Optimal Location of Railroad Wayside Defect Detection Installations

Yanfeng Ouyang
Assistant Professor
Department of Civil & Environmental Engineering
University of Illinois at Urbana-Champaign
Tel.: +1 217 333 9858; Fax: +1 217 333 1924
E-mail: yfouyang@uiuc.edu

Xiaopeng Li
Department of Civil & Environmental Engineering
University of Illinois at Urbana-Champaign
E-mail: yfouyang@uiuc.edu

Yung-Cheng Lai
Railroad Engineering Program
Department of Civil & Environmental Engineering
University of Illinois at Urbana-Champaign
E-mail: lai3@uiuc.edu

Christopher P.L. Barkan
Associate Professor
George Krambles Faculty Fellow
Railroad Engineering Program
Department of Civil & Environmental Engineering
University of Illinois at Urbana-Champaign
E-mail: cbarkan@uiuc.edu

Athaphon Kawprasert
Railroad Engineering Program
Department of Civil & Environmental Engineering
University of Illinois at Urbana-Champaign
E-mail: akawpra2@uiuc.edu

(September 4, 2008)

Word Count: 6,851 Words
Optimal Locations of Railroad Wayside Defect Detection Installations

Yanfeng Ouyang, Xiaopeng Li, Yung-Cheng Lai, Christopher P.L. Barkan, and Athaphon Kawprasert
Department of Civil & Environmental Engineering
University of Illinois at Urbana-Champaign

(September 4, 2008)

Abstract

Railroads have been using wayside inspection technologies for many years to improve the safety and efficiency of operations. There has been a recent proliferation of new, more sophisticated, but also more costly inspection systems that are capable of detecting a wide range of subtle defects before a problem has actually occurred. This new class of wayside detection systems has been termed "predictive" as compared to the older "reactive" detection technologies. Because of the different role these predictive detectors play compared to conventional detectors, fewer of them are needed and the deployment strategy also differs. The higher cost of these new technologies means that it is particularly important for railroads to maximize the benefit they derive from the investment. This paper develops a network optimization model that selects cost-effective installation sites for wayside defect detection systems over a railroad network. The objective is to maximize the total inspection benefits possible under any given investment budget. The paper also presents case studies with empirical data to illustrate the technique. The computational results show that the problem can be solved efficiently and the model has the capability of being applied to full-scale railroad networks at regional or national levels. There are a variety of ways that railroads can use this model to help them more efficiently invest in wayside inspection technology so as to maximize the safety and economic benefits of these technologies.
1. Introduction

Railroads use various different types of technology installed adjacent to railroad tracks to monitor the health and performance of passing rolling stock. The design and technology used in these "wayside detectors" varies widely depending on what part or characteristic of the railcar's performance or condition they inspect. The technological maturity of these systems also varies. For example, dragging equipment detectors were first developed in the 1930s (Post 1936, 1937, Burpee, 1945), and wayside hot bearing (hot box) detectors were developed in the 1940s and 1950s (Austin 1949, Gallagher and Pelino 1959), whereas others are still in the testing or early deployment stage (Blevins et al 2003, Resor and Zarembski, 2004, Barke and Chiu 2005, Lundgren and Killian 2005, Lagnebäck 2007).

Several factors affect the strategy regarding the number and location of various system installations. These include the problem or aspect of the car's condition the particular technology is designed to detect or monitor, the consequences if the problem is not detected, particularly the time between when a problem becomes detectable, and when action needs to be taken in response. In this context, Lagnebäck (2007) distinguishes between "reactive" and "predictive" detectors. He describes reactive systems as detecting faults on vehicles that may be hard to predict with the particular technology or have short latency between fault and failure. Detection of these faults protects the equipment from further damage and may prevent an accident. Dragging equipment and hot bearing detectors are widely used examples of reactive technologies. Predictive systems on the other hand are capable of measuring, recording and trending the performance of the vehicles and specific components. The information collected can be used to analyze the condition of equipment to predict possible failures and faults that may occur some time in the future thereby making it easier to plan and schedule maintenance activities and use and repair equipment more efficiently (Blevins et al 2003, Tournay and Cummings 2005, Tournay et al 2006, Lagnebäck 2007). The cost to install and maintain the technology and the benefits it provides are also important factors affecting deployment (Resor and Zarembski, 2004, Barke and Chiu 2005).
An example of a widely deployed reactive technology is the hot bearing or "hotbox" detector that uses infrared technology to detect overheated bearings (Gallagher and Pelino 1959). Bearings can go from a nearly undetectable problem to complete failure in a relatively short distance, i.e. less than 20 miles. The consequences of such a failure can be catastrophic and lead to derailments (English, 1996). The strategy with regard to deployment of this equipment is to place them along lines where the traffic density and operational practices justify them. They are typically spaced closely enough that a failing bearing can be detected in time to take remedial action before the bearing fails (CN 2008). Technologies such as this have distinct safety benefits and are widely used by railroads. However, dealing with the problems they identify is detrimental to railroad service quality and reliability because there is no advance warning that something must be done until fairly severe remedial action is required, such as setting out the car and/or severely restricting train speed.

