Secure Communications Based Train Control (CBTC) Operations

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

Mark W. Hartong
Master of Science
George Mason University, 2001
Master of Science
United States Naval Postgraduate School, 1985
Bachelor of Science
Iowa State University of Science and Technology, 1980

Director: Dr. Duminda Wijesekera, Professor
Department of Computer Science

Spring Semester 2009
George Mason University
Fairfax, VA

Dedication

This dissertation is dedicated to my beautiful and incredibly intelligent wife Rebecca, who has given me her love and support as well as her understanding and patience during those times when there was no light at the end of anything. She encouraged me and made me stick with it. Without her, there is no way I could possibly have accomplished this.

"Per Aspera ad Astra"

Acknowledgments

"If one does not know to which port one is sailing, no wind is favorable."
-Seneca (the younger)

Writing this has been a lonely and isolating experience, and was not possible without the personal and practical support of numerous people. I am very greatful to the members of my dissertation committee, Dr. Duminda Wijeskera (Chair), Dr. Stephen Nash, Dr. Edgar Sibley, and Dr. Jeremy Allnutt from George Mason University as well as Dr. Rajni Goel from Howard University. Their patience, support, advice, and encouragement made this work possible.

My advisor, Dr. Duminda Wijeskera, has had a most profound influence on my development as a researcher. His ability to identify critical issues and focus on the big picture are qualities to which I aspire. Dr. Wijeskera has never ceased to challenge me, always looking at every last detail and asking the hardest questions. I have found working with him tremendously rewarding.

Dr. Rajni Goel acted as a co-advisor for this effort. Her insight and guidance, along with that of Dr. Wijeskera, was invaluable in the development and acceptance of the peer reviewed work on which this dissertation is based. Without such deep involvement from both of them this project would have been much less interesting and fruitful.

I also wish to extend my appreciation to Mr. Grady Cothen, Deputy Associate Administrator for Safety Standards and Program Development at the Federal Railroad Administration, US Department of Transportation in supporting my academic pursuits.

Of course, despite all the assistance provided by Dr. Wijesekera and others, I alone remain responsible for the content of the following, including any errors or omissions which may unwittingly remain.

Table of Contents

List of Tables

List of Figures

Abstract

SECURE COMMUNICATIONS BASED TRAIN CONTROL (CBTC) OPERATIONS

Mark W. Hartong, PhD

George Mason University, 2009

Dissertation Director: Dr. Duminda Wijesekera

Communications Based Train Control (CBTC) provides positive train separation, over speed protection, and protection for roadway workers. Current system designs do not include trust management systems to provide support for security, rendering CBTC communications vulnerable to malactors. Traditional train control methods and the architecture of CBTC systems are studied to determine specific vulnerabilities of CBTC systems and the associated system security requirements. The security requirement are then used to derive an appropriate trust management system. Existing work on safe cross domain dispatch operations has not considered the impact of these trust management systems on allowable traffic delays and system velocity or related them to train dynamics. A relationship between train dynamics and trust management delay is presented to allow engineering estimates of the practicallity of potential trust management systems to support rail operations while preventing collsions. An algorithm for the safe and secure scheduling of trains through the interchange point between is provided. The algorithm supports positive train separation under a worst-case traffic scenario, allowing for safe and secure scheduling while reducing traffic delays. The approach presented is illustrated by an example, and is independent of the specific security management, CBTC, and dispatch systems.

Chapter 1: INTRODUCTION

Since the late 1980s, Communication Based Train Control (CBTC) Systems, also known as Positive Train Control (PTC) systems [1], for freight and passenger rail service have been under development in the United States [2,3]. These Supervisory Control and Data Acquisition (SCADA) systems have been advertised as offering significant enhancements in safety by ensuring positive train separation, enforcing speed restrictions, and improving roadway worker protection. Industry efforts to deploy these systems [4] have been accelerated by recent regulatory [5] and statutory initiatives [6].

When the railway network is owned and operated by more than one organization, CBTC system security is a critical non-functional requirement that has unique aspects related to cross-company rail operations. Each railroad company is an independent commercial entity that interchanges crews, locomotives, and their consists with other railroads. These personnel and equipment exchanges occur at fixed geographical points where the tracks from one company are interconnected with tracks from another. There are a limited number of these interchange points between any two companies, and they are geographically disperse. Because trains have a single degree of freedom with respect to their operations (that is they can only operate along the tracks), any delay of a train at an interchange point as it crosses from the operating domain of one railroad to the operating domain of another may delay the movement of subsequent trains operating along the same line to the same interchange point. Different delays may be encountered at the interchange point between different railroads if the scheduling and security management systems are not integrated. This dissertation provides a model for secure cross-domain authentication, authorization and scheduling.

1.1 Thesis Statement

It is possible to generate an engineering solution to safely and securely schedule trains at the interchange point of two railroads having:

- *Different public key based security management infrastructures,*
- *Different PTC systems,*
- *Different dispatch and scheduling systems, and*
- *Different communications infrastructures*

in the presence of electronic threats which attack any or all of these systems components to deny or cdisrupt legitimate users of the system, while attempting to minimize delay and optimize system throughput

1.2 Peer Reviewed Work

While extensive publically available bodies of work exist regarding key management, stratgeic (system wide) routing and scheduling, and to a lesser extent traditional signal system design, the same can not be said for PTC systems. Publically available research results regarding PTC systems security is virtually non-existant. Work in this disertation is therefore very unique. The results provides:

- a comparitive view of tradtional and PTC systems security and the resulting PTC system security requirments,
- a trust management system that supports sequre PTC sytem operations,
- an engineering relationship between trust management delays and train dynamics allowing for evaluation of feasibale safe and secure PTC system designs, and

- a railroad cross domain tactical scheduling algorithm.

that represents a starting point for further research refining the results obtained.
The following provides the corpous of publically available, vetted, PTC security specific, documents associated with the preceeding.

1. Mark Hartong, Rajni Goel, and Duminda Wijesekera, *Use Misuse Case Driven Forensic Analysis of Positive Train Control- A Preliminary Study*, Proceedings of the Second IFIP WG 11.9 International conference on Digital Forensics, Jan 29-Feb 01 2006, Orlando FL

   This work addresses identification of attacks and vulnerabilities against CBTC system as forensic issues definable in terms of Use and Misuse cases. It also provides further motivation for the need for a CBTC trust management mechanism.

2. Mark Hartong, Rajni Goel, and Duminda Wijesekera, *Key Management Requirements for Positive Train Control Communications Security*, Proceedings of the 2006 IEEE ASME Joint Rail Conference, 4-6 April 2006 Atlanta Georgia

   Starting from Use and Misuse Cases definitions of CBTC systems threats and their interrelationships, it examines and proposes requirements of the key management component of a comprehensive trust management system.

3. Mark Hartong, Rajni Goel, and Duminda Wijesekera, *Communications Based Positive Train Control Systems Architecture in the USA*, Proceedings of the 63rd IEEE International Vehicle Technology Conference, 7-10 May 2006 Melbourne, Australia,

This work outlines the fundamental architectures of CBTC systems, and provides implementation examples of these architectures. It provides a context for understanding attacks on CBTC systems.

4. Mark Hartong, Rajni Goel, and Duminda Wijesekera, *Communications Security Concerns in Communications Based Train Control*, Proceedings of the 10th International Conference on Computer System Design and Operation in the Railway and Other Transit Systems, 10-12 July 2006, Prague, Czech Republic

   This work identifies general threats to CBTC system architectures, identifies specific vulnerabilities and classes of attacks, and the need for security and trust management systems.

5. Mark Hartong, Rajni Goel, and Duminda Wijesekera *Mapping Misuse Cases to Functional Fault Trees for Positive Train Control Security*, Proceedings of 9th International Conference on Applications of Advanced Technology in Transportation Engineering, 13-16 Aug 2006 Chicago, IL

   This work addresses Use and Misuse Case relationships used in defining CBTC security, and their one possible mechanism for their translation into a formal notation.

6. Mark Hartong and Olga Catilda *Microprocessor Based Signal and Train Control- A New Regulatory Approach* Transportation Research Record, Journal of the Transportation Research Board, Issue Number: 1943 Transportation Research Board, National Academy of Sciences, 2006

   This addresses the regulatory requirements with which a framework any safe, secure cross-domain scheduling, authentication and authorization system must comply.

7. Mark Hartong and Olga Catilda, *Regulatory Risk Evaluation Of Positive Train Control Systems*, Proceedings Of 2007 ASME IEEE Joint Rail Conference & Internal Combustion Engine Spring Technical Conference March 14-16, 2007, Pueblo, CO, USA

   This work introduces the concept of probabilistic threats and varying interpretations as to the associated risk, motivating the need for quantifiable descriptions of the threats. It further explores the regulatory requirements that a frame-work addressing probabilistic threat must comply.

8. Mark Hartong, Rajni Goel, and Duminda Wijesekera, *Rail Infrastructure Security For Positive Train Control Systems*, Proceedings of FIP WG 11.10 International Conference on Critical Infrastructure Protection. Dartmouth College. Hanover, New Hampshire. March 18-21, 2007

   This explores the development of a general security framework for CBTC systems.

9. Mark Hartong, Rajni Goel, Csillia Faraka, and Duminda Wijesekera, *PTC-VANET Interactions to Prevent Highway Rail Intersection Crossing Accidents*, Proceedings of the 65th IEEE International Vehicle Technology Conference, 22 - 25 April 2007, Dublin, Ireland

   This work further explores the application of Use and Misuse cases to capture Vehicular Ad-hoc networks (VANETS) and PTC system requirements, then demonstrates the integration of different system security and performance requirements.

10. Mark Hartong, Rajni Goel, and Duminda Wijesekera, *A Framework For Investigating Railroad Accidents* Proceedings of Third IFIP WG 11.9 International conference on Digital Forensics, January 28 - 31, 2007, Orlando, FL

    Using Use and Misuse Cases relationships to define system behavior this work demonstrates the viability of translations for CBTC frameworks to relations as well as CBTC relations to framework.

11. Mark Hartong, Rajni Goel, and Duminda Wijeskera, *Security and the US Rail Infrastructure*, "International Journal Of Critical Infrastructure Protection, December 2008, Elsevier, Publisher

    This work addresses the general security situation of the US Rail system. It includes both communications based and non-communications based issues.

12. Mark Hartong, Grady Cothen, Olga Catadli, and Terry Tse, *Positive Train Control-Ready to Go?* Mass Transit, Volume XXXIII, No 8 December 2007-January 2008, Cygnus Inc, Pub, Beltsville, MD

    This article outlines the readiness of industry and government for wide scale deployment of CBTC. In addressing the readiness, it highlights the need for a near term effort to address technical issues such system security to implement CBTC.

13. Mark Hartong, Rajni Goel, and Duminda Wijesekera, *Cryptographic Protection And Recovery Of Railroad Event Recorder Data*, Proceedings of the Fourth IFIP WG 11.9 International Conference on Digital Forensics Kyoto University, Kyoto, Japan January 27-30, 2008

    This work provides further study into the application of security mechanism frameworks as well as the integration of security with system operations.

14. Jon Whittle, Duminda Wijeskera, and Mark Hartong, *Executable Misuse Cases for Modeling Security Concerns*, Proceedings of the 2008 International Conference on Software Engineering, Leipzig, Germany, May 10-18. 2008

    This work further explores formalization of Use and Misuse Cases relations, and their linkage to security issues common in trust management schema.

15. Mark Hartong, Rajni Goel, and Duminda Wijesekera *Trust-Based Secure Positive Train Control (PTC)* International Journal of Transportation Security, December, 2008 Springer Verlag, Publisher

    This paper ties together the application of Use and Misuse cases to define a trust management infrastructure and explores aspects of the relationship between a trust management system and locomotive scheduling.

16. Mark Hartong, Rajni Goel, and Duminda Wijesekera, Invited Chapter *Security and Dependability in Train Control Systems*, to appear in Vehicular Communication, Automotive and Beyond", John Wiley, Publisher

    This work provides a definitive reference on multi-modal vehicular communication safety and security. The invited chapter provides a reference on the role of PTC in train operations and their associated security issues.

17. Mark Hartong, Rajni Goel, and Duminda Wijeskera, *Integrating Secure Train Control and Scheduling* to appear, Proceeding of IFIP International Conference on Critical Infrastructure Protection, Dartmouth College, March 22-25 2009, Hannover, NH
    This work provides a discussion of secure, cross domain, train control operations.

1.3 Dissertation Organization

There are eight chapters in this document. Chapter 1 introduces the problem, presents the thesis statement, and identifies previous research on various aspects of the problem. Chapter 2 provides a background on traditional train control and methods of operations. This

provides the necessary context to evaluate the role played by PTC systems in advanced railroad operations. Chapter 3 describes the architecture and implementation of PTC systems. Chapter 4 addresses the security of PTC Systems. It identifies the security threats, the type of attacks, the required security attributes that a PTC system must posses. Chapter 5 develops a trust management system to provide the required security based rooted in Use and Misuse Cases. Chapter 6 presents a model of integrated trust management and scheduling, with the associated required performance requirements. Chapter 7 discusses the safety of the integrated model. Finally Chapter 8 draws conclusions and identifies potential areas for further work.

Chapter 2: TRADITIONAL TRAIN CONTROL AND OPERATIONS

A train is constrained to travel along on a single track, and cannot pass other vehicles operating on the same tracks, except where there are sidings, resulting in a system with a single degree of freedom. In order to control the movement of trains, various methods of operations began to be formalized, starting in the early 1820s when multiple trains began to share the same set of tracks. These methods of operations were designed to improve the operational efficiency of the railroad its and safety through the reduction of collisions, derailments, and the associated deaths. Today's methods of operations for the control of trains can be classified into four basic categories: verbal authority, mandatory directives, signal indications, and signal indications supplemented by cab signals, automatic train control, or automatic train stop systems.

2.1 Verbal Authority and Mandatory Directives

With verbal authority and mandatory directives (i.e. commands from the railroad dispatcher to the train crew), the aspect of wayside signals does not control train operations. Instead, trains are controlled by orders from the train dispatcher, who takes responsibility for knowing what trains are located where, and ensures that no two trains are issued authorization to occupy the same location of track at the same time. The dispatcher usually issues orders, mandatory directives, speed restrictions, as well as the location of any wayside work crews via two-way radio to the locomotive crew. The train crew then is responsible for ensuring that they obey them.

