The University of Kansas

INFORMATION & TELECOMMUNICATION TECHNOLOGY CENTER
The University of Kansas

Technical Report

Modeling for Analysis and Design of Communications Systems and Networks for Monitoring Cargo in Motion Along Trusted Corridors

Daniel T. Fokum and Victor S. Frost

ITTC-FY2010-TR-41420-25

May 2010

Project Sponsor:
Oak Ridge National Laboratory
Award Number 4000043403

Abstract

Exports from Asia to the United States have increased significantly in recent years, causing congestion at ports on the Pacific coast of the United States. To alleviate this congestion, some groups want to ship goods by rail directly from ports to inland intermodal traffic terminals. However, for such an effort to succeed, shippers must have "visibility" into the rail shipment. In this research we seek to provide visibility into shipments through optimal placement of sensors and network elements. We formally define the notion of visibility and then develop models to identify and locate network elements and containers on trains. Two models have been developed—one for use when all network elements are on the train and the other for use when some are located trackside—to determine sensor placements and network design. The models show that, under reasonable assumptions, sensor deployment reduces the overall system cost; therefore, sensor networks make sense for monitoring cargo. These models also enable the study of system trade-offs while achieving the desired level of visibility into cargo shipments.

CONTENTS

List of Figures

List of Tables

I. Introduction

In recent years exports from Asia to the USA have increased significantly, resulting in bottlenecks at certain key ports on the West Coast. Some groups involved in freight transportation have sought to get around the bottlenecks at Pacific Coast ports by using inland ports. To this end, they seek to offload cargo from ships directly onto trains destined for an inland intermodal traffic terminal. Once at the terminal, the freight can be processed by Customs and then distributed within the United States.

For the success of the scheme described above, shippers need to gain "visibility" into freight and cargo movement, particularly in intermodal 'black holes," where freight changes hands across modes and carriers. Visibility will only be possible through real-time integration of sensor data with carrier, shipper, broker, importer, exporter, and forwarder information. Unfortunately, different complex systems are currently used in the container transport chain [1].

To achieve the objective of providing visibility into cargo shipments, trains, railcars, and containers will be equipped with sensors and devices that communicate sensor status, sensor ID, and train location. Breaking a sensor on a container would generate a signal that is communicated to a reader over a network and then to train personnel and/or to an operations center as an alarm message in near real-time. In addition, location information will be sent with the alarm so that the geographic location of the breakage event can be identified. Shipment information from a Trade Data Exchange (TDE) [2] will be included in the alarm so that the the rail car, container, and its contents can be identified. While sensors will present a non-negligible initial cost, their use could allow the sensing system to demonstrate shipment integrity. It is also expected that the use of such systems may help reduce the risk of cargo theft, which the Federal Bureau of Investigation (FBI) estimates costs the U.S. economy \$15–\$30 billion dollars every year [3].

The objective of the research presented here is to develop models to find the "best" system design including communications network and locations for sensors in a rail-based sensor network, as well as to guide the design of future cargo monitoring systems. These models can also be applied to determine system trade-offs when monitoring cargo in motion.

A. Visibility

In this section we provide a formal definition of visibility. Events are recorded in the cargo monitoring sensor network whenever an attempt is made to open, close, or tamper with a seal. The seal also generates other events during normal operations. Informally the integrity of a cargo shipment has state. These states will include locking the container and closing the seal, opening the seal and then the container, and

$$\nu(j,\, t,\, \tau,\, \mathrm{TR}_j,\, P_\epsilon,\, E_j,\, P_\alpha,\, F_j) = \begin{cases} 1 & \text{if } (\Pr(t \leq \tau) \geq \mathrm{TR}_j \text{ AND } P_\epsilon \geq E_j \text{ AND } P_\alpha \leq F_j) \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

tampering with the seal. A critical event is generated whenever the integrity of a shipment changes states. Examples of critical events include messages indicating that seals are opened, closed, or tampered with. Messages are also generated during normal operation of the cargo monitoring system. These messages denote maintenance events and examples include alerts indicating an armed seal or a seal warning of a low battery. The set of messages will also include items such as a seal incorrectly reporting a tamper event, incorrectly being detected as missing, or an incorrect low battery report. This latter group of events will be considered false alarms. Important aspects of visibility include the likelihood of a sensor detecting an event at a container, the time taken by a sensor to notify decision makers of an event, and the likelihood of a false alarm from a sensor. We define visibility as a binary variable that relates the probability of detecting an event at a container with the time taken to report that event to the decision maker and the probability of false alarm for that container. More formally we define a container j, as visible if $\nu(j,\, t,\, \tau,\, \mathrm{TR}_j,\, P_\epsilon,\, E_j,\, P_\alpha,\, F_j) = 1$, where the visibility function is defined as in equation (1) and the parameters of the function are:

- An event can be detected at container j, and made known to the decision maker with a probability P_ϵ, that is greater than or equal to some threshold, E_j.
- The time t, taken to notify the decision maker of an event, must lie within an interval of length τ, with probability greater than or equal to some threshold TR_j, i.e., $\Pr(t \leq \tau) \geq \mathrm{TR}_j$.
- The probability of false alarm at container j, P_α, must be kept less than or equal to some threshold F_j.

The system design determines P_ϵ, P_α, and $\Pr(t < \tau)$. The combination of P_ϵ, P_α and $\Pr(t < \tau)$ can be mapped into a visibility space.

B. Problem Statement

In this section we introduce a generalized problem statement. We may be able to achieve visibility into a cargo shipment on a train by placing sensors, readers, and backhaul communication devices on every container on a train (as is done today for high-value cargo, e.g., hazardous material), or by deploying sensors on every container on the train and closely placing readers with backhaul communications capabilities at the trackside. However, the cost and system trade-offs for such approaches are unknown. As a result this research is aimed at answering the following system design question:

> Given a collection of containers and a collection of end-to-end information subsystems (including sensors, seals, readers, and networks); how do we design an end-to-end system that provides "visibility" (meeting given E_j, F_j, and TR_j constraints for all containers) while minimizing overall system cost?

In our specific rail scenario, our overall design question spawns the following issues:

1) How to map and analyze a "system" description of containers on railcars, train scenario—including train speed and trips per time unit—and associated communications infrastructure into the visibility space? Thus an appropriate system model needs to be developed.
2) How to assign a cost to every position in the visibility space?
3) How to use 1) and 2) to find minimum "cost" systems for providing visibility into a rail shipment?
4) How to use 1) and 2) to determine system trade-offs when seeking visibility into rail shipments?

C. Metrics

Metrics are needed to compare the "goodness" of two or more proposed system designs. In this section we present our metrics, which include:

- **System operational cost.** This metric is computed per trip, and it consists of each sensor's false alarm cost, the cost of deploying the sensors, repeaters and readers, network, and the backhaul communications devices, as well as the cost of reporting events. The costs of missing an event at a given container as well as the costs of a communications failure at a sensor are also components of this metric.
- **Visibility metric.** We declare that container j is visible if the sensor placed on j meets all the system designer-imposed constraints for visibility of the container.

The rest of this paper is laid out as follows: in Section II we present a scheme for identifying containers, sensors, and the locations that they occupy on trains. Section III introduces the parameters and variables in our model for analysis and design of communications systems for cargo monitoring. We describe our models for optimal sensor and communications system assignment in Section IV. Section V presents arguments for validating our models as well as discussing model growth. In Section VI we discuss related work. Concluding remarks are provided in Section VII.

II. A System Description for Identifying and Locating System Elements

Notation is needed to identify sensors and containers in models to analyze and design communications systems and networks for monitoring cargo in motion along trusted corridors. In this section we present one scheme for identifying containers and the locations that they occupy on a train. We focus first on container identification and then turn our attention to indexing container and sensor locations on trains.

A. Identification

The containers that are to be placed on the train typically come in a variety of lengths ranging between 20 ft. (6.1 m) and 53 ft. (16.2 m). We propose sorting the loads by length and numbering each container with a unique index j, that is sufficient for accessing the container's properties, such as its length, weight, value, and other intrinsic attributes of the container that might be necessary to solve the sensor placement problem and identify system trade-offs. For example, suppose we have two 20 ft. (6.1 m) containers, two 40 ft. (12.2 m) containers, and one 45 ft. (13.7 m) container, then j ranges from $1 \ldots 5$, with each container bearing a unique j index. The container types can then be identified by using a function that returns the length of a container when given an integer j, i.e., the container lengths could be stored in a vector L, so that L_j indicates the length of container j.

Every sensor, repeater, and backhaul communications device that is to be placed on the containers is identified with a unique index, i. This index, which starts off with value 1, is sufficient for reading the parameters, e.g, transmission range, associated with each communications element.

B. Location

Each railcar consists of one or more permanently attached units, where a unit is a frame that can support one or more slots [4]. Each unit is uniquely identified by an integer k, where k starts off with value 1. The index $k = 0$ is reserved for the locomotive.[1]

Review of the Association of American Railroads Loading Capabilities Guide [5] indicates that railcars used for intermodal transportation have at most two positions (layers) for carrying intermodal loads. Within each position, slots are available for holding containers. For example, [5] indicates that two 20 ft. containers can be placed in the bottom position and a 40 ft. container is placed over both 20 ft. containers in the top position, as shown in Fig. 1.