Consequently, railroads are interested in technologies that provide more advanced notice of an impending failure so that the information and corrective action can be managed more efficiently. Such systems have become increasingly feasible and cost-effective over the past two decades with the rapid advancement in sensor, information-processing and communications technology (Steets and Tse 1998, Lundgren and Killian 2005). Sophisticated systems have been developed to monitor more subtle indicators of railcar component health. An example is the expanding installation of Truck Performance Detectors (TPD) that provide various diagnostic information regarding the health of the truck and related components (Wolf and Peterson 1998, Tournay et al 2006). TPDs use an array of strain gauges mounted on the rails to measure vertical and lateral loads on the track structure as railcars pass through a particular type of track configuration. They integrate the temporal and spatial pattern of these loads so that certain patterns indicative of the condition of various elements of the wheels and truck (bogie) assembly can be assessed. Another example is the Association of American Railroads' (AAR) Fully Automated Car Train Inspection System (FACTIS) system, which is currently undergoing field testing by AAR and railroads (Lundgren 2007). Other advanced inspection technologies are also in the R&D stage, such as machine vision systems to detect the loading efficiency of intermodal trains (Lai et al., 2007) or to inspect safety appliances on railcars (Edwards et al. 2006).
In general, these technologies are not detecting problems for which failure is imminent. Instead, they are looking for irregularities in performance of a component that suggest some type of maintenance may be needed in the future. Related to the development of these technologies is the ability to integrate information from many detector sites as they inspect the same railcar at various points as it travels around the rail network. Systems such as InteRRIS consolidate and analyze this type of data and identify trends that may suggest a developing problem (Irani et al 2003a, 2003b).

The greater sophistication and capability of these systems comes at a considerably higher cost. Each installation may be on the order of three quarters of a million dollars, compared to $100,000 (or less) for more conventional systems such as hot bearing or dragging equipment detectors. Furthermore, the underlying philosophy of these advanced technologies is fundamentally different, (Lagnebäck 2007). Since they are not detecting imminent failure, deploying large numbers of them throughout the network is not as important as is maximizing the likelihood that the largest possible number of different cars are "seen" on a regular, but not necessarily frequent basis. Consequently, fewer of these detectors are needed compared to conventional detector systems and the characteristics of the deployment locations also may differ. Railroads are among the most capital intensive of all industries and thus they are acutely sensitive to the cost of equipment such as this. As the demand for deployment of these technologies increases, the importance of knowing how to locate them to maximize the safety benefits they provide becomes ever more important. There needs to be more consideration of where they should be placed in order to maximize the utility they provide to railroads in the most economically efficient manner possible.

There are a variety of criteria railroads can and do use to select locations to install detectors, but we are unaware of any comprehensive, formal quantitative framework that allows them to determine the optimal placement of predictive wayside inspection facilities. The field of optimization modeling and operations research offers the tools to solve this problem. Furthermore, the increasing computational power now widely available means that complex algorithms using very large amounts of data can be executed in a feasible time frame, thereby providing railroads with better information on how to optimize their installation strategies for new inspection technologies.
In this paper we develop a formal quantitative approach to address the problem of optimal placement of wayside rolling stock inspection facilities and present two potential solution methods. The model is designed to be flexible regarding the types of questions it can address, the weight assigned to different criteria that affect the outcome, and the type of data available to use the model.

Because North American railroads are made up of a complex system of intersecting lines and routes, understanding how to best locate wayside inspection systems is fundamentally a network-level problem. Furthermore, the approximately 1.3 million freight cars in the North American fleet make over 31 million trips over portions of this network each year (AAR 2006). Although certain cars such as those in unit or intermodal train service tend to travel together over fairly confined routes, a substantial number of cars involved in carload freight service travel much more widely across the network. Ensuring that all of these cars receive inspections on a regular basis is challenging. Development of an optimization approach for placement of wayside inspection facilities that takes all cars and routings into account is non-trivial and computationally intensive. The methodology we have developed is adapted from several well-known network analysis and optimization techniques that have proven useful for solving related problems.