This is the traditional means of controlling operations in the United States, and roughly 40 percent of all tracks in the United States are controlled in this manner. Verbal authority and mandatory directive operations are generally either one of two main types- Track Warrant Control (TWC) and Direct Traffic Control (DTC). TWC and DTC differ in the way the limits of authority are defined. In TWC, verbal instructions are given for the crew to proceed between stations or mileposts (a segment of track known as the authority limit). DTC is similar to TWC, but because the railroad is divided into pre-defined "blocks" DTC is simpler in execution. DTC movement authorities can only be specified in terms of the pre-defined blocks. The dispatcher authorizes a train to proceed in one or more of the blocks and does not have flexibility in the selection of authority boundaries. Although TWC is complex, TWC is replacing DTC due to the greater flexibility in operations it provides.

Both TWC and DTC are alternative verbal authorization systems defined by the General Code of Operating Rules (GCOR) and can be used as the sole means of dispatching and safety. TWC and DTC do not require wayside signals and can be used to supplement Automatic Block Signaling (ABS) to increase flexibility and traffic capacity. ABS was designed primarily for passenger operations, and has largely been replaced by an alternative form of signaling called Reverse Signal Centralized Traffic Control (CTC). When used as a supplemental mode of operation, DTC/TWC serve primarily as protection from errors in the movement authority and do not convey the authority to occupy the main track. That authority remains with the signal system, but the train crew requires a DTC/TWC authorization in addition to the signal to enter the main track.

The General Code of Operating Rules (GCOR) and the Northeast Operating Rules Advisory Committee (NORAC) Rules codify TWC and DTC operations. GCOR is used primarily by railroads in the western United States while railroads in the eastern United States use NORAC. The decision to use the GCOR or NORAC implementation of TWC or DTC is made by individual railroads based on what is most efficient for their operations. TWC is

used by most Western railroads and the Norfolk Southern railroad, while the Kansas City Southern, and the Union Pacific use DTC. CSX also uses DTC on portions of their routes. There are also variations of TWC called Form D Control System (DCS), which are used by Northeastern railroads that have adopted the NORAC Rule Book.

2.2 Control by Signal Indication

Train operations under signal indications makes up the remainder of the train control operations in the US. Track circuit based signal systems were first installed in the US in 1872, and in 1927 were centrally controlled in the first CTC system. CTC is not a separate control system, it uses block signal system and interlocking to control train movements (although radio communications between the dispatcher and train crews are available).

CTC, sometimes called Traffic Control System (TCS), has remained basically unchanged since the 1930s. In CTC, authority for train movements is provided by signal indications. The train dispatcher at the control center determines train routes and priorities, and then remotely operates switches and signals to direct the movement of trains. The CTC system is designed so that the dispatcher cannot grant conflicting authorities. For routine operations, most CTC systems can be programmed with selected predefined routes to reduce the dispatchers workload.

CTC involves the placement of block signals along the track at predetermined intervals and locations to indicate to each trains engineer operating a train on the condition of the track block ahead. It also uses electric track circuits to detect the presence of a train or a condition such as a broken rail. Switch point detectors are used to detect open track switches. In the simplest form, a block of track (usually several thousand feet in length) carries a low level electric voltage. Each block is separated from the next block by insulated joints between the rails that prevent current from flowing from one block to the next. When

a train or other vehicle with steel wheels and a steel axle connects the two energized rails, the circuit is completed between the two rails and the signal system responds by displaying a stop signal at the entrance to that block.

Some CTC systems have been enhanced to provide direct indications of wayside signals aspects to the locomotive engineer inside the locomotive cab. Signal aspect is the appearance of the signal, as opposed to a signal indication, which is the information conveyed by the appearance of the signal. These cab signal systems provide an on-board display of trackside signal indications through the transmission of signal aspect information in coded pulses along the track. The engineer controls the speed of the train with the signal information, and obtains authority to enter sections of track. Further refinements called automatic train stop (ATS) or automatic train control (ATC) systems automatically cause the train to stop or reduce speed where an engineer fails to respond appropriately to a trackside signal.

The design of traditional signal systems is based on the principle of fail safe design. The result is that most malicious tampering with the signals, such as cutting wires, will result in the display of the most restrictive signal aspect. This usually results in a false stop, which is the "fail-safe" state. However the false clear, a situation in which the signal displays any other aspect than its most restrictive, when it is supposed to display its most restrictive, is still possible. With false stops there is no violation of an operating authority, rather there is a nuisance factor, and the potential that an excessive number of system faults and warnings may be treated improperly. False clears, on the other hand, allow unsafe violations of operating authorities to occur. In the event of a total compromise of a traditional signal system, the railroad can resort to TWC or DTC mode of operation, with a corresponding reduction in operating efficiency.

2.3 Limitations of Current Technologies

Significant improvements in train operations have resulted in dramatic decreases in accident rates (Figure 2.1) [7]. This has been accomplished using the traditional methods of verbal, authority, mandatory directive, and signal indication despite significant increases in freight volumes. However the potential for an accident with catastrophic consequences still exists. The January 6, 2005 collision between two Norfolk Southern freight trains in Grantville South Carolina that resulted in a chlorine gas release is such an example. This accident resulted in the death of 9 people, the hospitalization of over 200 others, and the evacuation of an additional 5400 from their homes [8].

There are three technologies currently in general use to reduce the impact of failures in compliance with the railroad methods of operation as well as improve throughput. These are (a) cab signals, (b) ATS, and (c) ATC.

Cab signals simply relay the external signal indications to a visual display inside the cab of the locomotive, making it easier for the crew to note the signal aspect and the associated order it conveys. Unless operated with ATS or ATC, the cab signal systems do not provide speed or authority enforcement. Consequently no mechanism would exist to detect and prevent crew non-compliance with dispatcher orders and railroad-operating procedures.

ATS provides enforcement for signal indications. This can be done with or without a cab signals system in place. ATS however, does not provide speed enforcement. It only enforces the indication provided by the wayside signal in the event that the train crew fails to react. ATC, on the other hand, provides both signal indication enforcement and speed enforcement.

In general, ATS and ATC systems rely on the relay of information through audio-frequency (AF) current to transmit ATS or ATC related information along the track circuit [9]. This approach has some significant technical limitations. First, the location of trains can only be

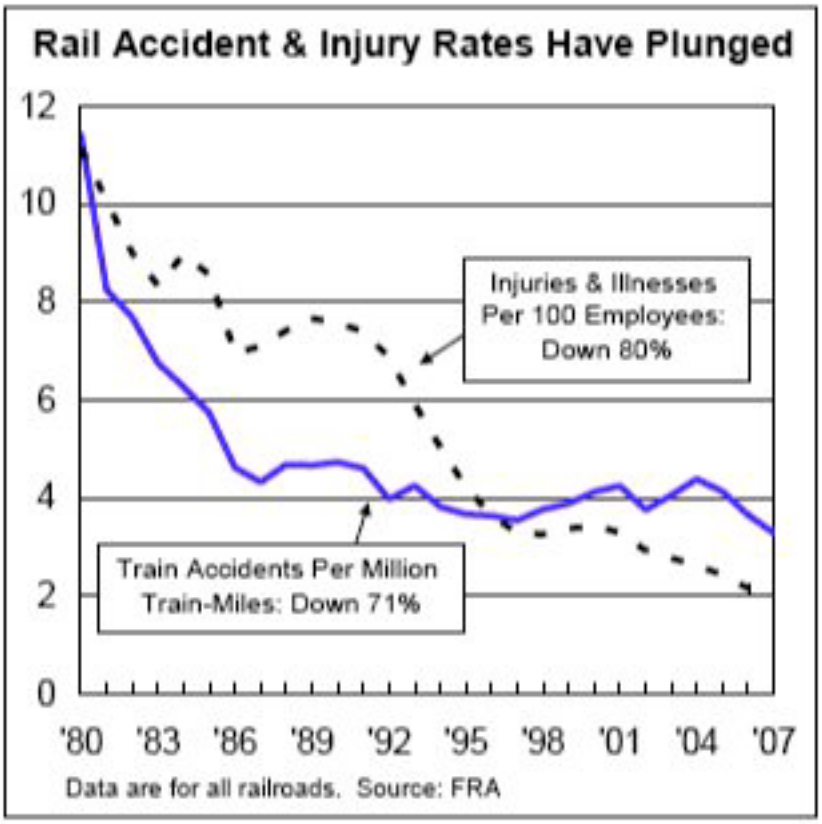

Figure 2.1: Accident Rate

determined to the resolution of the track circuits. If any part of a track circuit is occupied, that entire track circuit must be assumed as occupied. The track circuits length can be made shorter, but adding additional track circuits requires additional wayside hardware. This imposes additional costs, causing a practical (and economical) limit to the number of track circuits that a railroad can install. Second, the information that can be provided to a train through a track circuit is limited to a small number of wayside signal aspects or speed data.

In addition, the underlying signal system infrastructure to provide the required indications for cab, ATS, or ATC to operate are capital intensive. In 2003, the Class 1 railroads alone spent over $490 million in operations, administration, and maintenance of all types of communications and signaling systems with another $153 million in deprecation of the existing plant [10] on approximately 65,000 miles of track. Consequently the deployment of these technologies is limited to those areas where rail throughput needs to be high. Fewer than 5 percent of route-miles in the US [11] have systems in place that have signal indications in the locomotive cab or there is on-board enforcement of the signal indications, or both.

2.4 Railroad Dispatching

Railroad dispatch systems are decision support systems, providing train movement and authority recommendations to the dispatcher. Although dispatch systems provide recommendations, the dispatcher ultimately remains responsible for evaluating the dispatch system generated recommendations, and issuing authorities and/or directives as required. The complexity of dispatch systems varies widely. Simple systems may only keep track of train and vehicles position currently operating on the railroad. Complex dispatch systems may not only keep track of positions, but also determine optimal positioning and timing of train movements along a route (i.e. scheduling) to ensure the most efficient routing of the

trains through a railroads network. Dispatch systems are used to support both passenger and freight railroad operations.

The dispatching problem (optimization of position, schedule, and routes) has been the focus of extensive work by the operations research community. Depending on the functionality implanted by the system, multiple algorithms are used to provide the information required by the dispatcher. At the low end of complexity, positions of trains are defined in terms of the track block in which they are located, and the dispatch system provides recommended movement instructions in terms of the blocks. The operations research community developed algorithmic approaches for block position optimization starting in the early 1980s [12]. Early algorithmic solutions to the scheduling, and the routing problems were also identified at this time [13].

Integrating all three of these functions to determine not only an optimal local (positioning but global (scheduling, and routing) solution followed shortly thereafter [14]. Subsequently numerous alternatives for solving this integrated problem have been developed. Virtually all computer dispatch systems that address the positioning, scheduling, and routing problems in operation today from the major vendors (Alstom, Advanced Railway Concepts, Digital Concepts, GE Transportation, Siemens, Union Switch and Signal) are proprietary. As a result, the exact mechanisms by which they address these three issues are not known. Despite this, it is relatively safe to assume that they implement some variant of exact, heuristic, or simulation solution as defined by Suteewong [15].

While the positioning problem associated with freight and passenger trains is similar, there are significant differences between the scheduling and routing problem associated with freight and passenger service. Passenger service must run according to relatively fixed schedule and constant route. Freight, on the other hand, does not have the same restrictions. Freight traffic generally does not have the same level of time constraints, nor a route

as fixed as passenger service.

Modern dispatch systems play an important role in efficient rail operations. By considering the global rail network and providing the dispatcher global visibility of the rail network and its traffic, bottlenecks can be seen in advance, and traffic rerouted as necessary. This results in increasing system velocity (the average rate at which trains move through the network) and consequent increase throughput in the network. This improved utilization of the network directly translates to operational savings to the railroad. When coupled with PTC functionality further operational and safety efficiencies accrue. Electronic delivery of authorities, for example, eliminates the error prone and time consuming copy-read back exchange between locomotive crew and the dispatcher.

However, such efficiency gains have a price. Dispatchers and train crews generally operate as teams, and over time develop the capability of recognizing each others radio personal. When there is an unexpected change in the radio persona, either on the part of the crew or dispatcher, alternative methods may be used to establish communications to verify the authenticity of the parties and train orders. With electronic delivery of authorities in PTC systems, this recognition is lost. Given the susceptibility of the wireless networks used in PTC to attack, additional security measures are required. In order to determine the specific security measures, an understanding of PTC systems is required. PTC systems are the subject of Chapter 3

Chapter 3: CBTC SYSTEMS

The previous methods of operation are all supported by CBTC systems. Although CBTC systems are commonly implemented in the transit sector, recent completion of US Department of Transportation (DOT) initiatives [5] have just now provided a performance based regulatory framework to encourage industrial adoption of these systems for the general rail system. Various railroads in the United States are now experimenting with different implementations. The inability of cab signals, Automatic Train Stops (ATS), and Automated Train Control (ATC) to effectively incorporate collision and accident avoidance measures with the current methods of operations has been the primary motivation for the US National Transportation Safety Board (NTSB) recommendation to install PTC [16]. These CBTC systems can overcome the fundamental limitations of conventional ATS and ATC Systems.

3.1 The Regulatory Framework

The new regulatory framework supporting change consists of amendments by the Federal Railroad Administration (FRA) of the US Department of Transportation (DOT) to the Rules, Standards and Instructions (R S& I) for railroad signal and train control systems. These new regulations, amending Parts 209, 234, and 236 of Title 49 of the Code of Federal Regulations became effective June 6, 2005 and are known as the Standards for Development and Use of Processor-Based Signal and Train Control Systems. The new regulations were the result of a joint effort by the FRA and the railroad industry that started in 1997.

FRA and the railroad industry recognized that advances in technology in signal and train

control systems had overtaken the existing prescriptive signal and train control regulations, and that changes were needed. The advanced technologies coming into use had not been foreseen when the original Rules, Standards and Instructions (R S& I) were developed, and consequently these new advanced technologies were being regulated on a case-by-case basis. The new regulations eliminate case-by-case regulation. They specify an implementation-independent method of promoting the safe operation of trains on railroads that use processor-based signal and train control equipment. The new regulations are a performance-based standard with only two simple conditions: First, the new system must be at least as safe as what it replaces. Second, the implementer is responsible for demonstrating the safety claims of the new system.