[1]The issue of having to deal with several locomotives on a train is not part of this model. If more than one locomotive is present, all the locomotives are treated as one with respect to the goals of this system.

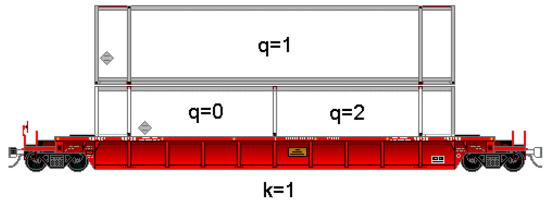


Fig. 1. Unit with Two 20 ft. Containers and One 40 ft. Container

In general, the first available slot in a unit, i.e., the first slot in the bottom position is marked with index $q = 0$. All other slots in the bottom position have even q indices. The first slot in the top position is always indexed with index $q = 1$, while all other slots in the top position will have odd indices.[2]

From above we see that the integer triple (j, q, k) is sufficient for identifying the unit and slot occupied by container j. For instance, using the five container example presented in Section II-A, the integer triple $(1, 0, 2)$ implies that container 1 is found in slot 0 of unit 2. Similarly, $(5, 1, 1)$ implies that container 5 is found in slot 1 of unit 1.

In this section we presented an orthogonal indexing system for containers on a train. This numbering scheme is based on assigning containers with a unique integer j, that is used to identify containers and to retrieve additional container attributes, an index k, that is used to identify railcars, and an integer q, used to identify slots on a railcar.

III. Parameters and Variables

In this section we use the container identification and location scheme from Section II to introduce the parameters and variables in two models for computing the system cost metric for a cargo monitoring system. Parameters will be given to the system designer for a specific placement problem, while the optimization process will assign appropriate values to the variables such that the objective is attained while satisfying any design constraints. To facilitate the presentation of the variables and parameters, throughout this section we use the five container example shown in Fig. 2. The rest of this section is laid out as follows: Section III-A introduces the parameters for the models. For the sake of completeness there is a very brief discussion on container assignment parameters in Section III-A1 while the communications system parameters are presented in Section III-A2. Section III-A3 uses probability distributions to deter-

[2]A review of the Association of American Railways Loading Guide [5] indicates that we will rarely have more than one container in the top position.

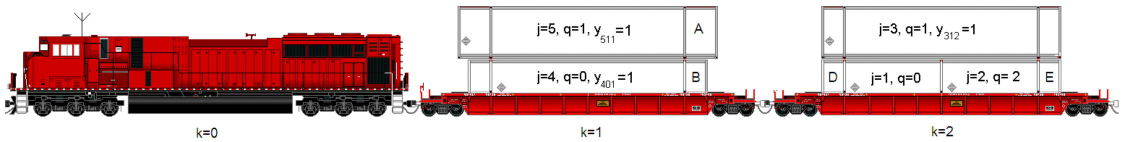


Fig. 2. Two Well-cars with Load Indices Identified

mine the likelihood of timely decision maker notification. The variables for the models are introduced in Section III-B.

Suppose that we have a train with a locomotive and two well-cars as shown in Figure 2. Furthermore, assume that we have two 20 ft. containers, two 40 ft. containers, and one 45 ft. container. Recall that each container (load) is uniquely identified by an integer j, while the railcars are uniquely identified by an index k. Assume that the containers are indexed[3] such that $j = 1$ refers to the most expensive 20 ft. container, while $j = 2$ refers to the least expensive 20 ft. container, $j = 3$ and $j = 4$ refer to the 40 ft. containers. Finally, $j = 5$ is used to denote the 45 ft. container.

A. Parameters

This section introduces the parameters for models. First, we discuss the container assignment parameters, which indicate valid container assignments as well as information on container and railcar attributes. Next, we discuss the communications systems assignment parameters. Finally, we present some distributions to model the time taken to notify decision makers of events on a train.

1) Container Assignment Parameters: This section introduces the container assignment parameters for our model. The length of the k^{th} unit is represented by U_k and the length of the j^{th} container is given by L_j. The binary parameter y_{jqk} indicates a given container's location on the train, and it is defined as:

$$y_{jqk} = \begin{cases} 1 & \text{if } j^{\text{th}} \text{ container is assigned to slot } q \text{ in unit } k, \\ 0 & \text{otherwise} \end{cases}$$

2) Communications Systems Assignment Parameters: In this section we introduce parameters that are necessary to address the system design, including the sensor assignment portion of the container assignment and sensor assignment problem.

[3]Note that we are not required to index the containers by value. The indices could be randomly assigned as long as each number is used exactly once.

Suppose each of the containers has a value v_j, Furthermore, suppose that each time the decision maker receives notification of an event at a container within a time interval τ_j we get a savings σ_j (Note that σ_j can be greater than v_j.). These savings can be viewed as the value of a detected event to the decision maker. We assume that all the repeaters and backhaul communications devices on the train are arranged in a linear topology.

Backhaul communications devices in our system are used to transmit event reports from the train to the decision maker (possibly via an operations center). We use the binary parameter B_{qk} to indicate when a backhaul communications device is placed in slot q of unit k. This parameter is defined as:

$$B_{qk} = \begin{cases} 1 & \begin{array}{c}\text{if backhaul communications device is placed} \\ \text{in slot } q \text{ of unit } k,\end{array} \\ 0 & \text{otherwise} \end{cases}$$

In our system, sensors have a limited transmission range, and they are interrogated by more powerful radios called “repeaters/readers.” The repeaters can communicate with each other over longer distances to get event reports to a backhaul communications device. We use the binary parameter A_{qk} to indicate when a repeater is placed in slot q of unit k. This parameter is defined as:

$$A_{qk} = \begin{cases} 1 & \text{if repeater is placed in slot } q \text{ of unit } k, \\ 0 & \text{otherwise} \end{cases}$$

There are other communications systems assignment parameters used in the optimal placement model. These parameters are shown as follows: Table I presents train-related parameters, sensor and communications equipment-related parameters are listed in Table II, Table III presents message-related parameters, communications system probability parameters are defined in Table IV, and all the cost parameters in the model are listed in Table V.

3) Distributions for Decision Maker Notification: In parallel with the modeling work being described here we have also built a Transportation Security Sensor Network (TSSN) for monitoring rail cargo in motion [6]. In the context of this work we envision that sensors will be placed on intermodal containers that are being shipped by rail, as shown in Fig. 3.

The sensors on the containers will allow the TSSN to detect events at shipping containers and report those that are important to decision makers using commercial networks. Experiments have been carried out with the prototype TSSN and empirical data has been collected [6], [7]. The empirical data will be used to enhance the models described here.

TABLE I

TRAIN-RELATED PARAMETERS

Parameter	Comment
D	Rail trip duration in hours.
d_T	Length of rail journey in kilometers.
$\dot{x}$	Train-speed in kilometers per hour.
t_f	Number of trains passing a given trackside reader per hour.
t_L	Number of trips per locomotive per hour.
ζ	Probability of event occurrence during a trip.

TABLE II

SENSOR AND COMMUNICATIONS EQUIPMENT-RELATED PARAMETERS

Parameter	Comment
LT_A	Useful lifetime of trackside reader in hours.
LT_c	Useful lifetime of cellular communications device in hours.
LT_s	Useful lifetime of satellite communications device in hours.
FP_2	Weight of sensor cost allocated to improving event detection.
FP_3	Weight of sensor cost allocated to improving timely reporting or successful communications in train-mounted and trackside cases, respectively.
FP_4	Weight of sensor cost allocated to reducing false alarms.
FP_5	Weight of sensor cost allocated to improving sensor transmission range.
FP_6	Weight of sensor cost allocated to reducing sensor read time.

TABLE III

MESSAGE-RELATED PARAMETERS

Parameter	Comment
l	Average message length in bytes between sensor and operations center.
λ_i	Message generation rate for sensor i.
RTT_c	Communications round trip time in seconds from train to operations center over the cellular link.
RTT_s	Communications round trip time in seconds from train to operations center over the satellite link.

One of the critical metrics for TSSN performance—and also for visibility—is the time between event occurrence and decision maker notification. Due to the proposed definition for visibility, we need to

TABLE IV

COMMUNICATIONS SYSTEM PROBABILITY PARAMETERS

Parameter	Comment
$\Pr(H)$	Probability of successfully transmitting a message from the train to the operations center over the cellular link.
$\Pr(I)$	Probability of successfully transmitting a message from the train to the operations center over the satellite link.

TABLE V

COST PARAMETERS

Parameter	Comment
C_α	Cost of one false alarm.
C_s	Cost of sending one byte by satellite.
C_c	Cost of sending one byte by cellular.
C_A	Acquisition cost of one reader/repeater.
C_F	Fixed cost of acquiring a sensor.
C_{BC}	Acquisition cost of one backhaul communications device (cellular).
C_{BS}	Acquisition cost of one backhaul communications device (satellite).
C_{HL}	Installation cost of one sensor/seal.
C_{AL}	Installation cost of one reader/repeater.
C_{AD}	Installation cost of one trackside reader.
C_{BD}	Installation cost of one trackside cellular communications device.