As mentioned above, an important factor affecting the objective function is the interval (whether measured in distance or time) between problem detection and failure. If it is a reactive detection system where the interval is short, such as a failing bearing, then the objective is to examine as many cars as possible, as frequently as possible under the conditions when detection of the fault is most likely. The solution to the detector placement problem under these circumstances is relatively simple. They should be spaced along the busiest sections of line where normal operating conditions are most likely to result in failures. In the case of hot bearing detectors on high-speed mainlines, they generally need to be placed closely enough to detect failure before they occur and are often spaced approximately every 20 miles for hundreds or even thousands of miles along the line. A simple analysis of system-wide traffic volumes and operating speeds will provide much of the information needed regarding their placement.

By contrast, with many of the new predictive technologies, it may be adequate to inspect a particular car only infrequently i.e. every six months or more. This fundamentally
changes the objective function of the optimization model, especially if installations are costly. Instead of wanting to maximize the overall number of inspections, the objective is to maximize the number of different cars inspected. If on a given route most of the traffic traverses its entire length, then multiple installations along that route provide almost no additional value. Instead, inspection facilities should be placed at other locations along other routes where the traffic volume and differential routing patterns will result in the maximum number of different cars being inspected. In short, the objective should be to maximize the number of cars inspected with the minimum number of installation sites.

This article presents a network design model that optimizes the installation placement of these wayside technologies so as to maximize the efficiency of their use. The exposition of the paper is as follows. Section 2 introduces the mathematical formulation and a solution algorithm based on Lagrangian Relaxation. Section 3 presents numerical studies with hypothetical and empirical data. Section 4 provides conclusions and future research.

2. Model Formulation

To systematically select optimal locations, we formulate the problem as follows. Consider the set of all candidate installation locations, each of which can potentially accommodate one and only one facility type at a characteristic cost (e.g., equipment acquisition, supply and maintenance). Overall, the installation costs for all facilities at all chosen locations cannot exceed a total budget.

Railcars having the same origin-destination pairs and travel paths are considered to be part of the same car flow. We have information on whether a flow passes a candidate location. Obviously, a railcar flow will be inspected for potential defects at a location if that flow passes the location and a facility exists at that location.

Suppose too that each railcar in flow will receive an economical benefit if inspected correctly. Computation of the benefits can be very flexible. Under different real-world situations, different formulas may be applied without affecting the optimization model and solution procedure.

Due to data confidentiality agreements, we cannot disclose detail of the input data and solution output.
We make decisions on where to install which type of technology via a mathematical program that maximizes the total inspection benefits subject to the budget constraint. The mathematical detail of the model is described in Ouyang et al. (2007).

The model can be used to identify the locations in the network that will yield the maximum benefit from inspection of different railcars, but a number of variations are possible. For example, a set of optimal solutions can be developed for railroads to guide their installation decisions under various different budget constraints. A related dual problem can find the locations with the minimum installation investment possible, while achieving any specified number (or percentage) of cars in the fleet of interest that need to be inspected in a given time period. The model can also be used to optimize the installation sequence over a period of years as budget or logistical constraints dictate, or to determine the optimal location of the next site(s) given the locations of a set of previously installed detectors. Other formulations are also possible based on various different questions of interest or constraints.

3. Case Studies

Analysis of all the railcars operating on the entire U.S. railroad network, or even those on one of the major Class 1 railroads, represents a huge computational effort. There are approximately 1.5 million railcars in service. One well known national rail network model, the Princeton Transportation Network Model (PTNM), has over 50,000 nodes.

3.1. Testing Examples with Hypothetical Data

3.1.1. Busy locations versus optimal locations

An empirical example was used to demonstrate the potential of the proposed model and why simply selecting the busiest locations in the network (i.e., those with the largest traffic volumes) yields suboptimal results. We selected a random sample of network locations and railcar trips for one of the Class I railroads, and flowed them over the rail network. Each year a railcar will typically travel in many trains over multiple routes. Therefore we identify each
railcar as a distinct flow unit. The set of candidate locations includes origin, junction, and termination locations of the railcar flows on the railroad network.

The locations where each railcar originates, passes by, and terminates can be identified using a network flow model, such as the PTNM. In PTNM the ‘NET3’ numbers indicate the individual nodes or locations on the network. In this case, we identified 8,920 individual railcars traveling around the network and 1,820 candidate locations. The network and candidate locations are illustrated in Figure 1(a).