Some parts of the new regulations, such as those dealing with software configuration management, are mandatory for all railroads. Others parts are only mandatory for railroads that are required by statute to implement PTC. The regulations provide four significant advantages that improve flexibility and cost effectiveness:

- First, only railroads that are required by statute to implement the new technology covered by the regulation are required to comply and bear the associated costs.

- Second, the regulations are technology neutral, so the railroad is free to pick the implementation technology best suited to their requirements.

- Third, railroads have the opportunity, within limits set by law, to select when and where to implement new technologies, allowing them to do so as their business case and finances support.

- Fourth, the regulations, being performance and risk-based, allow for customization. The solutions can be based on the probability and frequencies of occurrence of potential mishaps in the railroads operational environment

The new regulatory framework opens the potential for increased innovations by removing prescriptive design and technological limitations. These are reflected in the implementation of the basic architectural and functional requirements.

Recent accidents, most notably a collision between a Union Pacific freight train and a Metrolink Commuter train in September 2008 which killed 25 people and seriously injured another 132, have resulted in the US Congress passing a statutory mandate for the installation of PTC on all track carrying intercity or commuter traffic, all main lines (those that cary more than 5 million gross tons per year) belonging to the Class 1 Railroads, and any line over which toxic by inhalation material is carried [6]. The Federal Railroad Administration is currently developing new regulations for mandatory PTC system installation to supplement the existing voluntary regulations.

3.2 Basic Architectural and Functional Requirements

PTC systems are complex systems made up of distributed physical, but closely coupled, functional sub-systems. Their successful operation requires a well-orchestrated set of interactions. Understanding the basic PTC architecture, PTC functional requirements, and modes of operations assists in understanding a PTC system. All such PTC systems are derivations of a single basic functional architecture, with specific enhancements and modifications to both functions and modes of operations to support the unique requirements and operational needs of the individual railroad purchasing the system.

The basic functional architecture, illustrated in Figure 3.1, consists of three major functional subsystems: wayside, mobile, and dispatch/control. The wayside subsystem consists of elements such as highway grade crossing signals, switches and interlocks or maintenance of way workers. The mobile subsystem consists of locomotives or other on rail equipment, with their onboard computer and location systems. The dispatch/control unit is the central office that runs the railroad. Each major functional subsystem consists of a collection of physical components implemented using various databases, data communications systems, and information processing equipment.

The basic architecture implements a set of common requirements, and supports various optional requirements. The common functional requirements, known as PTC Level 1, are:

- Preventing train-to-train collisions, referred to as positive train separation.
- Enforcing speed restrictions, including civil engineering restrictions and temporary slow orders.
- Protecting roadway workers and their equipment operating under specific authorities.

The additional functionalities that augment PTC Level 1 are divided into Levels 2, 3, and 4, and the requirements are cumulative, as shown in Table 3.1. For example PTC level 3 includes all requirements for PTC Level 2.

3.2.1 Classification by Control

The PTC mode of operations can be further refined in terms of which subsystem is responsible for executing the majority of the operations required for the execution of PTC functionality. In mobile-based modes of operation, a control unit component in the mobile subsystem is responsible for the majority of the effort required to implement the various PTC functions. The wayside subsystem and dispatch/control subsystem communicate required control data to the mobile subsystem control unit. The mobile subsystem control

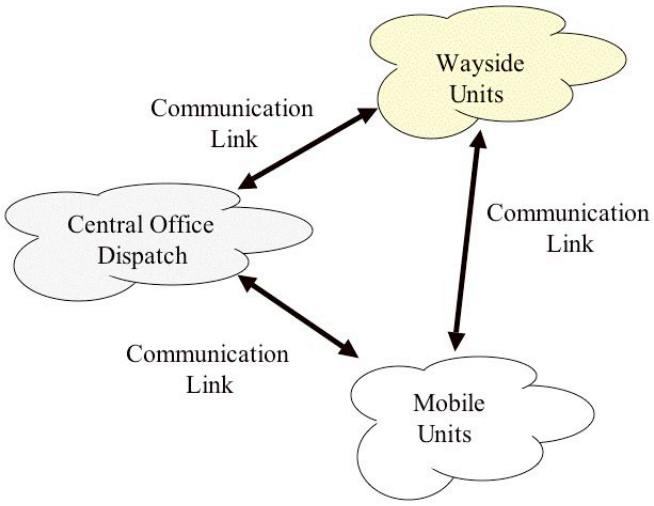

Figure 3.1: Basic Architecture

Table 3.1: PTC Functional Levels

PTC Level	Functionality
0	None
1	-Prevent train to train collision -Enforce speed restrictions -Protect roadway workers
2	PTC Level 1 + Automated Digital Dispatch of Authorities
3	PTC Level 2 + Wayside monitoring of the status of all switch, signal, and protective devices in traffic control territory
4	PTC Level 3 + - Wayside monitoring of all mainline switches, signals, and protective devices - Additional protective devices such as slide detectors, high water, hot bearings - Advanced broken rail detection - Roadway worker terminals for communications with dispatch and train

unit analyzes the received data, interprets it into actions for each subsystem and transmits the appropriate directives. The wayside subsystem components, the dispatch/control subsystem, or other components of the mobile subsystem then translates these directives into specific commands appropriate to the underlying hardware implementation that executes them.

In dispatch/control-based modes of operation, a control unit in the dispatch/control subsystem is responsible for most of the logical effort required to implement the various PTC functions. The wayside subsystem and mobile subsystem communicate required control data to the dispatch/control unit. The dispatch/control unit takes and receives data, analyzes it, interprets it into actions for each sub-system, and transmits the appropriate directives. The wayside subsystem components, the mobile unit subsystem components, or other components in the dispatch/control subsystem then translate these functional directives into specific commands appropriate to the underlying hardware.

A similar chain of relationships occurs in wayside based modes of operation- a control unit in the wayside subsystem is responsible for the majority of the logical effort required to implement the PTC functions. Mobile and office/dispatch subsystems communicate data to the wayside control unit. They or other components in the wayside subsystem receive functional directives for the underlying hardware in return.

In all three of the preceding modes of operation, the mobile office/dispatch, and wayside subsystems are self-monitoring and can act independently when failures and defects are detected. This assures fail-safe operation even when communications is lost.

3.2.2 Classified by Method of Operations

In addition to classification by functionality, PTC systems are also classified by the extent that they are used to augment the exsting method of railroad operations. This classification scheme also provides an example of the flexibility for both regulators and regulated entities with respect to enforcement and compliance issues. Full PTC systems completely change, or replace, the existing method of operations. Overlay PTC systems act strictly as a backup to the existing method of operations; which remains unchanged. The distinction between Full and Overlay in this classification schema, however, is undermined; over an argument of the extent that first and second order safety functionality is required to be directly associated with the term Overlay.

First order safety functionalities are those mandatory to ensure safe system operation. Their loss would potentially result in unsafe system operations. Second order safety functionalities are those that, when used in conjunction with another function, are mandatory to ensure safe system operation. Loss of a single second order function will not result in an inability to continue with safe system operations, unless coupled with the loss of an additional second order function.

One view suggests that Overlay systems do not require either first or second order safety functionality, bur rather that an Overlay system acts strictly as an aid to the train crew. With this type of Overlay system in place, the additional information provided to the crew can increase safety, since the crew is provided with additional information that better enables them to execute their responsibilities. In the worse case, a failure of the Overlay system, the train crew continues to operate under the same rules as before the installation of the Overlay with no loss of safety.

A different view suggests that certain limited aspects of Overlay PTC systems must have first order safety functionality and be treated accordingly. For example, one of the PTC

Level 2 functionalities replaces voice transmission of authorities between the dispatcher and crew with digital transmission of authorities directly to the onboard train control computer. In such a situation, the authorities would be sent and received, entirely by machine, without human intervention.

Given the role of the authority in safe railroad operations, this functionality and the implementing components would be classified as providing first order safety functionality. Since the new regulation for implementing PTC systems specifically requires that there must be no net reduction in safety, it would seem that failure to implement this in an Overlay system as a first order safety function would cause a net reduction in safety

Yet another view argues that crew over-reliance in an Overlay PTC system requires second order safety functionality. In the situation of crew over-reliance on the system, the system is no longer simply an aid to the crew. Therefore, although the system may not fit the criteria for first order safety functionality, the Overlay should be treated as providing second order safety functionality. Extending this reasoning, a crews over-reliance on the overlay may change over into their primary means of operation, relegating the nominal secondary mode to a primarily mode. Consequently, what first started as second-order safety functionality can evolve to become first-order safety functionality.

The technical challenge is demonstrating that there will be no reduction in safety over time based on the PTC Level implemented. An arbitrary association of required first and/or second order safety functionality for an Overlay will exist. Consequently, we advocate the position that the required order of safety functionality for an Overlay system must be evaluated on its individual merits.

3.3 Current US System Implementations

Today in the US there are 11 PTC systems either deployed or in development on over 3000 route miles on 8 railroads across 21 states. While they all provide the Level 1 PTC functionality, they are classified differently. Generally these systems function as designed, although they have encountered technical deficiencies. Although the vital (fail safe) or safety-critical functionalities have not been affected, the deficiencies potentially could adversely impact the efficiency of train operations. PTC issues encountered include: limited geographic communication coverage; appropriate and accurate location of infrastructure critical points; appropriate and accurate braking distance prediction; and train tracking.

The systems that are either operational or being deployed for revenue service are the Advanced Civil Speed Enforcement System (ACSES), Incremental Train Control System (ITCS), Communications Based Train Management (CBTM), Electronic Train Management System (ETMS) Version 1,Version 2, and METRA Configuration, Vital Train Management System (VTMS), Collision Avoidance System (CAS), Optimized Train Control (OTC), and Train Sentinel (TS). The remaining system, the North American Joint Positive Train Control (NAJPTC) System is not currently being deployed in revenue service

Developed for the US National Passenger Rail Corporation (Amtrak), ACSES is installed and fully operation on 240 route miles of the North East Corridor (NEC) between Boston, MA and Washington, DC. It supports Amtrak's ACELA, currently the fastest passenger service in the US, to speeds up to 150 miles per hour. ACSES is a track embedded transponder-based system that supplements the exiting NEC cab signal/automatic train control system.

Amtrak also operates the ITCS system to support high-speed passenger operations Niles, MI and Kalamazoo, MI. Operating on 74 route miles, ITCS currently supports speeds up to

95 miles per hour. It is unique from other PTC system implementations in that it includes advanced high-speed highway-rail grade crossing warning system starts using radio communication rather than track circuits. Depending on the reports received from the Highway Grade Crossing Warning (HGCW) system, the ITCS onboard imposes and enforces appropriate speed restrictions. Upon completion of the verification and validation of the software, maximum authorized speeds will be raised to 110 miles per hour.

CSX Transportation is preparing to field test the latest version of their CBTM on approximately 200 route miles of their Aberdeen and Andrews SC Subdivisions. Early versions were installed on their Blue Ridge and Spartanburg SC lines. Current CSX efforts are focused on harmonization of CBTM with the BNSF Railways ETMS Version 1 and 2, the Union Pacific (UP) Railroad Vital Train Management System (VTMS), and the Norfolk Southern (NS) Optimized Train Control (OTC) to interoperate freight trains.

BNSF Railways has undertaken an extensive PTC development and deployment effort to support their freight operations. ETMS Version 1 for low-density train operations has received full approval from the Federal Railroad Administration, and BNSF has stated deployment on 35 of their subdivisions. BNSF also has an enhanced version of ETMS, ETMS Version 2, to support high-density train operations under active test on their Fort Worth and Red Rock Subdivisions in TX.

A related configuration of ETMS Versions 1 and 2 is under development for the Commuter Rail Division of the Chicago Regional Transportation Authority (METRA). Created in response to a series of fatal accidents resulting from train over speeding or exceeding, the METRA implementation of ETMS is intended to support passenger commuter, as opposed to freight, operations. This system is under deployment on the Joliet and Beverly Subdivisions in Chicago, IL.

Unlike the METRA, CSXT, and BNSF variants of ETMS, which are Overlays, the UP and NS are developing Full (or vital) system variants of ETMS. The UP VTMS has begun test operations on 15 different UP subdivisions in Washington State in the US Pacific Northwest and the Powder River Basin of WY. The NS OTC variant, which integrates their new NS Computer Aided Dispatch (CAD) System with PTC and other specialized business functionalities, is under test on the NS Charlestown to Columbia SC Subdivisions. CBTM, ETMS, OTC, and VTMS are all developed by the same manufacturer, and share a common code base. They differ only in their specific hardware configurations.

The Alaska Railroad is undertaking installation of CAS on all 531 miles of their system. Also designed to be a Full PTC system, it is built to implement the same PTC functional architecture as other PTC systems using completely different hardware and software. CAS enforces movement authority, speed restrictions, and on-track equipment protection in a combination of Direct Traffic Control (DTC) and signaled territory. All of the wayside and office components have been installed and tested, and onboard system test operations are in progress on the to Portage and Whittier Subdivisions outside of Anchorage, Alaska.

The Ohio Central Railroad System (OCRS) version of a PTC system is the TS. TS is currently in use on various railroads in South and Central America. The OCRS version of TS is based on the TS installation currently operating in mixed passenger and high-speed freight service on the Panama Canal Railroad Balboa and Panama City in the Republic of Panama. The OCRS has completed installation of their office subsystem, and is conducting integrated office, wayside, and onboard subsystem between Columbus, and Newark, OH.

The NAJPTC, a joint effort of the FRA, the Association of American Railroads (AAR), and the Illinois Department of Transportation to develop an industry open standard high-speed passenger and freight service, was removed from service due to technical issues associated with communications bandwidth. The system was relocated to the US Department of

Transportation (DOT) Technology Transportation Center (TTC) Test facility in Pueblo, CO, for study and resolution of the communications issues associated with the standard in a controlled environment.

Finally, although no field-testing or deployment work has occurred, the Port Authority of New York and New Jersey (PATH) has begun design work on entirely separate and independent version from the CSXT Communications Based Train Management (CBTM) System. The PATH CBTM will provide PTC functionality to the Trans-Hudson River Commuter Rail Line running underground between New Jersey and New York City.

3.4 Overlay and Vital Implementation

All the preceding systems utilize the same basic architecture, albeit different hardware, to implement PTC Level 1 functionality using some variant of the four classifications discussed earlier. Of these systems, I will detail two, ETMS and ITCS, as representative implementations.

