliability Engineering. Springer London, London, pp. 345–370. http://dx.doi.org/10.1007/978-1-84800-011-7_14.
Yang, L., 2012. Design Principles of Integrated Information Platform for Emergency Responses: The Case of 2008 Beijing Olympic Design Principles of Integrated Information Platform for Emergency Responses: The Case of 2008 Beijing Olympic Games. doi:10.1287/isre.1110.0387.
Zhang, T., Andrews, J., Wang, R., 2013. Optimal scheduling of track maintenance on a railway network. Qual. Reliab. Eng. Int. 29, 285–297. http://dx.doi.org/10.1002/qre.1381.
Zhao, J., Chan, A.H.C., Burrow, M.P.N., 2009. A genetic-algorithm-based approach for scheduling the renewal of railway track components. Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit 223, 533–541. http://dx.doi.org/10.1243/09544097JRRT273.



https://dspace.org/

An autonomous system for maintenance scheduling data-rich complex infrastructure: Fusing the railways' condition, planning and cost

Durazo-Cardenas, Isidro; Starr, Andrew; Turner, Christopher J.; Tiwari, Ashutosh; Kirkwood, Leigh ...

2018-02-22
Elsevier
http://dx.doi.org/10.1016/j.trc.2018.02.010
Durazo-Cardenas I, Starr A, Turner CJ, et al., An autonomous system for maintenance scheduling data-rich complex infrastructure: Fusing the railways' condition, planning and cost. Transportation Research Part C: Emerging Technologies, Volume 89, April 2018, pp. 234-253

Downloaded from a DSpace repository. https://dspace.org/