We first consider installing one type of wayside technology to maximize the total number of distinct railcars inspected in a year. In addition, we assume that the budget is sufficient for 10 installations. The proposed algorithm yields optimal solution within 10 CPU minutes on a 2.3 GHz PC. At optimality, 4,951 cars (or 57% of the total railcars) would be detected by the 10 installations. It is interesting that, as discussed above, the top 10 optimal locations are not the 10 busiest nodes. In fact, only four out of the ten busiest locations are part of the optimal solution. If detectors were placed at the 10 busiest locations, only 4,160 cars would be detected, which is 19% less than if the optimal set of sites were used. Although in practice, an experienced railroad expert would do better than choosing the 10 busiest locations, this simple example demonstrates the importance of optimizing detector locations at the network level.

![Figure 1. The railroad network and candidate locations for the test problems.](image)
3.2. Empirical Application

We obtained empirical data from a major U.S. railroad for its network (nodes, links) and traffic (railcar shipment schedules) for 30, 60 and 90-day intervals. The standard maximum covering model and LR algorithm were applied to solve for a range of 1 to 20 installation locations that maximize the number unique railcars inspected.

The original data contain more than ten thousand candidate locations in the network, about half a million distinct railcars conducting about two million shipments per month. A stand-alone computer program was developed for this empirical study to find and display the best set of locations that inspect the maximum number of railcar flows. The software can also determine the subset of railcars that are inspected by any given set of locations. For more information about this software, see Li and Ouyang (2007).

For the same set of network locations and railcar flows, computational time increases slightly with the number of installations. However, on a personal computer with a 2.3 GHz CPU, our model can yield near-optimal solutions in about one hour for all computed cases. The objective function values (i.e., the number of inspected distinct cars) are quite close for 30, 60, and 90 days of traffic. The optimality gap can be further reduced by increasing computational time, but the marginal computational effort needed increases as the gap itself gets closer to zero.

The railroad also provided information on its current wayside detector installations. Compared with the existing installations on this railroad’s network, the solution from the proposed model (with the same number of installations) will improve the inspection benefit by a relative amount ranging from 20% to 60%.

4. Discussion: Potential Uses of the Model

The optimization model presented in Section 2 can be adapted to address several different types of practical problems that are of interest to the railroad industry. The following subsections discuss some of these applications, which are closely related from a modeling point of view (i.e., based on a similar underlying mathematical model). However, these
adaptations extend and enhance the model's utility to railroads and thus broaden its potential benefit.

4.1. Optimized Wayside Inspection Installation by Railroads

The model can be used by railroads wishing to optimize their investment in wayside inspection technology so as to maximize the utility they derive. A railroad will typically have a specified annual budget for investment in wayside detection installations. Use of the model described here would allow them to maximize the return on investment in terms of numbers of different cars inspected. Furthermore, installation of existing technologies is likely to take a number of years and as new technologies are developed, they will need to know the optimal deployment locations and schedules for these. The model can be used to identify the optimal sequence of facility installations over a multi-year period given a constrained capital spending budget, thereby helping ensure that the railroad receives maximum benefit in each successive year of the capital plan, and thereafter when all the installations are complete. As traffic patterns change, installation of new facilities should reflect these changes. New traffic data or projections can be used to rerun the model to check and see if existing facilities should be supplemented by new installations, or possibly moved so that the railroad continues to maximize the benefits it derives from them.

The same approach to a railroad optimizing its installation locations can be applied at the national network level. There are four principal railroads that operate principally in the US. Two of them have overlapping networks in the eastern one third of the nation and the other two cover the western two thirds (Figure 1). A substantial percentage of railroad freight shipments traverse more than one of these railroads between their origin and destination points. Consequently, the railcars involved in these shipments often pass wayside inspection sites on more than one railroad. North American railroads and car owners already cooperate so as to maximize the value of wayside inspection data collected from sites throughout the network in the form of the AAR's InteRRIS program (Irani et al 2003a, Hawthorne et al 2005). Data are collected and analyzed for individual railcars from wayside inspection facilities throughout the network and analyzed at a centralized location. Reports are then communicated to various individual railroads and car owners that subscribe to the
service. The consolidation and centralization of the data processing enables trending and potentially earlier warning to railcar owners for data for individual railcars than would otherwise be possible.

In the same manner that an individual railroad might want to optimize the placement of inspection facilities, the railroads collectively may wish to as well. It is almost certainly the case that optimizing at the national level will be more cost-effective overall than each railroad only considering locations on its own property. Consider two high-density mainlines on two different roads that interchange traffic at a common gateway. Optimizing locations individually, they might each locate an inspection facility on their respective side of the interchange point. If they are interchanging much of this traffic, then one installation used by both might be nearly as beneficial. If a single installation provides nearly as much benefit as two, then the optimization model can be used to identify another location where an installation would provide greater marginal benefit to the industry as a whole. Avoiding this kind of duplication frees up capital for installations at other locations or other uses. An interesting question is how different the set of locations might be when optimizing at the national level compared to the collective set of sites if each railroad optimized based on its own traffic and network alone, and if there are substantial potential savings from such a collective approach. This is a literal example of "global" versus "local" optimization.