ETMS is an Overlay system [17–19] designed and built by WABTEC Railway Electronics for freight trains. The system (Figure 3.2) consists of 4 segments- Onboard, Wayside, Communications, and Office (Computer Aided Dispatch System- CADS- and ETMS Server). ETMS provides for warning and enforcement of speed restrictions (permanent and temporary), work zone boundaries, and route integrity of monitored switches, absolute signals, and track (rail) integrity. During system operation, train crews are notified of potential violations when they are within a sufficient warning distance that allows them to take corrective action. If the crew fails to take corrective action, ETMS applies a full service brake application to stop the train. The method of operations does not change, however, and crews are responsible for complying with BNSF Railways operating rules at all times. ETMS has

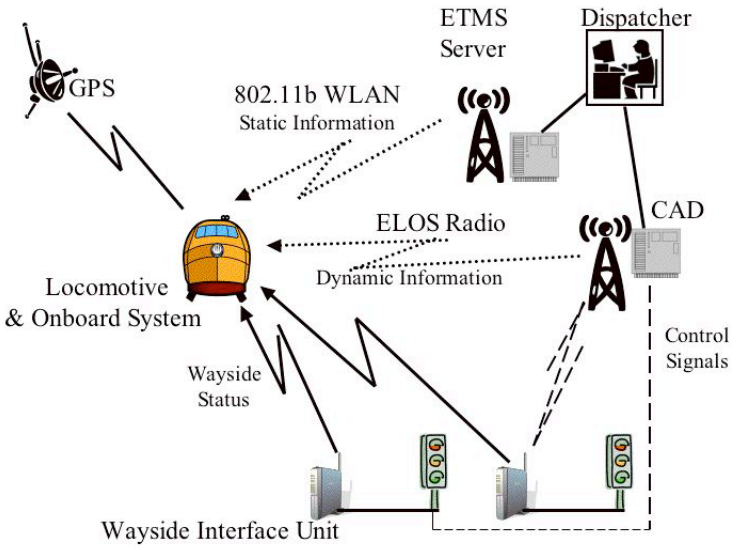

Figure 3.2: Simplified ETMS Architecture

been approved for operations on 35 subdivisions of the BNSF Railways. A second more advanced version is under test on the BNSF Red Rock and Fort Worth subdivisions.

The major components of the ETMS Onboard segment consist of the engineers color display, a brake interface, a radio, a differential GPS system and using a train management computer. The train crew obtains information by a series of complex graphics on the display of the track configuration and geometry, switch position, signal indication, authority limits, train direction and makeup, current speeds, max speed, distance to enforcement, time to enforcement, geographical location and text messages. These are augmented by the use of selective color highlighting and audible alarms. The text messages either describe enforcement action in progress, or advise of a condition or required action. In addition, all applicable active warrants and bulletins can be recalled from the onboard database.

The primary means of determining position is via differential GPS. The train management computer continuously compares its GPS position with the stored position of speed restriction zones, work zones, and monitored switches and signal from the track data base

in non volatile memory. As the train management computer determines that the locomotive position is approaching the position of speed restriction and work zones, the train management computer system automatically calculates and activates the brake interface as required. The braking enforcement curves are updated dynamically based on reported changes.

The Wayside segment consists of a set of interface units that act as a communications front end for switch position, signal indications, and broken rail indications. The onboard system monitors the indication transmitted by the wayside interface units in the trains forward direction of movement. The wayside interface unit provides the latest state of monitored devices, and the onboard system will accept changes in the indication (with the corresponding changes in required enforcement activity) up to a set distance before reaching the monitored device, after which point a change is ignored.

The communications system consists of a wireless 802.11b broadband network to transfer track database information and event logs at selected access points along the track, and an extended line of sight communications (ELOS) network for other data exchange. There is direct exchange of data over the Communications segment between the Wayside and the onboard system, as well as between the Office and Onboard system.

The Office system consists of the CADS and an ETMS server for providing train authorities, track data, consist data, and bulletins. Static information, such as track data is stored in the ETMS server portion of the ETMS Office System, while dynamic information, such as authorities are stored in the CADS portion of the ETMS Office System.

ITCS is a Full PTC system [20–22] designed and built by GE Transportation Systems-Global Signaling for both freight and passenger trains. The ITCS system also consists of 4 basic segments: Communications, Onboard, Wayside and Office. The system provides for

high-speed operations through wireless grade crossing activation and verification, warning and enforcement of speed restrictions (permanent and temporary), work zone boundaries, and route integrity of monitored switches and absolute signal integrity. The system design is such that a system failure results in a guaranteed enforcement. It is integrated with the existing Traffic Control System (TCS) where it obtains its signal indications. ITCS is designed to support passenger trains up to 110 mph, pending completion of software verification and validation it is operating at speeds up to 95 mph between Niles and Kalamazoo, MI.

The Communications segment consists of a radio network that allows communications between the Wayside segment components (which consists of the Wayside Server and Wayside Interface Units associated with each instrumented switch, crossing, and signal) and between the Wayside Servers and the Onboard segment. Also associated with the communications system are direct dial telephone lines from the office segment to the Wayside servers. These lines allow for the gathering of health and management information about the servers as well as posting of temporary speed orders.

The major components of the Onboard segment are an engineers display, differential GPS an on board computer and brake control interface and a track database, The engineers display is a simple LED display that indicates current speed limit, the actual speed, distance to the next enforcement target in the database, and time remaining to penalty enforcement augmented with audible alarms. An LCD is also provided to display simple text messages on software version and the locomotive type defining the braking enforcement curve.

Similarly to ETMS, the ITCS primary means of determining a train position is via differential GPS. The ITCS onboard computer also continuously compares the received GPS position with the stored position of switches, signal, and crossings and permanent speed restrictions in a non-volatile track database. The ITCS onboard computer also receives updates from the wayside servers of temporary speed order locations, interlock positions, and

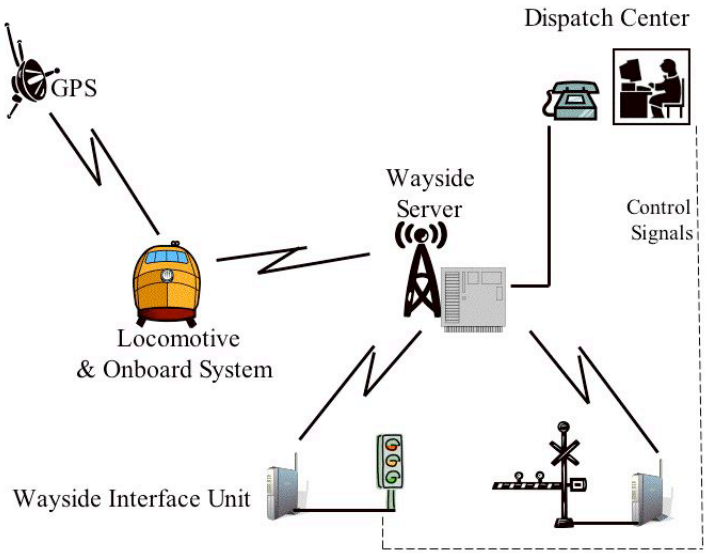

Figure 3.3: Simplified ITCS Architecture

signal indications. Using the received updates and its known position, the ITCS onboard computer automatically calculates warning and enforcement actions and activates the brake interface as required. The braking enforcement curves are not updated automatically- once a particular curve for a particular locomotive type is selected, the selection remains in force until another curve for a different locomotive type is manually selected.

The Wayside segment consists of individual interface units linked to a concentrating server. The individual wayside servers, which aggregate geographically similar wayside interface unit status and control information for communication to the Onboard System. The wayside server stores all work zones, temporary speed restriction, received switch positions, and received highway-grade crossing status indicators.

The Onboard system can actively control highway-grade crossings via the Wayside Segment. If the wayside segment reports a crossing is active, the onboard system signals the Wayside segment to arm the crossing and lower the gate based on the expected arrival time

of the train. The Wayside server signals the Wayside Interface Unit, which in turn orders the crossing to lower the gate. Once the crossing indicates the gate is down, it reports through the Wayside Interface Unit and the Wayside server to the onboard system. The Wayside segment monitors the crossing to ensure the crossing continues to report that it is in the down position. The Onboard system continuously evaluates the reported status from Wayside segment. In the event that a fault develops braking is automatically applied by the onboard system.

The Office segment is used to input temporary speed orders for transmission to the various wayside servers, and to display collected health and management data from the wayside servers.

3.5 PTC System Protection

In order to effectively address rail security issues, the security threat and consequences of successful exploitation of security vulnerabilities is required. Understanding the role and risks associated with PTC, and appropriate mitigations, requires an understanding of the entire threat environment as well as the vulnerabilities associated with the communication subsystem. Successful exploitation of non-communication vulnerabilities can aggravate the adverse consequences of communications vulnerabilites, just as successful exploitation of communications vulnerabilities can agravate the consequences of non-communications vulnerabilities. Both communication and non-communication vulnerabilities are the subject of Chapter 4.

Chapter 4: SECURITY THREATS

Railroads are a critical transportation asset and play a significant role in the United States economy. They operate in every state in the US except Hawaii, across a network that exceeds 140,000 miles (Figure 4.1). Use Cases capture functional requirements in terms of necessary interactions between an actor and the environmental constraints under which the system and its actors operate. Due to the recent trend of misusing and/or abusing systems defects and vulnerabilities by various mal actors, Use Cases have been augmented by Misuse Cases to specify and eliminate known undesirable interactions between Mal-actors and a system under design. The scale of rail operations is massive. In 2006 the seven major Class 1 railroads alone employed over 167,000 people at an average total compensation of over $94,000, moving over 1.7 trillion ton miles of freight, with revenues exceeding $40 billion [7]. The freight hauled was a diverse mixture (Table 4.1) of commodities that support all facets of the US industrial base. These include coal, industrial chemicals, ethanol, plastic resins, fertilizers, agricultural products; non-metallic minerals; food and food products; steel and other primary metal products; forest products, motor vehicles and motor vehicle parts; as well as waste and scrap materials (Figure 4.2)

The 559 freight railroads operating in the United States provide two different types of service, point-to-point and switching/terminal. Railroads providing point-to-point service primarily provide intercity pickup and deliveries, while railroads providing switching/terminal service primarily provide local pickups and delivery. Additionally, railroads providing point-to-point service are divided into three different classes based on their annual gross revenues. In 2006, Class 1 (Table 4.2) railroad gross revenue exceeded $346 million; Class 2 railroad gross revenues exceeded $40 million, whereas Class 3 railroads grossed less than $40 million.

Figure 4.1: Rail Network

Table 4.1: Shipments by Product Classification

Product	Tons (millions)
Coal	852,061
Chemicals	168,275
Farm Products	149,392
Non-Metallic Minerals	140,871
Intermodal Freight	105,433
Metal Products	62,256
Metallic Ore	60,601
Petroleum & Coke	55,449
Stone & Related	51,191
Waste & Scrap	48,280
Lumber & Wood	42,956
Pulp & Paper	37,225
Motor Vehicle	37,225
All Other	22,294

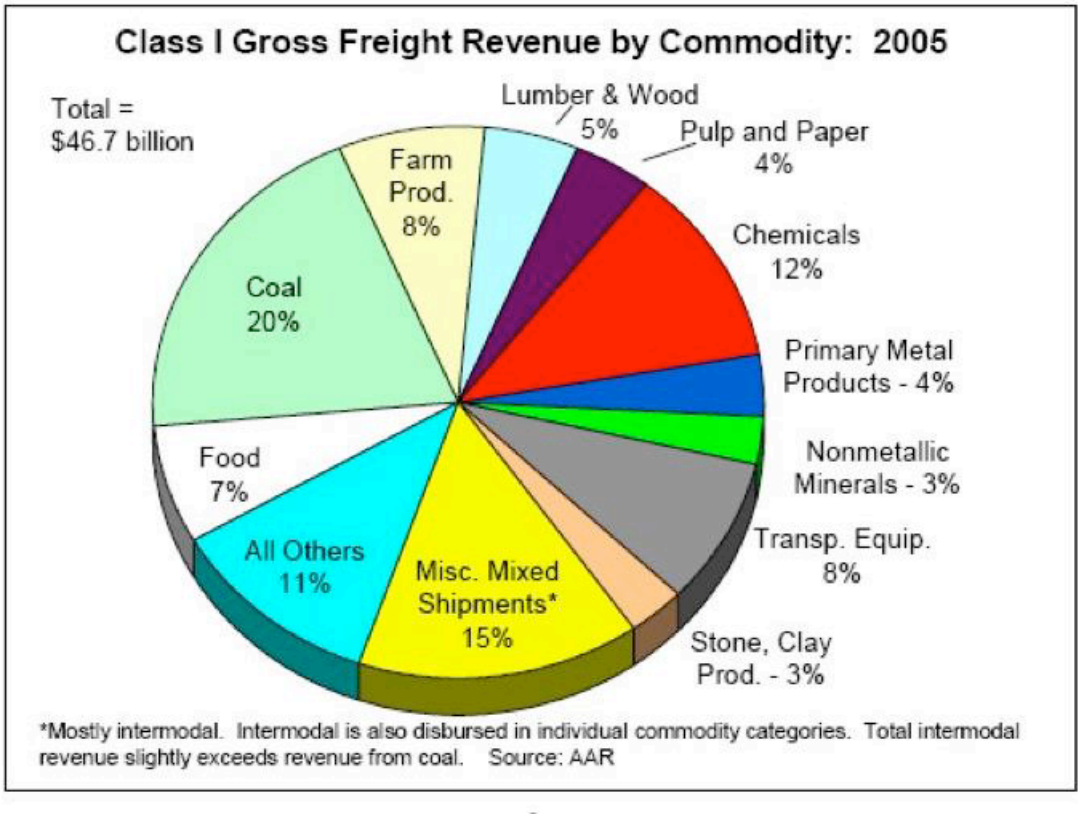

Figure 4.2: Shipments by Product Classification

Table 4.2: Class 1 Railroads

BNSF Railway
Canadian National (Grand Trunk Corporation)
Canadian National (Illinois Central)
Canadian Pacific (Soo Line)
CSX Transportation
Kansas City Southern Railway Company
Norfolk Southern Combined Railroad Subsidiaries
Union Pacific Railroad

Switching and terminal railroads are not subdivided by revenue.