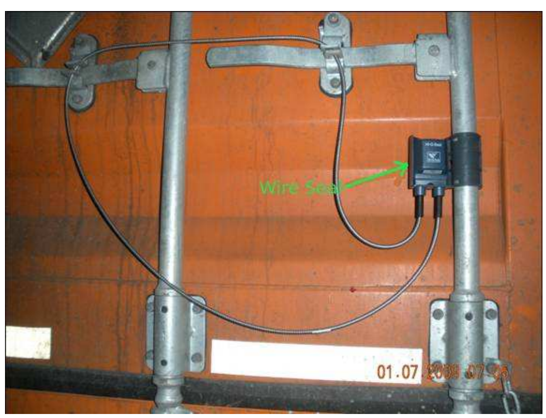


Fig. 3. Container Seal

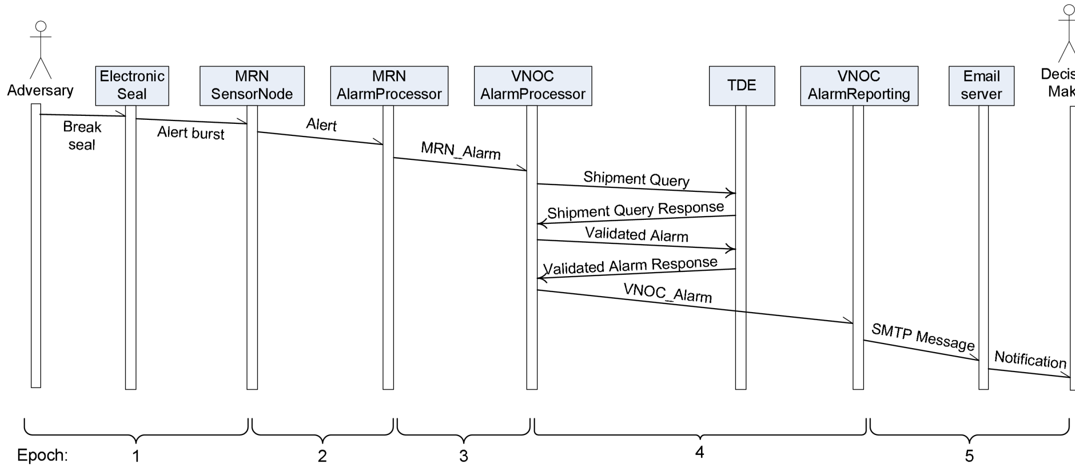


Fig. 4. Sequence Diagram with Messages Involved in Decision Maker Notification

create models that can predict the probability that a decision maker will be notified within a specified time interval. Please refer to Fig. 4 for a sequence diagram showing the messages that are exchanged within the TSSN to notify a decision maker.

In Fig. 4 the electronic seal, Mobile Rail Network (MRN) SensorNode, and MRN AlarmProcessor are on the train. The VNOC AlarmProcessor and Virtual Network Operating Center (VNOC) AlarmReporting services run on a server off the train. Finally, the Trade Data Exchange (TDE) is outside the shipper's network. As shown in Fig. 4 there are five epochs between an event taking place and decision maker notification on a mobile phone.

We would like to generate a distribution that measures the likelihood of timely decision maker notification of an event occurrence. It is reasonable to assume statistical independence of the epochs shown in Fig. 4 because the time taken to break a seal and generate an alert message is independent of the time taken to transfer a message from the MRN to the VNOC. Thus, the probability distribution of the time from event occurrence to decision maker notification is the convolution of the probability distributions for the five epochs listed above.

Based on our experiments we have calculated means and variances for the time taken to transmit a message in each of the epochs listed above. These statistics are summarized in Table VI.

The time epochs shown in Fig. 4 are random. We assume that these random variables can be modeled

TABLE VI

STATISTICS FOR TIME TAKEN IN SECONDS BETWEEN SEAL EVENTS AND DECISION MAKER NOTIFICATION FOR SHORT-HAUL TRIAL AND EMPIRICAL DATA

Epoch	Description	Min.	Max.	Mean	Median	Std. Dev.
1	Event occurrence to alert generation	0.81	8.75	2.70	2.13	1.86
2	Alert generation to MRN AlarmProcessor Service	0.01	0.08	0.02	0.01	0.01
3	One-way delay from MRN AlarmProcessor to VNOC AlarmProcessor	0.45	2.90	1.89	1.94	0.62
4	MRN_Alarm arrival at VNOC to AlarmReporting Service	0.01	3.01	0.17	0.05	0.32
5	Elapsed time from VNOC AlarmReporting Service to mobile phone	5.2	58.7	11.9	9.8	7.4

TABLE VII

ESTIMATED GAMMA DISTRIBUTION PARAMETERS FOR TIME TAKEN BETWEEN SEAL EVENTS AND DECISION MAKER NOTIFICATION

Epoch	Description	$\hat{\alpha}$	$\hat{\theta}$
1	Event occurrence to alert generation	4.01	0.60
2 + 4	Alert generation to MRN AlarmProcessor Service and MRN_Alarm arrival at VNOC to AlarmReporting Service	1.13	0.13
3	One-way delay from MRN AlarmProcessor to VNOC AlarmProcessor	13.95	0.14
5	Elapsed time from VNOC AlarmReporting Service to mobile phone	10.44	1.00

using Gamma[4] probability density functions. The parameters for the distributions are estimated from the collected data and shown in Table VII, where $\hat{\alpha}$ and $\hat{\theta}$ represent the shape and scale parameters respectively. Using the results from [8] we see that there is 99.9 % chance that a decision maker is notified of an event within 4 minutes.

[4]This distribution assumption is based on the positive values and the asymmetric histograms of the observed data. However, this assumption is based on a limited number of samples and additional experiments are needed to validate this claim.

TABLE VIII

TRAIN-MOUNTED DEPLOYMENT VARIABLES

Variable	Comment
α	Probability of false alarm for sensor.
ϵ	Probability of event detection by sensor.
φ	Probability of successful end-to-end communications from sensor to operations center.
C_H	Acquisition cost of one sensor/seal.
Γ_k	Cost of false alarms per railcar.
Δ_k	Cost of missed detection per railcar.
Ξ_k	Cost of reporting an event outside of desired deadline for container visibility.
Λ_k	Cost of transmitting messages generated by all the sensors on a railcar.
Ψ_k	Cost of acquiring and installing sensors on each railcar.
Υ_k	Cost of acquiring and installing repeaters on each railcar.
Ω_k	Cost of acquiring and installing one backhaul communications device on each railcar.

B. Communications Systems Assignment Variables

This section presents the variables used to indicate communications system assignment in our models. These variables are either integers or positive real numbers. Appropriate values will be assigned to these variables such that the best objective function value is attained. First, we present the variable that is common to the trackside and train-mounted cases. Next, we present the other variables that are unique to each case. In general, whenever a variable or parameter is indexed by $q = 0$, and $k = 0$ it is assumed that we will be referring to the locomotive. For example, $A_{00} = 1$ and $B_{00} = 1$ will indicate that a reader and a backhaul communications device, respectively, are located on the locomotive. In our discussion, a "sensor" refers to the combination of sensing and communication devices, e.g., the seal shown in Fig. 3. The binary variable S_{ijqk} indicates when sensor i is assigned to the j^{th} container. This variable is defined as:

$$S_{ijqk} = \begin{cases} 1 & \text{if sensor } i \text{ is attached to } j^{\text{th}} \text{ container in slot } q \text{ of unit } k, \\ 0 & \text{otherwise} \end{cases}$$

1) Train-Mounted Deployment Variables: There are other variables, in addition to S_{ijqk}, used for the case when the sensors and related communications infrastructure are on the train. Table VIII presents these variables and equations (2)–(9) show how the variables are computed.

$$\Gamma_k = C_\alpha \sum_{\forall\, i,j,q} \alpha S_{ijqk} y_{jqk} \tag{2}$$

$$\Delta_k = \zeta \left(\sum_{\forall\, j,q} \sigma_j y_{jqk} - \sum_{\forall\, i,j,q} \epsilon \sigma_j S_{ijqk} y_{jqk} \right) \tag{3}$$

$$\Xi_k = \zeta \left(\sum_{\forall\, j,q} \sigma_j y_{jqk} - \sum_{\forall\, i,j,q} \varphi \sigma_j S_{ijqk} y_{jqk} \right) \tag{4}$$

$$\Lambda_k = D(\Pr(H) C_c + \Pr(I)(1 - \Pr(H)) C_s) l \sum_{\forall\, i,j,q} \lambda_i S_{ijqk} y_{jqk} \tag{5}$$

$$C_H = C_F + \epsilon \mathrm{FP}_2 + \varphi \mathrm{FP}_3 + (1 - \alpha) \mathrm{FP}_4 \tag{6}$$

$$\Psi_k = \sum_{\forall\, i,j,q} (C_H + C_{HL}) S_{ijqk} y_{jqk} \tag{7}$$

$$\Upsilon_k = \sum_{\forall\, q} (C_A + C_{\mathrm{AL}}) A_{qk} \tag{8}$$

$$\Omega_k = \left(\frac{C_{\mathrm{BC}}}{t_L \times \mathrm{LT}_c} + \frac{C_{\mathrm{BS}}}{t_L \times \mathrm{LT}_s} \right) \sum_{\forall\, q} B_{qk} \tag{9}$$