4.2. Optimized Inspection of Certain Railcar or Cargo Types

A more specialized application of the model's potential use involves particular types of inspection technologies that may be particularly important for certain types of cars or traffic. For example, certain railcar types are particularly susceptible to problems with the ability of their trucks to steer well in curves. These cars can cause derailments in certain curve alignments even though their performance is within AAR specifications. A detector placed ahead of such curves might identify a problem and the train could be stopped. However, although this could prevent a derailment, it would delay the train, and all the cars in it, thereby disrupting service. A better solution would be to ensure that the performance of the components affecting steering are being monitored and repairs or maintenance take place well in advance of any potential to cause a problem at a time and location that could be better
planned and managed. This is part of the underlying philosophy of InteRRIS, but it also applies to the strategic placement of wayside detectors. Since the car types that are prone to these problems are known, along with their typical routing patterns, their identity can be a parameter in the model used to select locations for the particular type of wayside installation that monitors the type of problem they are susceptible to.

A related example is cars transporting cargoes whose on-time delivery is particularly important. They may be transporting high value commodities or time-sensitive hazardous materials where the delay due to repair of a failing component would be particularly costly, troublesome to the customer, or poses risk. Or, it might be a car transporting a hazardous material where an accident has the potential to cause severe damage or injury. In these cases there may be an additional premium for detecting and correcting components in time for planned maintenance to occur before they cause a serious delay or accident. Again, the model could be used to optimize the placement of relevant types of detector taking into account the traffic routing and density patterns of these particular railcars and commodities.

Delays of such shipments, particularly due to accidents will often be affected by problems on other cars. An alternative version of the model might be constructed in which the objective function is to minimize the likelihood of accident occurrence or minimization of delays in general, and then compare the solutions to one developed based on the objective function described in this paper. If they were different, then it might suggest an alternative placement strategy, and if they were similar, it would enhance confidence in the value of the solutions developed by the model using the objective function as described here.

4.3. Assessing the Performance of Inspection Facilities

The model may also have utility in providing railroads with objective performance assessments of their overall inspection activity. Railroads have already deployed wayside inspection equipment at various points around their systems based on a variety of criteria. These are providing useful information and additional installations are underway. However, it is difficult for railroads to know how the benefit they are deriving compares to some objective expectation of the maximum they might be able to achieve from their investment in
this technology. The model can be used to provide an objective, quantitative metric of the potential performance that they can compare to the performance of their current installations. As mentioned above, if they are unsatisfied with the results from such a comparison, they can use the model to guide further, new installations or redeployment of existing equipment in a more efficient manner.

5. Conclusions

Wayside inspection technology is widely used and critically important to railroads to maintain safe and efficient operations. There is rapid development and deployment of new technologies that have greater capabilities than earlier generations of detectors (Barke and Chiu 2005, Lagnebäck, 2007). These new technologies are more capable of identifying performance metrics that are predictive of failure and useful for longer-term trending analyses of individual railcar performance. However these new technologies are considerably more costly and railroads are often capital constrained, so it is important that railroads optimize both the locations and sequencing of these installations so as to maximize the utility they derive.

In this paper we present a network design model to optimize the installation placement of wayside technologies. Numerical examples show that it yields near-optimal solutions in reasonable computational time. Based on the model, a stand-alone computer program was developed that solves this problem for any large-scale railroad network.

In the future, the stand-alone software will be further enhanced so that railroads can use it directly for various purposes. Metaheuristics that have the potential to further improve the solution will be considered and included in the software. On the modeling side, several assumptions are currently made in the paper (e.g., on technology performance probability and independence). Future work shall consider possible relaxation of these assumptions. Also, future railcar flows (e.g., over a longer time horizon) may be unknown and stochastic factors can be incorporated into the model. Overall, the research results should help railroads better allocate limited resources and coordinate with each other to maximize the utility they derive from installations of new, advanced-technology, wayside inspection systems.
Acknowledgment

This research was supported in part by research grants from the Association of American Railroads (AAR) Technology Scanning Program and CSX Transportation Inc. We thank Alain L. Kornhauser and ALK Technologies, Inc. for providing the PTNM software, and Dharma Acharya, Natasha Zuraikat and CSX Transportation Inc. for data support.

References