The criterion for classification of passenger railroad service is the distance over which the service is provided. Commuter railroads are passenger operations conducted over short distances and intercity passenger services cover longer distances. In the United States, there is one intercity rail service provider, AMTRAK, and 22 commuter rail [23] providers. Commuter railroads carry the majority of passenger traffic. Roughly 40% of all intercity freight goes by rail, including 67% of the coal used by electric utilities to produce power. Railroads also operate the 30,000 miles of the Department of Defense Strategic Rail Corridor Network (STRACNET) for the movement of Department of Defense munitions and other materials [24].

Disruptions in railroad services can have a significant adverse impact on the US economy as well as military preparedness. In 1998, for example, service problems in Texas on the Union Pacific resulted in direct costs of $1.093 billion and an additional $643 million in additional costs to consumers [25] More recently, commuter rail and CSX freight rail service along the East Coast of the United States experienced cancellations and delays of up to 24 hours. Neither of these was the result of any deliberate attack. In the case of the former, it was the inability to position and move equipment. In the later it was the result of the accidental introduction of a computer virus that disabled the computer systems at the CSX headquarters [26].

The effect of disruptions can be more than inconveniences or lost revenue. The freight moved by the railroads includes 1.7 to 1.8 million carloads of hazardous material [27]. One extremely hazardous subset of this material is known as Toxic by Inhalation (TIH) material. TIH material are gases or liquids that are known or presumed on the basis of test to be so toxic to humans as to pose a health hazard in the event of a release during transportation [28]. Although TIH materials constitute only 0.3% of all hazardous material shipments by rail, this still equates to over 21.6 million ton miles of TIH movements per year [29].

Although a statistically rare occurrence, the effects on public health from the release of hazardous substances during rail transportation are potentially catastrophic. [30]. For example, each year 8500 tank cars of chlorine move by rail through the middle of Washington, DC passing within 2 blocks of the U.S. capital. In a worst-case scenario, the complete release of the contents of just one 90-ton car of chlorine in the center of Washington, DC has the potential to kill or injure 100,000 people [31]. The level and duration of chlorine gas exposure that results in death is as low as 430 parts per million for periods of 30 minutes. Death is by slow suffocation as the chlorine gas reacts with moisture in the lungs, forming hydrochloric acid. Exposure, even if not fatal, can result in lung congestion, pulmonary edema, pneumonia, pleurisy, or bronchitis [32].

4.1 Vulnerabilities

The disperse nature of railroads provides an extremely large number of points where an attack can be made. Attack opportunities include both the physical as well as the communications infrastructure [33, 34].

4.1.1 Physical Vulnerabilities

The critical physical infrastructure elements can be broken into three general categories, track, signal, and motive power and equipment. Each element implements different functions, resulting in unique vulnerabilities.

The track infrastructure includes what the train rides on (the railway), rides through (tunnels), and rides across (bridges and culverts). The railway is composed of ballast, which supports crossties, which in turn support the rails. It functions to restrain the track ties from movement both under static and dynamic loading while providing drainage for the track, and keeps crossties and rails at the proper elevation and alignment. Crossties and rails form the track. Crossties provide support for the rails, communicate, and distribute rail loads to the ballast. They also maintain the rails at the proper gauge and alignment. Rails support the wheels of the train, and provide a smooth surface for the wheels to run over.

A specialized safety device known as a derail or derailer is another element of the track infrastructure. Attached to a rail, it provides protection for a location by directing motive power and equipment off the track before it enters the protected location by derailing the equipment as it rolls over or through the derail. Any derailment is damaging to equipment and track, and requires considerable time and expense to remedy. However, when properly used derails can prevent worse damage than that which would occur if the equipment were allowed to proceed uninterrupted. Derails work by lifting the wheel flange of a locomotive or car from its normal position on the inner surface of the rail to the outside of the rail.

Attacks on the track infrastructure not only require the least level of effort on the part of an attacker, but also provide the greatest window of opportunity for attack. Removal of the spikes that attach the rail to the cross-ties is the simplest type of attack. When unrestrained from the crossties, the rail can move laterally, changing the amount of separation

(gauge) between the rails. If sufficiently unrestrained, the change in gauge can derail the train. Execution of this attack requires nothing more than access to the railway, and a crow bar to remove the spikes. Another simple attack is to misuse a derail. Normally the derail is a last resort safety mechanism to prevent the inadvertent incursion of a train into a location by diverting the train from the rails. When misused, the result is an unplanned, as opposed to planned, diversion from the rails.

Switches provide another avenue of attack. All switches have a manual mechanism at each switch that allows local repositioning. While locks secure the manual switch mechanism, an attacker can forcibly remove the locks. This allows an attacker access to the manual positioning mechanism to reposition the switch. A train entering the switch at speeds greater than the design speed of the switch, or upon a partially repositioned switch, may derail.

Tunnel, bridge, or culvert attacks require a significantly greater level of effort on the part of the attacker, and generally require the use of some sort of explosive or pyrotechnic device. Unless the train and tunnel, bridge, or culvert attacks are collocated in space and time, the collateral damage to the railway from the tunnel, bridge, or culvert attack results in damage to the train. For a tunnel, this might be debris from a collapsed part of the tunnel blocking the tracks. For bridges and culverts, it might be the collapse of the part, or all, of the structure that causes damage to the railway. In bridges, or culverts, constructed of a flammable material, the use of more readily available incendiary devices can be substituted for explosives. Culverts may also lend themselves to attack by hand where the culvert is under mined so that it fails under static or dynamic loads imposed by the train, and the track geometry changes sufficiently to derail a train.

Signal systems are equally vulnerable to attacks. The purpose of the signal system is to pass information regarding the condition of the track ahead of the train to the engineer and control the trains movement. Signal systems usually consist of a set of colored lights,

mounted above or adjacent to the railway connected to a wayside bungalow, which in turn is connected a control or dispatch center. The configuration of lights seen by the engineer is called the aspect of the signal. Different light configurations have different meanings. The meaning of a specific configuration of lights is known as the indication. Indications include the occupancy of the track ahead, movement authorizations, speed authorization, and switch position. Signals do not convey maximum authorized speed, civil speed restrictions, or temporary speed restrictions. Aspects, and the associated indications, may be displayed directly onboard the locomotive to the engineer, where they are known as cab signals.

Track circuits, the key-operating component of signal systems, are conceptually very simple. In a track circuit an electric current is sent from a source, down one rail, through a relay, then back up the other rail to the source. The length of rail over which this circuit is made is a block, and each block is electrically isolated from adjacent blocks. The current flowing through this path sets the relay, making a separate circuit that controls the signal aspect to indicate the track block ahead is clear. When a train enters the track block, it establishes an alternate path between the energized rails through the train's wheels and axle. This reduces the current flow through the relay, opening the relay and changing the signal aspect to indicate the track block ahead is occupied. The relay is configured so it is fail safe, in the absence of sufficient current the signal aspect is set to indicate the track ahead is occupied. Attacks on a signal system are more difficult to execute than attacks on the track/rail infrastructure. Successful signal system attacks accomplish one of two things. The attacker configures the signals allowing block entry to allow two trains to be in the same place at the same time. Alternatively, they configure the aspect of a signal controlling a turnout so it is not in correspondence with its associated position or operating speed. The former attack results in a collision. The later attack may result in a derailment if the train speed sufficiently exceeds the maximum turnout speed, track conditions beyond the switch are insufficient to support rail operations, or the switch is in the wrong position for the train movement.

Accomplishing either of these attacks requires a high degree of technical sophistication on the part of the attacker. The attacker must have a detailed understanding of the track circuits involved since the attack must overcome their failsafe design. Positive control of the aspects is required, so the appropriate aspects can be set to cause a collision or derailment. Control must be obtained in a manner that cannot be detected by the dispatcher. Coordinated control of a sufficient number of signals must be gained to arrange for two trains to be in the same block at the same time to force a collision. In the case where the signals are geographically disperse, this requires that the attacker have sufficient personnel who can communicate with each other to provide the necessary signal aspects at the correct time. Forcing a collision also requires a detailed understanding of the railroad timetable, the position of the trains to collide, and their velocities so that the appropriate time to take control can be pre-computed. A derailment, while not requiring quite the same level of coordination, still requires the attacker to successfully accomplish the same tasks and have the same level of knowledge.

Signal system failures, known as false proceeds that result in signals passed at danger (SPAD) result in behaviors similar to signal system attacks that would also allow a SPAD. In the case of a false proceed a signal aspect which indicates that the locomotive engineer may proceed into the next block is displayed. This display occurs despite the fact the block that the train is already occupied. The signal system has shown itself to be very robust to disruptions that result in false proceeds. In 2007 on the Class 1 railroads there were only 43 such occurrences in the over 1.7 trillion ton-miles of freight moved.

The third major physical infrastructure element, locomotive and equipment, provide several avenues of attack. The first, and perhaps the most simple, is to gain physical possession of the asset. Because large numbers of rail cars aggregate in rail yards, where they are received, sorted, and stored until made up in a departing train, access to a wide range

of hazardous material, in large quantities can be gained. The proximity of many of these yards to densely populated areas means that the loss of control of the car and release of its contents may affect large numbers of people.

An attackers access to a rail yard is relatively easy. This allows the attacker freedom to control the material release. Most yards are not fenced, and even if fenced, the fence may be easily by-passed. Railroads also do not employ sophisticated physical intrusion detection systems. While larger yards have railroad police conducting regular security sweeps, the limited number of officers available for this task and size of the yard reduces the effectiveness of these patrols. In addition to mass groupings of cars in rail yards, smaller individual collection (called cuts) of cars can be found located in geographically disperse industrial sidings and spurs. Often lightly guarded, if at all, they too are susceptible to attacks wherein the attacker gains physical possession of the car and its contents.

Targeting of specific cars is made easier by the regulation [35] that requires cars containing hazardous material to display placards indicating the type and effects of the material they contain. The objective of this regulation was to assist first responders in determining the positional scope of a situation and the action to be taken, and also identifies the cars whose contents can cause the greatest damage to the surrounding population.

If attackers cannot obtain physical access to a rail car containing a hazardous material, or if the yard or siding location does not allow for a sufficient numbers of casualties, another option would be to hijack the train. This allows the attacker the potential to move the car to the location that would maximize casualties. Hijacking also has the potential to allow attackers to relocate hazardous material cars to a location that may inflict the most damage. A successful hijacking and material movement is difficult. Movement of the train beyond certain geographical limits requires gaining control of the switches and/or associated signals, a time consuming activity. Also, while gaining physical access to a locomotive

may be relatively simple, successfully exploiting that access is not.

Attempts to move a locomotive and its attached consist require great skill. Incorrect movements result in the generation of intra train forces that exceed the coupler knuckle strength. This may result in broken couplers, dividing the train into separate segments. Segmentation of the train precludes the theft of the entire train, and potentially eliminates the relocation of hazardous material cars. Even if an attacker successfully hijacks the train and subsequently tries to induce a hazardous material release through collision or derailment, the odds of success are low. Releases, as a result of derailment or collision, however, are relatively rare. In 2007 only 49 hazardous material releases of all types occurred as a result of 197 collisions and 1849 derailments [36].

4.1.2 Communication Vulnerabilities

Although the communications links between the various PTC subsystems may consist of both wired and wireless links, it is the wireless component of the links that offers the greatest susceptibility to attack relative to the wired component of the links. This is due to the ease of access that an intruder has to the wireless link with respect to the hardwired links. This is, of course, not to say that successful attacks could not be made on a CBTC system through a hardwired communications link, only that the wireless links offers a significantly easier target to exploit.

Recent research has examined security and possible problems in the rail infrastructure and surveyed systems in use [37, 38]. Completion of recent regulatory initiatives, coupled with accelerated industry efforts in the deployment of CBTC systems, have increased the level of risk that the public may potentially be exposed to as a result of the greater use of wireless technology. The most significant source of risk in wireless networks is that the technologys underlying communications medium, the airwave, is open to intruders.

Changes in malicious hacker activity have shifted from conventional fixed wired systems to wireless networks. These networks have included not only traditional telecommunications systems, but also industrial control systems. Studies by the National Research Council and the National Security Telecommunications Advisory Committee [39] show that hacker activity includes the ability to break into wireless networks resulting in the degradation or disruption of system availability. A recent Government Accountability Office study [40] has indicated that successful attacks against control systems have occurred.

While these studies were unable to reach a conclusion about the degree of threat or risk, they uniformly emphasize the ability of hackers to cause serious damage. The resources available to potential intruders are significant [41]. Intelligence is already widely available on the Internet that enables intruders to penetrate any sort of traditional computer network and wireless systems. Detailed vulnerability information is publicly discussed on newsgroups. Tutorials are available that describe how to write automated programs that exploit wireless systems vulnerabilities. Large numbers of automated software tools have been written that enable launching these types of attacks. Publicly available Web sites whose sole purpose is to distribute this data have been established, often ensuring wide spread distribution of the information before public access can be terminated.

The Information Assurance Technical Framework Forum (IATFF), an organization sponsored by the National Security Agency (NSA) to support technical interchanges among U.S. industry, U.S. academic institutions, and U.S. government agencies on the topic of information assurance, has defined five general classes of information assurance attacks- passive, active, close-in, insider, and distribution [42] (Figure 4.3

The danger of a passive attack is a result of the surreptitious way information is gathered. It is the easiest type of attack to execute, and the hardest to defend against. Since the attacker is not actively transmitting or disturbing the transmitted signal of the signal

Table 4.3: IATF Attack Class Definitions

Attack Type	Definition
Passive	Passive attacks include traffic analysis, monitoring of unprotected communications, decrypting weakly encrypted traffic, and capture of authentication information. Passive intercept of network operations can give adversaries indications and warnings of impending actions. Passive attacks can result in disclosure of information or data files to an attacker without the consent or knowledge of the use
Active	Active attacks include attempts to circumvent or break protection features, introduce malicious code, or steal or modify information. Active attacks can result in the disclosure or dissemination of data files, denial of service, or modification of data.
Close-In	Close-in attack consists of individuals gaining close physical proximity to networks, systems, or facilities for the purpose of modifying, gathering, or denying access to information. Close physical proximity is achieved through surreptitious entry, open access, or both
Insider	Insider attacks can be malicious or non-malicious. Malicious insiders intentionally eavesdrop, steal or damage information, use information in a fraudulent manner, or deny access to other authorized users. Non-malicious attacks typically result from carelessness, lack of knowledge, or intentional circumvention of security for benign reasons
Distribution	Distribution attacks focus on the malicious modification of hardware or software at the factory or during distribution. These attacks can introduce malicious code into a product, such as a back door to gain unauthorized access to information or a system function at a later date.

owner, the signal owner (defender) has no means of knowing that their transmission has been intercepted. This kind of attack is particularly easy for two reasons: 1) frequently confidentiality features of wireless technology are not even enabled, and 2) because of the numerous vulnerabilities in the wireless technology security, determined adversaries can compromise the system.