The cost of false alarms per rail car is given by equation (2). This is given by the cost of each false alarm times the sum of probabilities of false alarm for all the sensors that are currently used. Assume that if an event is detected and reported in a timely manner, then there is no loss to the decision maker. In addition, assume that the probability of an event occurring at a container is independent of the probability of a sensor detecting that event or reporting it in a timely manner. Equation (3) computes the cost of a missed detection per railcar, which is given by the probability of event occurrence times the savings that are lost if an event is not detected. Similarly, equation (4) computes the cost of reporting an event outside the required deadline for container visibility. This cost is given by the probability of event occurrence times the savings that are lost if the event is not reported in a timely manner. Equation (5) computes the cost of transmitting messages generated by all the sensors on a railcar. This cost is given by the rail trip duration times the mean cost of transmitting one byte times the sum of message generation rates for all sensors in use. The unit cost of acquiring a sensor for the train-mounted deployment is captured in equation (6). This cost is given by adding up the fixed cost of acquiring each sensor, plus the cost of getting a sensor with specified probabilities of detection, timely reporting, and false alarm. The cost of acquiring and installing the sensors on a railcar is given by substituting equation (6) into (7). Repeater acquisition and installation costs per unit are computed with equation (8). Finally, equation (9) calculates the cost of acquiring and installing a backhaul device on each rail car. We assume that the backhaul

TABLE IX

TRACKSIDE DEPLOYMENT VARIABLES

Variable	Comment
α	Probability of false alarm for sensor.
ϵ	Probability of event detection by sensor.
ρ	Probability of successful communications between trackside reader and sensor.
β	Rate of change of probability of unsuccessful communications with train speed.
η	Probability of unsuccessful communications between trackside reader and sensor when both are stationary.
θ	Real number that specifies the minimum sensor transmission range in meters.
t_{Read}	Real number that states the maximum time in seconds available to read the sensors.
C_H	Acquisition cost of one sensor/seal.
Γ_k	Cost of false alarms per railcar.
Δ_k	Cost of missed detection per railcar.
Ξ_k	Cost of unsuccessful communications between a trackside reader and the sensors on a railcar.
Λ_k	Cost for transmitting messages generated by all the sensors on a rail car.
Ψ_k	Cost of acquiring and installing sensors on each railcar.

devices are reused for several trips, thus we amortize this cost over the expected number of trips in the device's lifetime.

2) Trackside Deployment Variables: The variable S_{ijqk}, which is defined above, is also used when the sensors are mounted on the train and the readers are at the trackside. The rest of the variables for the trackside deployment case are defined in Table IX and equations (2), (3), (7), and (10)–(13) show how the variables are computed.

$$\rho = 1 - (\eta + \dot{x}\beta) \tag{10}$$

$$C_H = C_F + \epsilon \text{FP}_2 + \rho \text{FP}_3 + (1-\alpha)\text{FP}_4 + \theta \text{FP}_5 + \frac{\text{FP}_6}{t_{\text{Read}}} \tag{11}$$

$$\Xi_k = \zeta \left(\sum_{\forall\, j,q} \sigma_j y_{jqk} - \sum_{\forall\, i,j,q} \rho \sigma_j S_{ijqk} y_{jqk} \right) \tag{12}$$

$$\Lambda_k = \left(\frac{d_T}{\dot{x}} \right) C_c l \sum_{\forall\, i,j,q} \lambda_i S_{ijqk} y_{jqk} \tag{13}$$

Suppose that we are given that the probability, ρ, of successful communications from a sensor to a reader varies with train speed according to equation (10). In the trackside case the optimization process will determine appropriate values for α, ϵ, ρ (including η and β), θ, and t_{Read}. The cost of acquiring one sensor for the trackside case is given by equation (11). The cost, Ψ_k, of acquiring and installing sensors on each railcar in the trackside case is given by substituting equation (11) into (7). The values for the Γ_k and Δ_k variables are computed using equations (2) and (3), respectively. As we did above we assume that the likelihood of an event occurring at a container is independent of a sensor detecting that event or the timely notification of that event. In this case we assume that events will get to the operations center in a timely manner if the sensors are read by a trackside reader. The cost of trackside reader failing to read a sensor is given by equation (12). This cost is given by the probability of an event times the cost of a trackside reader failing to read a sensor. Equation (13) computes the cost of transmitting all the messages generated by all the sensors on a railcar. This cost is given by the rail trip duration times the cost of transmitting one message times the message generation rates for all the sensors on a railcar.

IV. Model Descriptions

In this section we present two models for computing the cost metric for a system that uses sensors for cargo monitoring. The models that we develop here are robust enough to handle the following sensor deployment cases:

- A deployment of sensors and a backhaul communications device on the train. This case can be further divided into two subcases:
  - The sensors cannot engage in multihop communications. Instead, they can only communicate with the repeaters or the backhaul communications device. We call this the hierarchical deployment case.
  - The sensors can engage in multihop communications to forward messages to the backhaul communications device. As a result, this case does not contain any dedicated repeaters. We call this the ad hoc deployment case.
- A deployment of sensors to the train, while the readers and backhaul communications devices are at the trackside. This case can also be split into two subcases for when the train speed is fixed and when it is allowed to vary.

The first model, which is presented in Section IV-A, is used when the backhaul communication devices and repeaters are placed on a train. Section IV-B presents the second model, which is used when the train's speed is fixed and backhaul communication devices and readers are placed trackside. Section IV-C shows how the trackside model can be applied in the case where the train speed is allowed to vary. The models discussed in this section are presented using the following general optimization problem formulation:

$$\begin{aligned} &\text{minimize } f_o(x;p) \\ &\text{subject to } f_i(x;p) \leq b_i, \qquad i = 1, \dots, m \end{aligned}$$

The objective function, $f_o(x;p)$, will be the system cost metric function, which depends on a vector of variables, x, and a vector of parameters, p. The constraints of the optimization problem are defined by the m f_i equations. When necessary we provide comments relevant to the equations inline.

In Section IV-D we show how the hierarchical sensor deployment can be mapped to the ad hoc sensor deployment case. Our analysis in the next four subsections assumes that the containers on the train are already placed in fixed locations on the train. Thus, Section IV-E briefly mentions an optimization-based approach which can be used to place containers on trains.

A. Train-mounted Deployment

In this subsection we present a model to minimize the system cost metric of a cargo monitoring system when the sensors and backhaul communications device are on the train. First, we present the objective function and then we discuss the model's constraints, which define valid container and sensor placements.

1) Objective Function: Equation (14) computes the system cost metric over the duration of a trip:

minimize

$$\sum_k \left(\Gamma_k + \Delta_k + \Xi_k + \Lambda_k + \Psi_k + \Upsilon_k + \Omega_k\right) \tag{14}$$

The objective function sums the cost of false alarms over a rail journey, cost of missing a detection at a given container, the cost of a sensor failing to communicate in a timely manner, the cost of communications across a rail journey, the material and installation costs of sensors and repeaters, respectively. Finally, the last term in the sum computes the material and installation cost of the backhaul communications device.

2) Constraints: The following constraints must be valid for any given optimal deployment of sensors to containers on a train.

subject to

$$\sum_{\forall j,q,k} S_{ijqk} \leq 1 \quad \forall\, i \tag{15}$$

$$\sum_{\forall i,q,k} S_{ijqk} \leq 1 \quad \forall\, j \tag{16}$$

Certain attributes (for example, transmission range, detection probability, and false alarm rate) of the sensors, repeaters, and backhaul communications devices are unique to the network elements. Thus, if a given sensor, for example, is placed on a certain container that same sensor cannot be used on another container. Equation (15) ensures that each sensor cannot be simultaneously assigned to more than one container, while equation (16) ensures that each container has no more than one sensor.

$$\varphi = \Pr(t \leq \tau) \tag{17}$$

$$\sum_{\forall\, i,q,k} \varphi S_{ijqk} y_{jqk} \geq \mathrm{TR}_j \qquad \forall\, j \tag{18}$$

In equation (17) we use the probability distribution defined in Section III-A3 to look up the probability of timely notification. Equation (18) enforces one of the visibility requirements for container j. In (18) we require that t, the time taken by a sensor to notify a decision maker of an event, must lie within an interval τ, with probability exceeding some threshold TR_j.

$$\sum_{\forall\, i,q,k} \epsilon S_{ijqk} y_{jqk} \geq E_j \qquad \forall\, j \tag{19}$$

Equation (19) requires that events are detected at container j with a probability ϵ, that exceeds some threshold E_j.

$$\sum_{\forall\, i,q,k} \alpha S_{ijqk} y_{jqk} \leq F_j \qquad \forall\, j \tag{20}$$

Equation (20) enforces the third component of the visibility requirement. In (20) we require that the probability of false alarm at container j, α must be kept lower than some threshold F_j. Equations (18)–(20) ensure that only solutions in the visibility space are considered.