Active attacks that can be launched against a wireless network come from a broad continuum. In its simplest form, active attacks use some mechanism disabling the entire communications channel between the sender and the receiver. With the original sender and receiver unable to recognize transmissions between each other, they cannot exchange information, and are unable to communicate. No detailed knowledge of the message parameters between sender and receiver is required, only a device capable of blocking communications operating over the entire channel.

More sophisticated forms of active attack are the Denial of Service (DOS) or Distributed Denial of Service (DDOS). The DOS and the DDOS differ primarily in the location of the origin of the attacks. The DOS originates from only one location., the DDOS from multiple locations. The specific mechanisms of a DOS and DDOS are very communications protocol and product implementation dependent, since these attacks exploit weaknesses in both the communications protocol and the products implementation of the protocol.

Other active attacks are based on exploitation attempts associated with the sender (identity theft, where an unauthorized user adopts the identity of a valid sender), weakness associated with the receiver (malicious association, where unsuspecting sender is tricked into believing that a communications session has been established with a valid receiver,), or weaknesses associated with the communications path (man in the middle, where the attacker emulates the authorized receiver for the sender- the malicious assertion, and emulates the authorized transmitter for the authorized sender- identity theft.). These attacks are primarily geared

at disrupting integrity in the form of user authentication (assurance the parties are who they say they are), data origin authentication (assurance the data came from where it said it did), and data integrity (assurance that the data has not been changed).

Close-In, Insider and Distribution Attacks describe the nature of system access, as opposed to the passive or active nature of the attack. Close-in, insider, and distribution attacks make use of some form of either an active or passive attack whose effectiveness is enhanced by the degree of the attackers access to the system. Insider and distribution attackers usually will utilize their specialized knowledge or access to carry out some form of a passive or active attack.

4.2 Attack Mitigation

The basic security mitigations for information and information processing systems attacks in the United States have been codified [43]. Specifically these are confidentiality, integrity, and availability. Confidentiality is concerned with ensuring that the data and system are not disclosed to unauthorized individuals, processes, or systems. Integrity ensures that data is preserved in regard to its meaning, completeness, consistency, intended use, and correlation to its representation. Availability assures that there is timely and uninterrupted access to the information and the system.

Closely related to these three are authenticity, accountability, and identification. Authenticity is the ability to verify that a user or process that is attempting to access information or a service is who they claim to be. Accountability enables events to be recreated and traced to entities responsible for their actions. Authenticity and accountability require the ability to identify a particular entity or process uniquely, as well as the authorizations (privileges) that are assigned to that entity. Identification is the specification of a unique identifier to each user or process.

The preferred mitigation methods for passive attacks are access control and confidentiality. Access control mechanisms are used to prevent unauthorized users accessing services and resources for which they have not been granted permission and privileges as specified by a security policy. Confidentiality should prevent the gain of information about from the content of the messages exchanged. Mitigation methods against active attack include access control, availability, accountability, authentication, and integrity. The access control and availability countermeasures must maintain or improve data availability. The system must be able to ensure the availability of both data and services to all components in the system. In the event that a PTC platform cannot handle its computational and communication load, it must provide graceful degradation of services and notify the operator that it can no longer provide the level and quality of service expected to prevent an unintentional denial of service.

The use of a Cyclic Redundancy Codes (CRC) is sometimes claimed as a means of providing data integrity. A CRC does not provide protection against malicious errors. This is because of the CRC many to one relationship between its input and output. It is possible for multiple inputs to check sum to a single CRC value. As a result, a data substitution can be made, with a correct CRC, and remain undetected. Consequently ensuring message integrity in a hostile environment requires the use of a Message Authentication Codes (MAC) or their associated Hashed Message Authentication Code (HMAC), Unlike a CRC, a MAC or HMAC establishes a unique one to one relationship between input and output, where each data input generates a unique MAC or HMAC. Any change in the input results in a change in the MAC or HMAC, which is detected at the receiver when the MAC or HMAC calculation is carried out and the received MAC or HMAC does not correspond to the calculated MAC or HMAC value [44].

Authentication mechanisms provide accountability for user actions. User authentication and data origin authentication differ in that user authentication involves corroboration of

the identity of the originator in real time, while data origin authentication involves corroboration of the source of the data (and provides no timeliness guarantees). User authentication methods range from so called time invariant weak authentication methods such as simple passwords to time variant strong cryptographically based authentication methods. In non-hostile environments no or weak user authentication may be acceptable, while in hostile environments strong user authentication is essential to provide authenticity. Data origin authentication provides assurances regarding both integrity and authentication. They rely on the use of digital signatures and can be either symmetrical or asymmetrical digital signature methods.

Ensuring integrity, authentication, and confidentiality, places restraints on availability and they have performance costs. Signing and or encrypting messages in transit may impose unacceptable delays in environments where near real-time response is required. These restrictions must be carefully considered in the development of any mitigation framework. more critically, they require a trust management system to exchange and control the necessary keying material for the system to work.

Chapter 5: TRUST MANAGEMENT

The successful application of mitigations for communication vulnerabilities that could be exploited by mal-actors must be done in the framework of trust relations between actors. This chapter addresses the Use and Misuse cases to capture trust relations, outlines how one such system can be constructed using Over the Air Rekeying (OTAR), and illustrates how its performance may be estimated.

5.1 Use and Misuse Cases

Use Cases, a construct of the Unified Modeling Language (UML) [46], are commonly used to capture systems requirements. They specify the interactions that the system is supposed to carry out with the external entities (so called actors). Consequently, a system that is developed according to the Use Cases provides all those interactions. Naturally, Use Cases assume that the actors interacting with the system are not acting with malicious intent.

Such an assumption does not adequately reflect the present world, where systems are often mis-used by malicious entities, called mal-actors, who are intent on disrupting the intended use of the system. Consequently, the design process needs to ensure that the constructed systems are not defeated by mal-actors. In order to formalize the various attacks that can be undertaken, a UML construct similar to Use Cases known as Misuse Cases has been proposed. First introduced by Sindh and Ophah [47], Misuse Cases provide a UML mechanism for capturing the relationships between actors and mal-actors and their impact on the behaviors of the system. As first defined, Misuse Cases add two additional constructs to UML: "prevents" and "detects". In the "prevents" relationship, the functionality of a

Misuse Case or mal-actor prohibits the execution of the functionality of the related Use Case. In the "detects" relationship, the construct reflects the functionality of a Use Case discovering a Misuse Case or mal-actor functionality capable of preventing the execution of the functionality of the Use Case.

These relationships were further extended by Alexander [48] to include "threatens", "mitigates", "aggravates", and "conflicts with", resulting in increasing the descriptive power of Misuse Cases. In the "threatens" relationship, the functionality of the Use Case is not eliminated by the actions of the mal-actor, as is the case with "prevents", rather is placed in jeopardy from executing properly. In the "mitigates" relationship, the mitigating Use Case counters the action of a mal-actor or another Misuse Case. Its functional inverse, "aggravates", the aggravating misuse case worsens the adverse impact of an existing adversarial relationship. Finally the "conflicts with" relationship states the mutual exclusivity of two of the related use cases.

An alternative formalization to capture adverse influences has been proposed by McDermott and Fox [49]. Known as Abuse Cases, they also state real world behaviors of systems subjected to adverse influences. They differ from Misuse Case models in two ways however. First, Abuse Cases are not shown on a Use Case diagram and Use Cases are not shown on an Abuse Case diagram, where Misuse Cases are blended with the Use Case models, forming an integrated statement. Second, Abuse Cases, by definition, are limited to eliciting security requirements, unlike Misuse Cases, that are intended to elicit any negative relationships. However, sans intent, Misuse Cases and Abuse Case are synonymous.

Both Use and Misuse Cases use a common graphical notation. Actors and mal-actors are represented by stick figures with a short descriptive name. Unlike the actor, however, the mal-actors head is solid black, as opposed to clear. The use of colors to assist in the

interpretation of UML objects is not with precedent. Couad et al [50] suggested pink, yellow, blue, and green to represent time dependent, role dependent, simple descriptions, and other elements. The Use Case and Misuse Case are represented by an ellipse with a short name that contains an active verb and noun phrase either in, or below, the ellipse, followed optionally by a list of properties, with Use Cses ellipses being clear, and Misuse Cases being black, The normal associations between actors and Use Cases, Use Cases and Use Cases, mal-actors and Misuse Cases, and Misuse Cases and Misuse Cases follow standard UML 2.0 notation and are joined by solid lines. The relationship of mal-actors and Misuse Cases follows a similar pattern to that of actors and Use Cases, just as actors are external entities to the system and interact only with Use Cases, mal-actors are external entities to the system and only interact with Misuse Cases.

While compelling, graphical representations of Use Cases, Misuse Cases and their associations have a significant drawback that is caused by not capturing complex nuances and details of textual descriptions. Consequently textual descriptions of Use Cases and Misuse Cases are necessary, and these must specify details such as sequential behavior within a binary relationship, parallel behavior between binary relationships, etc. Different templates have been proposed for use case descriptions, for example [51–54]. Each of these recommends various styles, content, and formats, along with different approaches for developing the material to be specified. The templates for Misuse Cases have been well developed and are adaptable to either the causal development approach [51] or the "facade" development approach [53]. The textual fields in the templats are generally self-explanatory and are similar for use and misuse cases.

Critical trait attributes are captured from the textual misuse cases from an analysis of nouns and verbs by a technique called noun-verb extraction [55]. By using noun-verb extraction analysis, specific characteristics that could represent evidence (i.e. behaviors that

may be directly observed or conclusively inferred from observed behaviors) or invariant elements are identified. This process can be done entirely by hand by an engineer, or through the aid of tools [56], When both the textual and graphical notations are used together, they capture the requirements for mitigation of communications vulnerabilities.

PTC Misuse Cases can be classified into primary (platform independent) and secondary (platform dependent) misuses. Authentication and authorization of PTC inter-operation can be misused by claiming invalid trackage rights or impersonation. Crews or locomotives claiming to be from railroad A or railroad B respectively, may not be authentic, and the combination may not be authorized to enter a specific segment of company C's tracks. The secondary misuses exploit the vulnerabilities of the underlying (wireless and wired) communication infrastructure and protocols. Authorization in a distributed system is quite different from that in centralized systems. Trust management systems unify the notion of security policy, credentials, access control, and authorization, allowing direct authorization of security critical actions in a distributed environment. Blaze [57] provides a general discussion of the use of trust management in distributed system security. Our distributed trust management system framework with online key exchanges prevents the primary misuses and detects the secondary Misuse Case, yet enables the central Use Case. The framework consists of certificates, policies, and distribution protocols. We approximate a calculation of the security overhead that is acceptable in ensuring the safety required for railway interoperation.

5.2 Trust Management for PTC Systems

The detailed design and implementation of a secure Trust Management System (TMS) is a complex and difficult process that is best based on a philosophy of risk management for the

life cycle of the system at a number of levels. This design process must address not only the selection of appropriate standard cryptographic protocols at the device level, but also the selection and management of the keying material as well as the supporting operational and management infrastructure. Inappropriate decisions will result in little or no security. Even if appropriate decisions are made, without the support of senior management, their execution may fail. Our proposed TMS illustrates a distributed authentication and authorization scheme with the associated operational implications and is based on guidance provided in the National Institute for Standards and Technology (NIST) Federal Information Processing Standard (FIPS) and NIST Special Publications (SP).

The distributed authentication and authorization TMS has three major components: certificates, policy decisions, and distribution protocols [45] each with associated operational implications.

5.2.1 Certificates

A certificate is a statement issued by a trusted authority, called a Certificate Authority (CA), stating that selected attributes have specific values. Each service requestor must present the certificate to obtain service and/or furnish proof of authenticity. Prior to granting the service, the service provider is expected to verify the certificate"s authenticity and validity with the issuer. This requires every user to belong to a *trust domain* represented by a certificate authority (CA). The key management certificates specify the details about the keys used to (en/de)-crypt information, key validity periods, and their revalidation procedure. Other policies and service level agreements (SLAs) are encoded as values of named attributes in certificates. These are digitally signed by CAs to enforce integrity and are cross-checked prior to the service being granted. Similarly, all requests and responses are signed to evade alteration and injection of spurious messages.

Company A's crew, operating in company B's locomotive seeking to enter company C's tracks, must present C's entry checkpoints with certificates of service entitlements and authenticity for A's crew (issued by A's CA) and certificates of B's locomotive's authenticity (issued by B's CA). The CA of C must validate the certificate with respective issuers. If the entities are authenticated and certificates validated, trackage rights are granted. This solution enables the central Use Case while evading the impersonation Misuse Case.

We propose using X.509 public key and attribute certificates [61], binding the subject's public key and privileges to their identity via X.509 signature (DSS) and X.509 key management (KEA) certificates. X.509 certificates are based on the ITU-T X.500 standard, and are the most widely used standard for digital certificates. The X.509 attributes pertinent to enforcing secure PTC inter-operation are as follows:

[Version (Type: integer)] Purpose: enable parsing

[Serial Number (Type: integer)] Purpose: Cert. ID

[Signature Algorithm (Type: integer, algorithm name)] Purpose: identify algorithm used for signing and hashing

[Issuer (Type: text)] Purpose: Issuer ID

[Validity Period (Type: time interval)] Purpose: (begin, ending) time of certificate validity

[Subject Name (Type:text)] Purpose: Holder ID

[Name (Type: integer)] Purpose: User ID

[Public Key Info (Type: integer/text)] Purpose: Public Key and Extensions

At the time the security service is designed, the issued certificates contain authorizations, such as those issued to engineer Casey Jones shown in Figure 5.1. Figure 5.1 shows a sample certificate identifying its holder as engineer Casey Jones of the CSX Railroad, stating that BNSF CA will verify his identity for the period 1 July 2006 to 7 July 2007. The pair `(Issuer, Serial Number)` provides a system-wide unique ID for a certificate.

Service Level Agreements (SLA) would contain trackage rights negotiated and approved prior to creating the certificates by the Surface Transportation Board of the US Department of Transportation under 49 USC 11323. The SLA's are then encoded in the X.509 V3 certificate extensions. In the example of Figure 5.1, Engineer Jones was issued an X.509 version 3 certificate. The extensions authorize him to operate trains owned by the UP or BNSF railroads (train symbols 1234, W3F4, and TY65) on the Anna and Bess blocks of the BNSF Beardstown subdivision, and only from 24 August to 15 October 2006. This certificate is given to the train dispatching system of a railroad, where its recipient must verify the certificate validity with BNSF's CA.

5.2.2 Policy Decisions

In addition to certificates, the trust management system behaves as specified in meta-policies. While a full discussion of these is beyond the scope of this paper, a minimal overview of the following must be addressed by the trust management system:

Admission to a trust domain:

- Procedures for physical identity validation
- Physical identity-to-key binding
- Registration attributes

Certificate:
Data:
Version: 3
Serial Number: 1
Signature Algorithm: SHA1withRSAEncryption
Issuer: C=US, ST=KS, O=BNSF OU=Signal, CN=CA
Validity: Not Before 1 July 2006 0001Z
Not After 1 July 2007 0000Z
Subject Name; C=US, ST=KS, O=CSX, OU=OPS, CN=Casey Jones
Public Key Info
Algorithm: RSA Encryption
RSA Public Key: <Public Key>
X.509 v3 Extensions
Basic Constraints: Critical
Basic Constraints Shared Authorization
Train Owner: C= US ST= KS O= BNSF; C= US, ST=NE, O=UP
Train ID: OU=OPS, CN=1234; CN=W3F4; TY65
Track Block: Anna, Bess Subdivision: Beardstown RR: BNSF
Time: Not Before 1200Z 24 Aug 2006
Not After 1200Z15 Oct 2006
Signature Algorithm:

Figure 5.1: X.509 Certificate

Initializing cryptographic material:

- Key generator and cryptographic algorithms
- Configuring algorithmic preferences
- Definition of domain and trusted parameters
- Identifying trusted parties

Key management:

- Installation, establishment, key generation and distribution of active keying material
- Key update process, transporting initial keys and crypto periods
- Archival and recovery of storage and protection
- Accountability: Access requirements to keying material generating-distributing-destroying-and conducting audits

Survivability of Critical Infrastructure:

- Continuity of operations plans
- Backup and recovery mechanisms for trust breaches and system-wide failures.

The following illustrates proposed policy decisions for the PTC trust management infrastructure:

Admission and Authentication:
A role-based access control (RBAC) policy can be used for human accesses. A subset of biometric information, passwords and PINs may be a used as identifying attributes. Devices and systems such as Onboard, Wayside, Office/Dispatch can be bound to their public keys with a certificate containing a hardware level address such as the MAC address.

Initializing Cryptographic Material:

The PTC cryptographic material uses specific key generation algorithms for private/public key pairs, and seeds for symmetric keys. Algorithms such as RSA, AES, and SHA1 can be used for encryption and hashing operations.

Key Management:

The key management policy uses a PKI and communicates original keys by means of an independent secure channel. Archival and recovery procedures can require a user to replace elapsed certificates with current ones.

Survivability:

In the event of a disaster, CA operations are first re-established. If the CA is physically damaged, copies of signatures, keys and certificates are replaced.

5.2.3 Distribution Protocol

The operation of a TMS requires a mechanism for distribution for cryptographic keying material. The keying material distribution mechansism requires either direct physical contact with each device being keyed, or a mechanism for doing so remotely. The former is practical for small numbers of geographically concentrated devices infrequently keyed, but does not lend itself to large numbers of geographically distributed devices that must be frequently keyed as is the case in the rail environment. Since railroads operate a large number of trains that are geographically dispersed and on the move (seven Class 1 railroads alone operate almost 24,000 locomotives [7]), manually rekeying just the locomotives by hand is not practical. Add the wayside equipment that must be rekeyed that is spread out over 145,000 miles of track, and the situation becomes untenable. To overcome this situation, we propose that the rekeying be done remotely using Over-the-Air Rekeying (OTAR) [70].

OTAR is a suite of protocols (first used by the US Department of Defense [58]) for remote

key management, renewal, and distribution, including key changeovers sent from the Key Management Facility/Key Distribution Center (KMF/KDC) to end units. OTAR allows remote rekeying to be done at a regular interval (referred to as Crypto periods), eliminating the need for an individual to physically interact with the device that must be rekeyed. The railway scenario is similar to that of the US Department of Defense: very large numbers of geographically dispersed elements that required periodic rekeying. Using a combination of Traffic Encryption Keys (TEKs) to hash data, and Key Encryption Keys (KEKs) to encrypt and exchange other KEKs or TEKs, OTAR algorithm allows for the transmission of secure information over an insecure communications path.

The use of OTAR, while providing for remote management of geographically disburse keys, introduce some operational issues of its own that must be considered by the system designer. Implementing the OTAR requires the purchase of specialized infrastructure components to support key generation, distribution, and management. This may involve significant capital investment, the scope of which depends on the capabilities of the existing communications infrastructure. OTAR does not eliminate the need for creation and implementation of key management policies within, and between railroads. Organizations must establish appropriate operation and personnel polices to ensure protection of the keying material. Agreements between railroads must be made to control their own or shared keys and define their respective responsibilities. OTAR also does not eliminate the need to provide end-to-end communications between the communicating entities (office/dispatch, onboard, wayside).

OTAR also still requires establishing Key Management Facilities (KMF), storage and accountability for active and contingency cryptographic keying material both at the KMF and the field units as well as capabilities for addressing key compromise or revocations. OTAR also does not eliminate the need for training of the personnel using the system. While the use of OTAR can reduce the level of effort required to address these issues, it still does not eliminate them entirely.

5.3 Secure PTC Interoperation

Authentication and/or Authorization are fundamental in a TMS creating safe and secure PTC inter-operation. Interoperability is the ability of the PTC subsystems associated with one railroad to exchange information with the PTC subsystems of another railroad in such a way that the change from one system to another is transparent to the end users. Accomplishing this objective requires that the subsystems share common data formats, protocols, and interfaces to allow for the exchange of data as well as common interpretations of how the data is to be interpeted to provide for an exchange of information.

Although conceptually simple, in practice implementing secure interoperability is non-trival. There are many good reasons for a vendor (or vendors) to limit interoperability, such as ensuring or increasing market share or limiting or eliminating competition. In the case of the railroad industry, there are mutually competing PTC vendors attempting to market their products to the oligopoly formed by the railroad companies. Even within the oligopoly formed by the railroads, the individual railroads act out of self-interest to define interoperability in terms of systems they have already procured as a standard for use on their railroad, requiring other railroads to make additional capital expenditures.

We will limit our discussion of interoperability to technical requirements, leaving the political and economic implications as separate issues outside the scope of this document.

5.3.1 Authentication

At a minimum, non-moving wayside equipment needs to periodically authenticate with their domain controller. This provides a check of communications connectivity between the wayside device and the domain controller along with assurance of the identity of each communicating party. This process can be extended to include the health information messages of the wayside equipment. The wayside device can be designed to push health

status information when a change occurs, the domain controller can be designed to periodically pull health information on demand, or a combination of the two may be implemented.

We select the Needham-Schroeder-Lowe [63] public key protocol as representative of an authentication protocol that can support the preceding communications connectivity and identity validation. It reduces the problem of key explosion, reduces the number of required shared key pairs, and the brevity makes it ideal for situations of limited communications bandwidth devices with restricted computational power that have very infrequent connectivity requirements, as might be encountered with wayside equipment located in remote areas. One is not limited, however, to the use of the Needham-Schroeder-Lowe protocol. Other protocols may be equally acceptable depending on the communications and equipment capabilities of the entities involved.

The Trusted Management process assumes that the wayside device (W) and the domain controller (DC) share a common trusted source (Certificate Authority-CA) and that this CA maintains a single Certificate Revocation List (CRL) (Figure 5.2). The process begins when W requests the public key of DC (Pub_{DC}) from the CA. The CA checks the CRL to ensure that W and DC's certificates and associated keys have not been revoked. If the certificates and keys are not on the CRL, the CA sends the public key Pub_{DC} of the DC to W by encrypting the DC's public key (Pub_{DC}) with a common shared key between W and the CA (K_{CA-W}). The notational description of the identity of the domain DC and its public key encrypted by the common shared key is (DC,Pub_{DC}):(K_{CA-W}). (DC,Pub_{DC}):(K_{CA-W}) is returned to W. W recovers the public key of the domain controller, allowing W to communicate directly to DC. W sends a message to the domain controller consisting of a Nonce N_1, and its identity encrypted with the public key of the domain controller. Notationally this is represented by (N_1, DC): (Pub_{DC}).

To complete the process, W and the DC must mutually authenticate each other and verify

communications connectivity. The Domain Controller decrypts the message from W by using its private key (Pri_{DC}). Next, DC queries the CA for W's public keys, and the CA returns the identity of W and the public key of W encrypted with a shared common key between DC and the CA ((W,Pub_W): K_{W-CA}). Using the public key of W, the DC encrypts a message back to W consisting of the nonce N_1 received from W, a nonce N_2 generated by the DC, and the DC's identity (N_1, N_2, DC):(Pub_W). Finally, upon receipt of this message by W, W decrypts the message using the private key to recover three items: N_1, the identity DC, and nonce N_2. W validates the received N_1 matches the transmitted N_1. If they match, W replies to DC with a message containing its identity, the received nonce N_2 encrypted with the public key of DC (W, N_2, Pub_{DC}). The DC verifies the transmitted nonce N_2, and the received nonce N_2 match, thus authenticating each other.

Next, a security policy that allows unrestricted use of the cached keys for a period of time (the crypto-period) must be implemented. After this time period, W and the DC must repeat the process of re-polling the CA for the public key and verifying that the certificates and keys have not been compromised or changed. The use of cached certificates reduces the communications bandwidth required, an important consideration, as previously noted, in bandwidth or power constrained environments.

This authentication scheme could be extended to allow the exchange of symmetric session keys for subsequent communications checks. By doing so, and caching the resultant session symmetric session keys, subsequent communications checks could be accomplished significantly faster, with lower overhead costs in terms of both computational complexity and communications bandwidths. This is in fact, crucial where frequent message are exchanged in the high speed, high traffic density environments in which PTC systems would provide their optimal safety benefits.

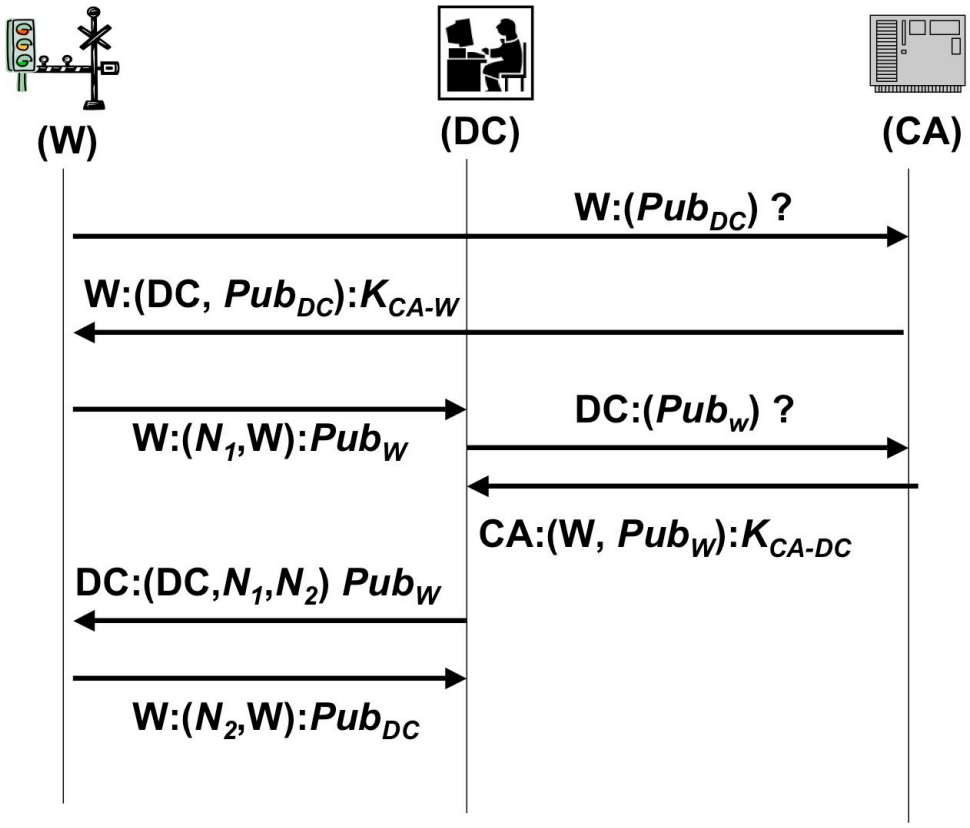

Figure 5.2: Wayside-Domain Controller Authentication

5.3.2 Authorization

Authentication alone, while satisfactory for checking a communications path, does not consider the exchange of authorizations or other informational data. A typical example is the previously identified Central Use Case. Not only must the identity of the entities communicating be clearly established, authorization for the entities for a service must be requested and received. For example, in the scenario postulated where engineers from company A approach an entry point belonging to company C, A and C may have mutually authenticated but C may have a policy that does not allow A to pass beyond the entry point C.

We consider a scenario that reflects current cross-domain railroad dispatching practices between two different railroads. In order to facilitate cross-domain dispatching today, the dispatchers of one company accept the dispatchers of the other company as the authoritative source regarding train movements in their respective domains. Assume that Company A's crew only has access to their trusted source, Certificate Authority A (CA-A) and Company C's entry point has only access to their trusted source, Certificate Authority C (CA-C). Company A and Company C have implemented a security policy where by information exchanged between CA-A and CA-C is trusted. If CA-A receives a request for information associated with Company C, it will pass the request onto CA-C, and treat CA- C's response as authoritative. Similarly, CA-C treats information from CA-A as authoritative. Finally we assume that CA-A and CA-C maintain a CRL only for members of their respective domains.