B. Trackside Deployment with Fixed Train Speeds

In this subsection we present a model to minimize the system cost metric of a cargo monitoring system when the backhaul communications devices and readers are trackside. In this case the train's speed is fixed; however, the probability of successful communications from the sensors to the readers varies with train speed. We intend to study the system trade-offs that exist when monitoring rail-borne cargo. This second model facilitates exploration of the trade-off space by capturing the metrics of a different cargo monitoring methodology, which can be compared with the metrics of the first model. As was done above, the objective function is presented first followed by a discussion of the constraints for this model.

1) Objective Function: Equation (21) computes the system cost metric for a trackside-based freight monitoring system over the duration of a trip:

minimize

$$\sum_{k}\left(\Gamma_k+\Delta_k+\Xi_k+\Lambda_k+\Psi_k\right)+\left(\left(\frac{C_A+C_{\mathrm{AD}}}{t_f\times \mathrm{LT}_A}+\frac{C_{\mathrm{BC}}+C_{\mathrm{BD}}}{t_f\times \mathrm{LT}_c}\right)\times\left\lfloor\frac{d_T}{d_A}\right\rfloor\right) \tag{21}$$

The sum in the objective function captures the cost of false alarms over a rail journey, the savings that are lost when a sensor either fails to detect that an event has occurred at a container, the cost of communications across a rail journey, and the material and installation costs of sensors. Finally, the last term captures the cost of setting up trackside readers along a given route.

2) Constraints: Equation (15) holds in this case because no sensor can be placed simultaneously on more than one container. We also require that each container can have no more than one sensor, thus, equation (16) is also valid in this case. In addition, equations (19) and (20) are visibility requirements, thus they are also applicable in this case. Finally, the following constraints must also apply:

subject to

$$2\theta-\dot{x}t_{\mathrm{Read}}\geq 0 \tag{22}$$

Equation (22) says that the minimum time that a sensor is within range of a trackside reader must be greater than the time taken to read a sensor. This constraint allows the train's speed to be limited such that the trackside reader has enough time to read the sensor.

$$2\theta-\dot{x}_{\mathrm{Max}}t_{\mathrm{Read}}\leq 0 \tag{23}$$

Equation (23) states that sensor view time must be less than or equal to the read time if the train is passing the trackside reader at the maximum speed at which a sensor can be read.

$$d_A\leq 2\dot{x}\rho\frac{(\tau-\mathrm{RTT})}{(2-\rho)} \tag{24}$$

The trackside readers are spaced according to equation (24) so that the expected time for end-to-end communications from any sensor plus the time taken to cover the distance between trackside readers must be less than the message reporting deadline.

C. Trackside Deployment with Variable Train Speeds

In this subsection we present a model to minimize the system cost metric of a cargo monitoring system when the backhaul communications devices and readers are trackside and the train speed can be varied based on sensor parameters. In this case we are optimizing over sensor locations, train speed, and reader separations. This change can be accommodated using equation (21) as the objective function.

Constraints: Equations (15), (16), (19), (20), (22), and (24) also hold in this case for the same reasons advanced in Section IV-B2. On the other hand equation (23) does not hold since the train speed is not fixed.

D. Extending the Sensor Placement Models

The presence of repeaters in any system deployment for cargo monitoring adds one more layer of complexity. In Section IV we claimed that a deployment where the sensors can only communicate with repeaters or a backhaul communications device on the train is related to a deployment in which the sensors can engage in multihop communications to forward messages to the backhaul communications device. In this section we discuss how to map the hierarchical deployment case to an ad hoc deployment. In demonstrating this mapping we make the following assumptions:

- The sensor deployment in the hierarchical case is dense enough to have, in the ad hoc case, a fully connected network of sensors with multihop communications capabilities.
- The visibility constraints are the same in all cases and these constraints determine which containers get sensors.
- The probabilities of detection, timely reporting, and false alarm for the sensors do not change as we go from the hierarchical to the ad hoc deployment case.
- Each case contains the same number of backhaul communications devices.
- The ad hoc deployment case does not contain any repeaters.

Suppose that CM_{Hier} and CM_{AD} represent the cost metrics for the hierarchical and ad hoc deployment cases respectively. Observe that no changes need to be made to the objective function because it simply returns a cost metric when presented with sensor and communications infrastructure locations and their characteristics.

Definitions: Let C_H and C_{HL} represent the acquisition and installation costs for the sensors used in the hierarchical case, while C'_H and C'_{HL} represent the acquisition and installation costs for the sensors used in the ad hoc case. Let J_{Hier} and J_{AD} represent the sets of containers assigned sensors in the hierarchical and ad hoc deployment cases respectively. Let I_{Hier} and I_{AD} represent the sets of communications devices (sensors, repeaters, and backhaul communications) which are assigned in the hierarchical and ad hoc deployment cases, respectively. Furthermore, define S_{Hier} and S_{AD} as the set of sensors in the hierarchical and ad hoc deployment cases. B_{Hier} and B_{AD} and R_{Hier} and R_{AD} are the sets of backhaul communications

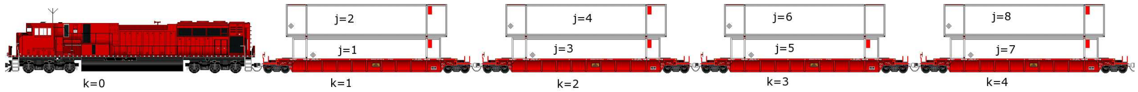


Fig. 5. Example Train With Sensors Assigned

(B_{XX}) and repeaters (R_{XX}) for the hierarchical and ad hoc deployment cases. Then:

$$I_{\text{Hier}} = S_{\text{Hier}} \cup R_{\text{Hier}} \cup B_{\text{Hier}}$$

$$I_{\text{AD}} = S_{\text{AD}} \cup R_{\text{AD}} \cup B_{\text{AD}}$$

We claim that, given the assumptions above, the hierarchical sensor deployment case can be mapped to the ad hoc case. This mapping is based on the assumption that sensors which were previously assigned in the hierarchical deployment case are not moved to other containers in the ad hoc case. This mapping is shown in equation (25).

$$\begin{aligned}\text{CM}_{\text{AD}} = \text{CM}_{\text{Hier}} &+ C_\alpha \sum_{\substack{i \in I_{\text{AD}} \setminus I_{\text{Hier}} \\ j \in J_{\text{AD}} \setminus J_{\text{Hier}} \\ \forall\, q,k}} \alpha S_{ijqk} y_{jqk} - \zeta \sum_{\substack{i \in I_{\text{AD}} \setminus I_{\text{Hier}} \\ j \in J_{\text{AD}} \setminus J_{\text{Hier}} \\ \forall\, q,k}} \sigma_j S_{ijqk} y_{jqk} \Big(\epsilon + \varphi\Big) \\ &+ D(\Pr(H) C_c + \Pr(I)(1 - \Pr(H)) C_s) l \sum_{\substack{i \in I_{\text{AD}} \setminus I_{\text{Hier}} \\ j \in J_{\text{AD}} \setminus J_{\text{Hier}} \\ \forall\, q,k}} \lambda_i S_{ijqk} y_{jqk} \\ &+ \sum_{\substack{i \in I_{\text{AD}} \\ j \in J_{\text{AD}} \\ \forall\, q,k}} (C'_H + C'_{\text{HL}}) S_{ijqk} - \sum_{\substack{i \in I_{\text{Hier}} \\ j \in J_{\text{Hier}} \\ \forall\, q,k}} (C_H + C_{\text{HL}}) S_{ijqk} - \sum_{\forall\, q,k} (C_A + C_{\text{AL}}) A_{qk}\end{aligned} \tag{25}$$

When proving our claim we will use the example train shown in Fig. 5 to illustrate the proof. The train consists of a locomotive and four well cars, with each car bearing two containers. The savings resulting from detecting an event at a container is 8,000 units. The following components are deployed in hierarchical mode for cargo monitoring: a backhaul communications device, a repeater, and seven sensors. The small rectangles on each of the containers in Fig. 5 indicate sensor assignments, while a repeater is on container 6 on railcar 3, and the backhaul communications device is in the locomotive. Finally, assume that we are given the parameter values shown in Table X.

Proof:

1) If we map the hierarchical sensor deployment case to the ad hoc deployment case, then we must get rid of any repeaters in the deployment (Recall that the ad hoc deployment case does not contain any repeaters.). Therefore, $R_{\text{AD}} = \emptyset$, and any repeaters in the hierarchical case are replaced with sensors in the ad hoc case.

TABLE X

PARAMETERS USED IN VALIDATING MODELS

Parameter	Value	Comments
D	20	Rail trip duration in hours.
ζ	0.2	Probability of event occurrence during trip.
F_j	3×10^{-3}	Visibility requirement for probability of false alarm at a container.
E_j	0.85	Visibility requirement for probability of detection at a container.
TR_j	0.85	Visibility requirement for making a timely event report to decision makers.
α	1×10^{-3}	Probability of false alarm for each sensor.
ϵ	0.90	Probability of detection for each sensor.
φ	0.90	Probability of timely event reporting for each sensor.
l	690	Message length in bytes.
λ_i	9.0×10^{-2}	Message generation rate for a sensor. This results in 90 messages every 1,000 hours.
$\Pr(H)$	0.90	Probability of train being in cellular coverage.
$\Pr(I)$	0.90	Probability of train being in satellite coverage.
C_c	5×10^{-5}	Cost in units of sending one byte over a cellular link.
C_s	2×10^{-4}	Cost in units of sending one byte over a satellite link.
$C_{\text{HL}} + C_H$	46	Cost to acquire and install each sensor in the hierarchical case.
$C'_{\text{HL}} + C'_H$	51	Cost to acquire and install each sensor in the ad hoc case.
$C_A + C_{\text{AL}}$	101	Cost to acquire and install each repeater.
C_α	20000	Cost per false alarm.
	14.6	Amortized cost of backhaul communications device.

Using Fig. 5 as an example we assume, without loss of generality, that sensor 1 is assigned to container 1, sensor 2 is assigned to container 2, etc. Then, in the hierarchical deployment $R_{\text{Hier}} = \{6\}$, i.e., the repeater with id code 6 is assigned, while $R_{\text{AD}} = \emptyset$ in the ad hoc case. In addition, assume that in both the hierarchical and ad hoc cases $B_{\text{Hier}} = B_{\text{AD}} = \{9\}$.

2) Since we assume that the same visibility conditions hold in both cases, then we can conclude that the ad hoc deployment case contains at least as many sensors as the hierarchical case, with equality being achieved if the hierarchical case did not contain any repeaters. This condition is captured below:

$$|S_{\text{Hier}}| + |R_{\text{Hier}}| \leq |S_{\text{AD}}| \tag{26}$$

Referring to Fig. 5, the set of sensors assigned in the hierarchical case is $S_{\text{Hier}} = \{1, 2, 3, 4, 5, 7, 8\}$. The set of sensors assigned in the ad hoc case is $S_{\text{AD}} = \{1, 2, 3, 4, 5, 7, 8, 10\}$. Thus, we see that the claim from equation (26) holds with equality.

3) The set of containers that that has sensors in the ad hoc deployment case, but which was not assigned sensors in the hierarchical case is defined as:

$$J_{\text{AD}} \setminus J_{\text{Hier}} \tag{27}$$

Observe that this set is empty if no additional containers are assigned sensors in the ad hoc deployment case. Similarly the set of communications devices used in the ad hoc deployment case, but not in the hierarchical case is defined as:

$$I_{\text{AD}} \setminus I_{\text{Hier}} \tag{28}$$

As with the containers, this set is empty if no additional communications devices are used in the ad hoc deployment case. Note that, since we assume that both cases contain just one backhaul communications device while the ad hoc deployment case contains no repeaters, then equation (28) simplifies to:

$$S_{\text{AD}} \setminus S_{\text{Hier}} \tag{29}$$

Using Fig. 5 as an example, then $J_{\text{AD}} \setminus J_{\text{Hier}} = \{6\}$ since container 6 is the only container that has a sensor assigned in the ad hoc deployment case, but which did not have a sensor in the hierarchical deployment. Similarly, $S_{\text{AD}} \setminus S_{\text{Hier}} = \{10\}$, since sensor 10 is the new sensor assigned in the ad hoc case.

4) From equations (14) and (21) we observe that false alarm and communications costs increase as additional sensors are added, while the savings lost due to missed detections and late event reports decrease. The cost metric of the ad hoc case is the cost metric of the hierarchical case plus the false alarm costs of any new sensors minus the costs of missed detection and untimely reporting due to the new sensors plus any savings from detecting and reporting an event in the desired notification window. To this sum we add the increase in communications costs for the new sensors as well as the installation and material costs for the new sensors. Finally, we subtract the material and installation costs of the repeaters and sensors that were included in the hierarchical deployment. This mapping is summarized in equation (25).

Returning to the example train shown in Fig. 5 let us assume that we are given the parameter values in Table X and that all the containers on the train have low values. Then, the cost metric for the initial hierarchical deployment is 14,178.4 units while the cost metric for the ad hoc deployment is 6,983.5 units. The following costs can be computed for the additional sensor in the ad hoc deployment: false alarm cost for the additional sensor is 20 units, additional savings in event detection due to the new sensors is 3,600 units, savings resulting from decision maker notification in a timely manner is 3,600 units, the additional communication cost is approximately 0.11 units, cost of acquiring and installing the eight new sensors is 408 units, and the amount gained by not deploying a reader is 101 units. It can be shown that $6983.5 = 14178.4 + 20 - (3600 + 3600)$ $+0.11 + 408 - 322 - 101$, which confirms equation (25).

■

E. Container Placement

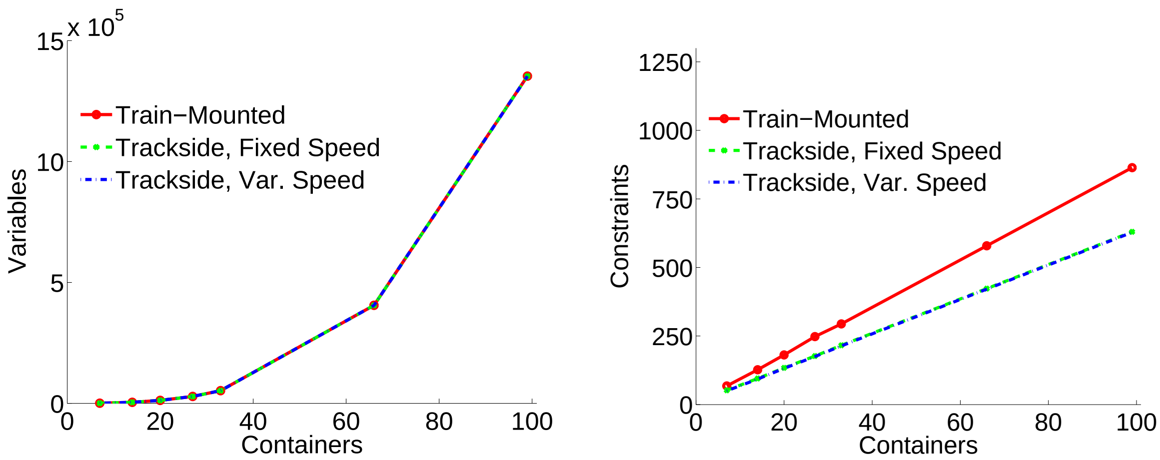
For the purposes of this research we assume that containers have been placed in fixed locations on the train such that the aerodynamic efficiency of the train is maximized. We assume that container placement is done using Lai *et al.*'s [4] method. Please consult [4] for details on the objective function and constraints for this container placement methodology.

V. Model Growth and Validation

In this section we review model validation and the growth of the sensor placement problem with train size. Model validation seeks to determine if a given mathematical abstraction matches a real system. This task is generally hard to accomplish. Kleindorfer *et al.* [9] provides a more complete discussion on validation of models, especially simulation models. By validating our models we can have greater confidence in the optimization results reported by our models.

A. Model Growth and Computational Complexity

In this subsection we examine the growth of our models with different problem inputs. The optimization models described in Section IV have been solved using the Bonmin [10] solver running on the NEOS optimization server [11], [12]. Both models have been run for trains with 7, 14, 20, 27, and 33 containers (this translates to 3, 6, 9, 12, and 15 units respectively). The computational complexity of our models depends on the number of variables and constraints, with the problem becoming more complex with more variables and constraints. The growth in the number of variables and constraints is summarized in Fig. 6. From Fig. 6a it is clear that the train-mounted and trackside models have about the same number of variables. Note that the trackside model with fixed train speeds has additional variables, e.g., sensor transmission range and sensor read time, that are not found in the train mounted model. From Fig. 6b we see that the number of constraints in all three models increases gradually with train size. This growth



(a) Variables

(b) Constraints

Fig. 6. Problem Growth in Number of Variables and Constraints

is partially due to the fact that there is one instance of equation (15) for every sensor and one instance each of equations (16), (19), and (20) for each container. The rapid growth in the number of variables motivates us to consider using heuristics to assign sensors and related communications infrastructure. In our future work we specify a heuristic for assigning sensors to containers in fixed positions on a train. In the rest of this paper and our future work we only consider the train-mounted and the trackside model with fixed train speeds.

B. Model Validation

In this subsection we construct arguments for validating the train-mounted and trackside models by studying trends in the behavior of the optimization models at the boundaries of the visibility space. For the sake of discussion we will use an example train to illustrate our claims. We use the parameter values from Tables X and XI in our discussion.

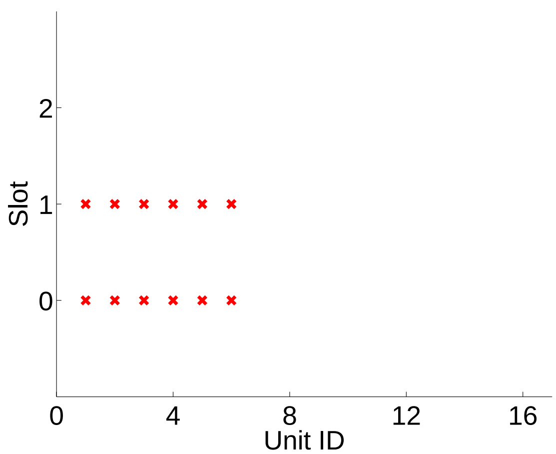
1) Train-Mounted Model: Suppose we have a train with 15 units and 33 containers; where 20 of the containers have a low value, 9 have a medium value, and 4 have a high value. If the train-mounted model achieves an optimal result, it returns the cost metric at the optimal solution as well as the final sensor assignment.

Assume that there are initially enough sensors for each of the containers. Suppose that the visibility conditions on the containers are relaxed such that: $\text{TR}_j = 0.0, E_j = 0.0,$ and $F_j = 1.0$, for some of the containers. In addition assume that there are exactly enough sensors available to satisfy the visibility constraints. Fig. 7a shows the slot and unit locations when only 12 of the 33 containers are visible.

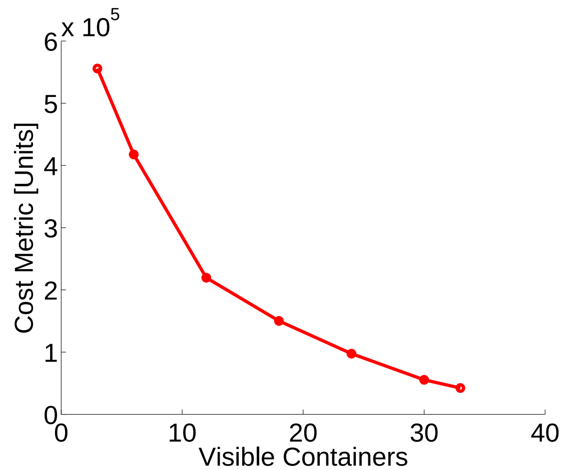
TABLE XI

ADDITIONAL PARAMETERS USED IN VALIDATING MODELS

Parameter	Value	Comments
σ_j	$200,000$	Average savings resulting from event detection at high value container. Reference [13] indicates that in 2006 the average container entering the US had a value of 66,000.
σ_j	$100,000$	Average savings resulting from event detection at medium value container.
σ_j	$20,000$	Average savings resulting from event detection at low value container.



(a) Sensor Locations, where Locomotive = Unit 0, slot 0 is in the bottom level of the railcar, and slot 1 is in the top level.



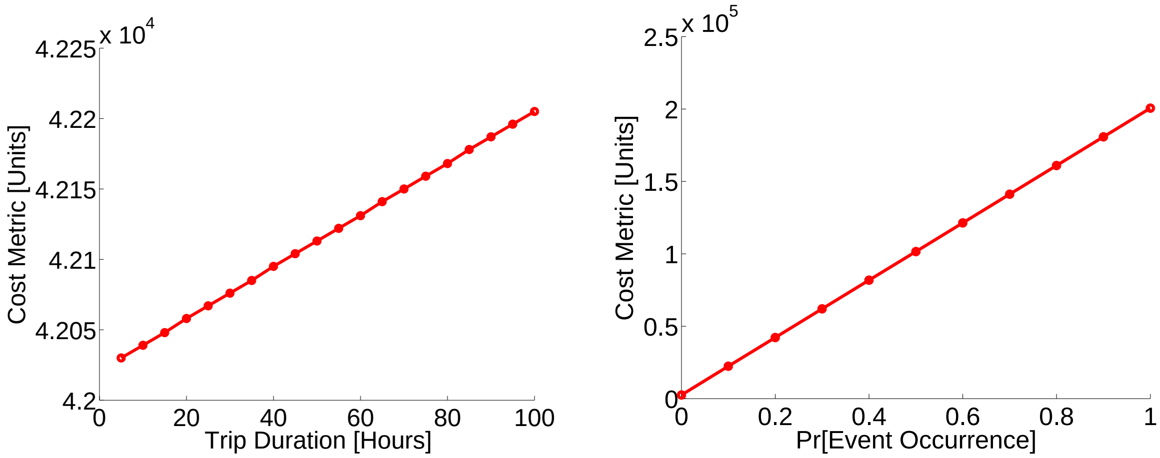
(b) Visibility vs. Cost

Fig. 7. Train-mounted Model: Sensor Locations and Cost Metric Variation with Number of Visible Containers

Fig. 7b shows the relationship between the number of visible containers and the cost metric. As we have fewer sensors the cost metric per trip increases as more containers are not "protected" by any sensors.

As the rail trip duration is increased the cost metric per trip should increase as there is greater opportunity for messages to be transmitted. Fig. 8a shows that as the rail trip duration is increased the system cost metric also increases. Fig. 8b shows the relationship between the probability of event occurrence and the system cost metric. As events become more likely, the system cost metric per trip also increases. Figs. 7b and 8 show that the train-mounted model exhibits correct trends.

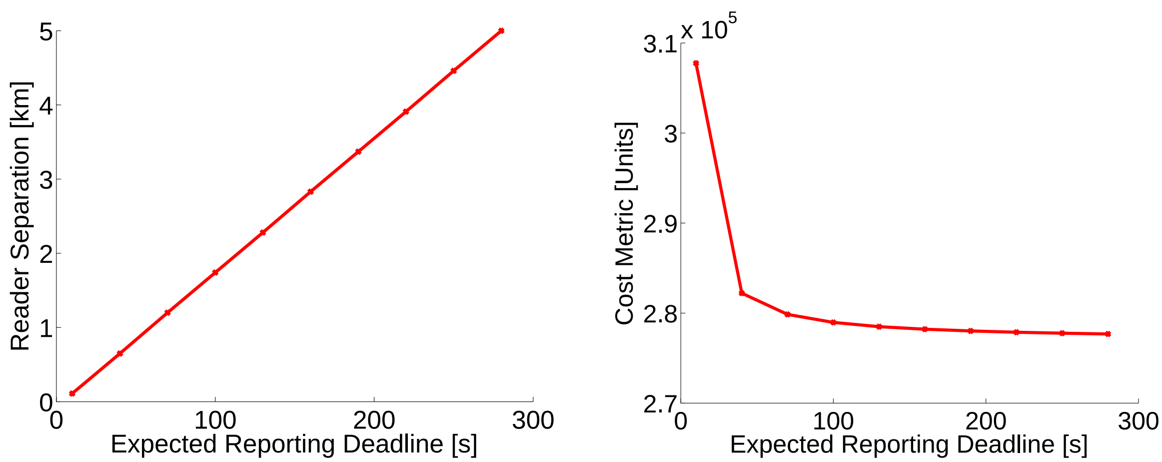
2) Trackside Model with Fixed Speeds: As stated in Sections III-B and IV-B the outputs of the trackside model include the system cost metric, sensor locations, maximum sensor read time, and minimum sensor



(a) Trip Duration vs. Cost

(b) Prob. of Event Occurrence vs. Cost

Fig. 8. Train-mounted Model: Trip Duration and Pr[Event Occurrence] versus Cost Metric



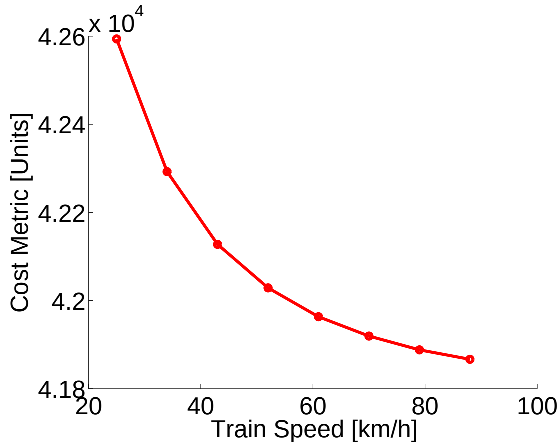
(a) Deadline vs. Reader Separation

(b) Deadline vs. Cost

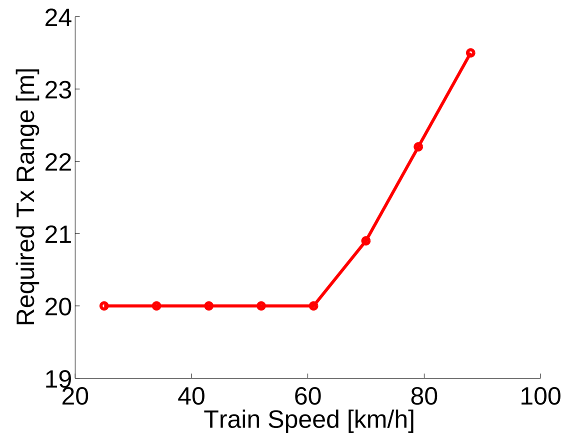
Fig. 9. Trackside Model: Reporting Deadline versus Reader Separation and Cost Metric

transmission range. In addition we also compute reader separation given the reporting deadline and probability of successful communications from a sensor to a trackside reader.

For the trackside model the cost metric for the entire system will increase, as was the case for the train-mounted model, as fewer sensors are available to be used on the train. This is because more and more of the containers are not protected by sensors. Assume that we have the same train configuration mentioned in Section V-B1, with each container being assigned a sensor while the readers are at the trackside.



(a) Train Speed vs. Cost Metric



(b) Train Speed vs. Sensor Transmission Range

Fig. 10. Trackside Model: Train Speed versus Cost Metric and Sensor Transmission Range

Fig. 9 shows the effect of changes in the expected reporting deadline on reader separation and system cost metric when the train speed is fixed at 45 km/h. Fig. 9a shows the relationship between reporting deadline and reader separation. As the reporting deadline is reduced the trackside readers need to be placed closer together. Since more readers are required, the cost metric increases significantly as the reporting deadline is shortened. Fig. 9b shows the change in cost metric with the reporting deadline.

Fig. 10a shows that the cost metric decreases as the train speed is increased. As the train speed is increased the train can cover the distance between its origin and destination in a shorter time implying that the trackside readers can be placed further apart while satisfying the reporting deadlines. Finally, suppose that the system specifications state that each sensor is read in at most 3 s. As the train speed is increased equation (22) shows that the sensor transmission range must increase so that each sensor can be read in the specified interval. Fig. 10b shows that the sensor transmission range increases as expected. This relatively simple example shows that equation (22) correctly captures system operation for the trackside model.

In this section we have shown that our optimization models exhibit correct trends matching a real system. Therefore, we can have confidence in our results.

VI. Related Work

In this section we provide an overview of solution techniques for mixed integer linear and mixed integer nonlinearly constrained problems; two classes of optimization problems that we have encountered in our modeling work.

A. *Mixed Integer Linear Programs*

A mixed integer linear program (MILP) is an optimization problem where the objective function and all of the constraints are linear functions, while some of the variables are integer-constrained [14]. Darby-Dowman and Wilson [15] state that integer program models are generally harder to solve than linear program models of the same size, while Bonami *et al.* [16] state that MILP are $\mathcal{NP}$-Hard problems. Mixed integer linear programs are either solved by branch-and-bound, branch-and-cut, or branch-and-price methods.

When solving integer programs a tree of the entire solution space is created, the root node of the tree is the entire state space, S, while all other nodes represent smaller partitions of the solution space. With branch-and-bound the branching is done by selecting a variable x with a fractional value k and then creating two sub-problems with the additional constraints $x \leq k$ and the other $x \geq k + 1$. This is called the Linear Programming Relaxation (LPR) At a selected node of the tree the integer program LPR is solved. If there is no feasible solution to the problem at that node, the node is eliminated. Otherwise if the solution of the linear programming relaxation is integer feasible and the objective function solution is less than the previous upper bound then the objective function value for this subproblem is set as the new upper bound for the objective function. Branching continues until the best integer feasible solution found is shown to be optimal. With branch-and-cut at each stage in the development of the solution space tree an equation called the cut is added to the set of constraints when carrying out the linear programming relaxation. The cut has the added requirement that it must not exclude any integer solutions at that node or any of its descendants; however, it may exclude integer solutions for preceding nodes. With branch-and-price an auxiliary problem is solved to identify which columns should be added to the linear programming relaxation. The relaxation is optimized and more columns are identified for addition to the LPR [15].

B. *Mixed Integer Nonlinear Programs*

A mixed integer nonlinear program (MINLP) is an optimization problem with some integer-constrained and continuous variables as well as nonlinear constraints and/or objective function. If all the variables are continuous, then we have a nonlinear program. MINLPs are a superset of mixed integer linear programs, where the reduction to MILP takes place when all of the functions in the optimization problem are linear [14].

Mixed Integer nonlinear programs are worse than $\mathcal{NP}$-Hard [14]. However convex MINLPs can be solved using the following techniques: branch-and-bound, extended cutting plane, outer approximation, generalized Benders decomposition, LP/NLP-based branch-and-bound, and branch-and-cut [14], [16]. This section provides an overview of each of these techniques. More detailed explanations of the solution methods are found in [16]. Branch-and-bound for MINLPs is done just as for mixed integer linear programs, except that a nonlinear program is now solved at each node of the tree [14]. The extended cutting plane method constructs a mixed integer linear program relaxation and solves it. If the solution is not feasible, then a cutting plane of the most violated constraint at the optimal solution is added to the relaxation and the problem is re-solved and the process is repeated [14], [16]. Outer approximation (OA) is based on the observation that a MINLP is equivalent to a MILP of finite size. The MILP can be generated by linearizing both the objective and constraint functions. The linearized function is then solved and the integer solution from this step is used as a bound on the optimal value of the NLP. This process is repeated until the upper and lower bounds of the optimal value of the non-linear program are within a specified tolerance [16]. Generalized Benders decomposition is very similar to the outer approximation method except that it has only one continuous variable [14]. LP/NLP-based branch-and-bound is an extension of the OA method. It uses LPR to find an integer solution in a branch-and-bound tree and then solves the nonlinear program to get upper bounds on the solution [14], [16]. Branch-and-cut has been adapted to solving MINLPs [14]. This method is similar to branch-and-bound, but it adds cutting planes at each node of the tree to strengthen the NLP relaxation [14].

VII. Conclusion

This paper presented two models that can be used to find the optimal cost metric for a rail-borne cargo monitoring system. We presented the parameters and variables for our models. The models presented in Section IV are suitable to enable quantitative evaluation of the trade-offs that can be made when monitoring rail-borne cargo. In addition this paper has also shown that we a hierarchical deployment of sensors can be mapped to an ad hoc sensor assignment, given that the sensors assigned in the initial case are not moved to other containers. Finally, this paper has shown that there is a large number of variables involved in the models for sensor assignment. As a result, future work will determine if heuristics can yield near-optimal performance for sensor assignment.

Acknowledgments

This work was supported in part by Oak Ridge National Laboratory (ORNL)—Award Number 4000043403. This material is also partially based upon work supported while V. S. Frost was serving at the National Science Foundation.

REFERENCES

[1] European Conference of Ministers of Transport, *Container Transport Security Across Modes*. Paris, France: Organisation for Economic Co-operation and Development, 2005.

[2] KC SmartPort. (2008, Nov. 10) Trade Data Exchange—Nothing short of a logistics revolution. Digital magazine. [Online]. Available: http://www.joc-digital.com/joc/20081110/?pg=29

[3] Federal Bureau of Investigation. (2006, Jul. 21) Cargo Theft's High Cost. Headline. Federal Bureau of Investigation. [Online]. Available: http://www.fbi.gov/page2/july06/cargo_theft072106.htm

[4] Y.-C. Lai *et al.*, "Optimizing the Aerodynamic Efficiency of Intermodal Freight Trains," *Transportation Research Part E: Logistics and Transportation Review*, vol. 44, no. 5, pp. 820–834, Sep. 2008.

[5] Intermodal Committee, *Loading Capabilities Guide*, Association of American Railroads Std., Jun. 26 2003. [Online]. Available: http://www.aar.org/AARPublications/~/media/AARPublications/FreePubs/AAR%20Loading%20Capabilities%20Guide.ashx

[6] D. T. Fokum *et al.*, "Experiences from a Transportation Security Sensor Network Field Trial," University of Kansas, Lawrence, KS, Tech. Rep. ITTC-FY2009-TR-41420-11, Jun. 2009.

[7] M. Kuehnhausen and V. S. Frost, "Application of the Java Message Service in Mobile Monitoring Environments," University of Kansas, Lawrence, KS, USA, Tech. Rep. ITTC-FY2010-TR-41420-18, Dec. 2009.

[8] S. Nadarajah, "A Review of Results on Sums of Random Variables," *Acta Applicandae Mathematicae: An International Survey Journal on Applying Mathematics and Mathematical Applications*, vol. 103, no. 2, pp. 131–140, Sep. 2008.

[9] G. B. Kleindorfer *et al.*, "Validation in Simulation: Various Positions in the Philosophy of Science," *Manage. Sci.*, vol. 44, no. 8, pp. 1087–1099, Aug. 1998.

[10] P. Bonami. (2010, May) Bonmin. Project wiki. [Online]. Available: https://projects.coin-or.org/Bonmin

[11] (2010, Feb.) NEOS Solvers. Solver listing. Argonne National Labs. Argonne, IL, USA. [Online]. Available: http://neos.mcs.anl.gov/neos/solvers/index.html

[12] J. Czyzyk *et al.*, "The NEOS Server," *Comput. Sci. Eng., IEEE*, vol. 5, no. 3, pp. 68–75, Jul.-Sep. 1998.

[13] S. E. Flynn. (2006, Jan.–Feb.) Port Security Is Still a House of Cards. Article. Council on Foreign Relations. [Online]. Available: http://www.cfr.org/publication/9629/

[14] A. N. Letchford, "Mixed-Integer Non-Linear Programming: A Survey," presented at the 1st LANCS Workshop on Discrete and Non-Linear Optimisation, Southampton, United Kingdom, Feb. 2009.

[15] K. Darby-Dowman and J. M. Wilson, "Developments in Linear and Integer Programming," *The Journal of the Operational Research Society, Special Issue: Applications and Developments in Mathematical Programming*, vol. 53, no. 10, pp. 1065–1071, Oct. 2002. [Online]. Available: http://www.jstor.org/stable/822966

[16] P. Bonami *et al.*, "Algorithms and Software for Convex Mixed Integer Nonlinear Programs," Computer Sciences Department, University of Wisconsin-Madison, Madison, WI, USA, Tech. Rep. 1664, Oct. 2009.