Figures 5.3, 5.4, and 5.5 illustrate an integrated authentication and authorization process using local access control. To enable A to communicate with entry point C. A's crew requests C's public key (Pub_C) from its trusted source CA-A, who first ensures that A's certificates and associated keys have not been revoked. CA-A then queries CA-C for the public key of C. If CA-C verifies C's certificates are still valid, CA-C will pass the public key of C (Pub_C) to CA-A. CA-A encrypts company C's entry point identity (C) and C's public

key (Pub_C) using a common shared key between A and CA-A (K_{A-CA}). The encrypted information (C, Pub_C): (K_{A-CA_A}) is retuned to A, who recovers the public key of C from the message received from CA-A; this allows direct authorized communication from A to C.

The communication proceeds by A sending a message to C's entry point consisting of a nonce N_1, and A's identity, encrypted with the public key of C (N_1, A): (Pub_C). Upon receipt of the message from A, C decrypts the message using its private key (Pri_C).

Communications from C to A requires C to obtain A's public key. Just as A queried CA-A for C's public keys and the request was passed to CA-C, C queries CA-C, to determine the public key of A and the request is passed to CA-A. CA-A returns the public key of A (A, Pub_A) to CA-C. CA-C then encrypts the identity of A along with the public key of A using a shared common key between C and CA-C ((A,Pub_A):K_{C-CA_C}) and returns it to C. After obtaining the public key of A, C may communicate directly with A.

Finally, A and C must mutually authenticate each other. Using the public key of A, C encrypts a message (N_1, N_2,C: Pub_A) back to A consisting of the nonce N_1 received from A, a nonce N_2 generated by C, and C's identity. On receipt, A decrypts the message using its private key, recovering N_1, the identity C, and C's nonce N_2. Finally A replies to C with a message (A,N_2):(Pub_C) containing A's identity and the received nonce N_2 encrypted with the public key of C (Pub_C). Upon receipt of this message by C, and C's verification of N_2, A and C have mutually authenticated each other.

After mutual authentication, A and C continue to establish a common shared symmetrical key with each other. As indicated earlier, the importance of symmetrical caching keying schemes increases significantly when the frequency of communications increases, the periodicity between communications decreases, and the timeliness of response is essential. While

asymmetrical keying scheme could be used, their increased computational complexity as compared to symmetrical keying schemes [62] requires more powerful processing equipment (with its associated increased power requirements) to minimize encryption and decryption times. Asymmetrical encryption is at least two orders of magnitude slower than equivalent symmetric schemes of equivalent.

The symmetrical key establishment process begins when A prepares a message to C consisting of A's identity, A's component of a common shard key k_A, a nonce N_1, and a CRC calculated over A's identity, N_1, and the key k_A. This information is encrypted using the private key of A, then encrypted again using the public key of C. The doubly encrypted message is then transmitted to C, who decrypts it to reveal the identity of A, N_1, k_A and the CRC. C then verifies that the CRC is correct, and if it is correct, the process is reversed. C prepares a message to A consisting of C's identity, C's component of a common shared key k_C , a new nonce N_2,the received nonce N_1 and a CRC over the key component, the nonces, and C's identity. C's component of the shared key, the nonce's N_1 and N_2, and the CRC are first encrypted using the private key of C, then encrypted again using the public key of A. Upon receipt of this message from C, A decrypts the received message, first using A's private key, and then again using C's public key.

This reveals the CRC, k_C, N_1 as returned from C, and C's new nonce N_2. After first verifying the CRC is correct, and the N_1 received from C is the same as the nonce N_1 sent by A, A concatenates k_A and k_C to form a common shared key k_{AC} with C. A returns its identity, the received nonce N_2, and a CRC to C. This information is first encrypted using the private key of A and then again using the public key of C. Upon receipt, C first decrypts the received message using C's private key, then again using A's public key, revealing the received N_2 and CRC. C compares the received N_2 with the previously transmitted N_2 and validates the correctness of the CRC. If N_2 received by C is the same as N_2 first transmitted by C, C concatenates k_C and k_A to form the common shared key k_{AC}. Once a common

shared symmetrical key is established it may be cached allowing A and C to exchange service messages directly with each other until the common key k_{AC} expires or is otherwise voided.

After establishing the common key, service requests are processed using the following process. When requesting a service, the service requester (in the example Company A's crew) prepares a service request message R. The service message specifies the service requested, and includes a copy of A's certificate containing A's service authorizations, a message number, and a time of transmission. The service request is hashed using cryptographic function F with the common shared key K_{AC}. This yields a MAC *(*M_{AC}. The MAC is then encrypted using A's private key. The encrypted MAC is appended to R and the result transmitted to C. Upon receipt, C divides the message into the received service request R and the encrypted MAC. C first decrypts the encrypted MAC using the public key of A to ensure the message came from service requester A . The received message R is then cryptographically hashed by C using the same function F and the common key K_{AC}. The resulting MAC generated by C is compared with the received MAC. If they are identical, C accepts the service request from A as unchanged and authentic.

After verifying the service request, C examines the service request R to determine if the service requested by A is an element of the set of A's authorizations $a_1, a_2, \ldots, a_N$ and is also an element of the services *(*$c_1, c_2, \ldots, c_N$*)* that C is authorized to provide. C queries CA-C for a copy of A's certificate. Assuming CA-C has not cached the certificate, CA-C passes the request to CA-A. CA-A checks the certificate request against the CRL, and if not revoked returns a copy of A's certificate to CA-C. The certificate is then passed from CA-C to C. C validates that the certificate received from A and the certificate received from CA-C match. If the certificates match, the requested service is an element of both A's authorized services and C available services, C provides the service to A. The set of services ($c_1, c_2, \ldots, c_N$) that C provides to a requester depends upon the security policy

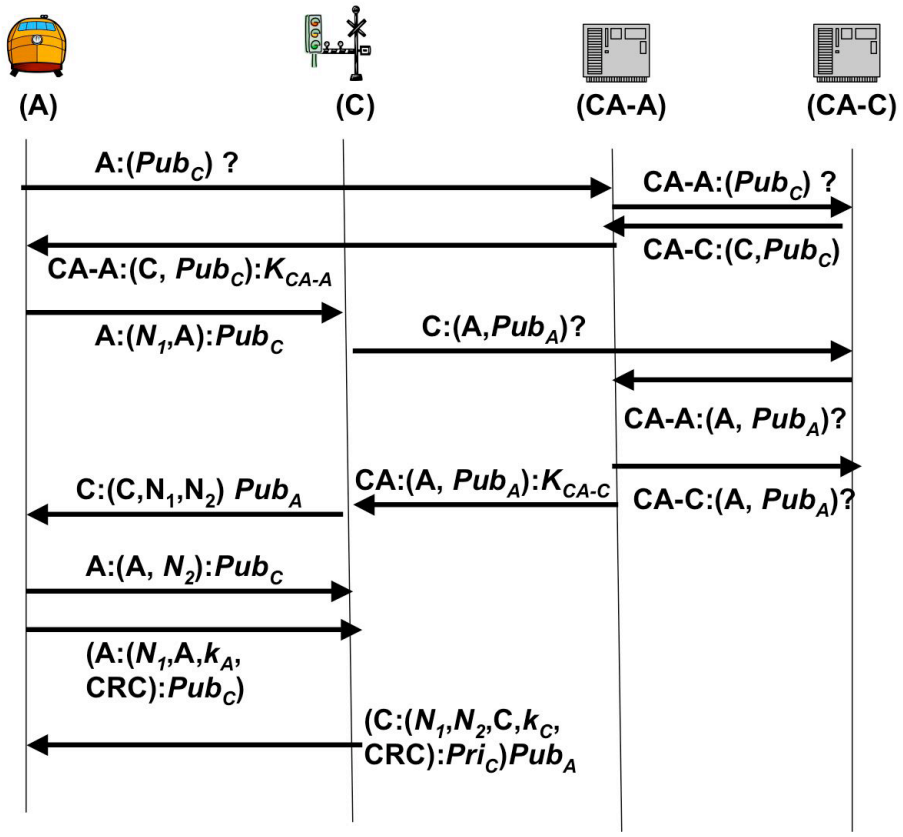

Figure 5.3: Local Access Control Decision Function-Key Exchange

of the system and the individual capabilities of the particular C. A similar process can be repeated in the case of Company B and C or Company A and B exchanging service requests.

A distributed authorization process extends the described local authorization process. When a Service Level Agreement (SLA) between A and C exists, C must further determine if the requested service R is covered by that SLA. If A's requested service R is covered within C's capabilities, then C validates the SLA. If this is successful, then C provides the service R to A.

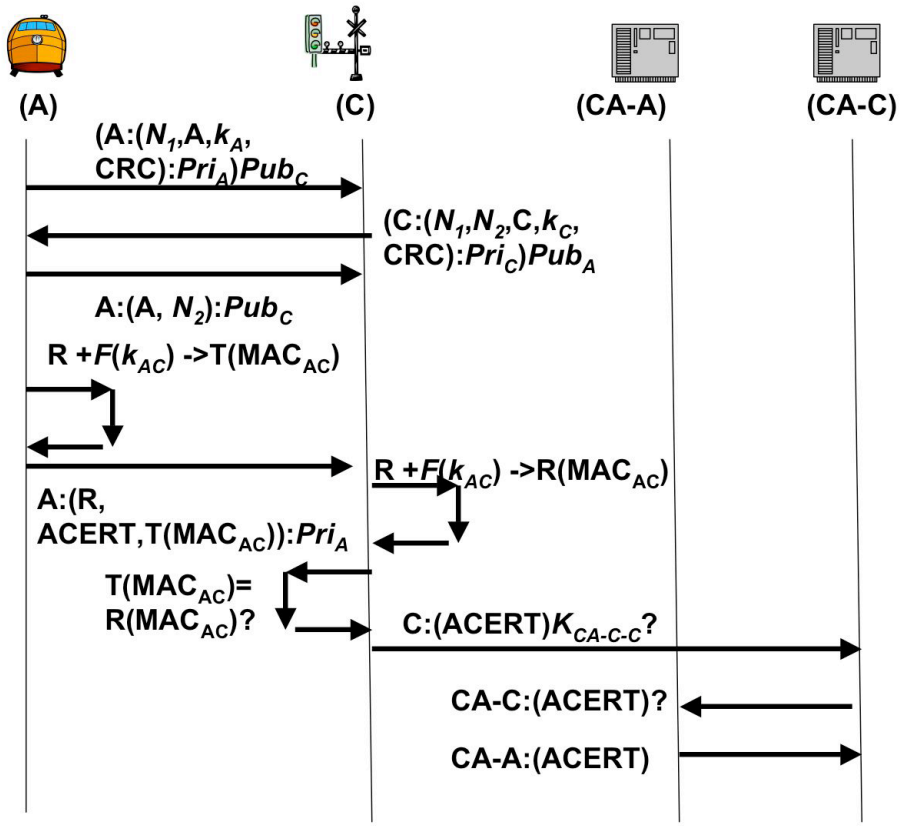

Figure 5.4: Local Access Control Decision Function-Service Request Presentation

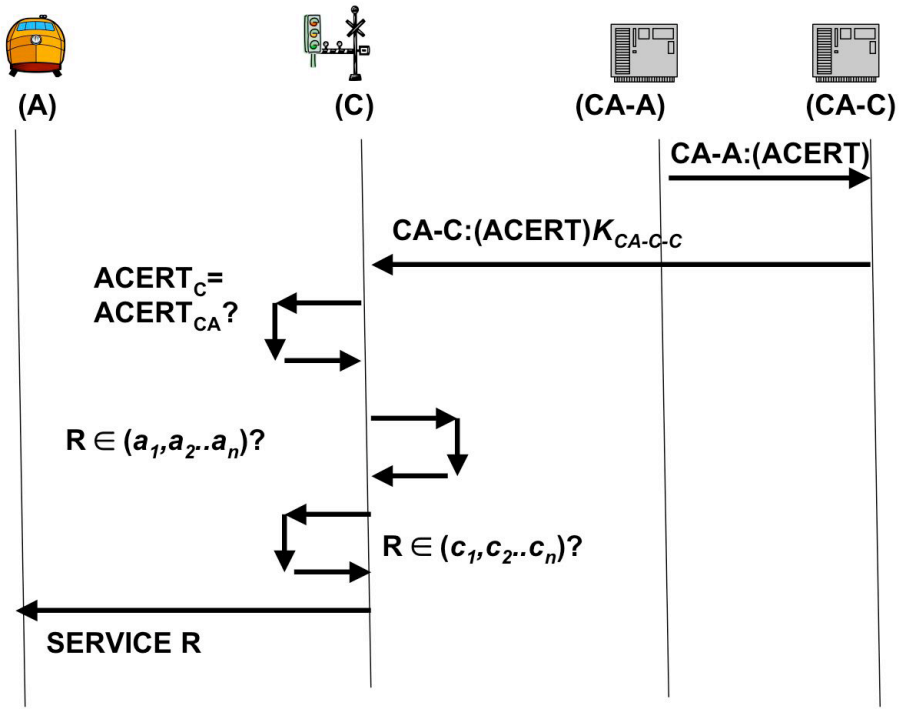

Figure 5.5: Local Access Control Decision Function-Service Request Service

A trust management system with the authentication, authorization and validation of the SLA: (1) satisfy the central Use Case, (2) prevent the primary Misuse Case, and (3) detect the secondary Misuse Case. Message alteration is detectable (but not preventable) by receipt of an invalid MAC, while spurious messages can be detected by receipt of an invalid signature. Message drops (benign or otherwise) can be detected by including a message number in the service request that is hashed. The extent of a delay can be evaluated by including the time of transmission, comparing it against the time of receipt, but is not prevented with the proposed algorithm.

5.4 Performance

Adding a distributed trust management over PTC creates time overhead affecting the PTC session due to added challenge-response steps and space overhead due to expanded headers and padding in security-enhanced protocols. As a consequence, the system designer has to make appropriate tradeoffs [69] in meeting safety and security requirements specific to interoperating PTC systems, because per-packet delays increase when more nodes are added.

The overhead on an information exchange between two communicating entities is directly related to the communications protocol used, as well as any safety and security protocol. In its most abstract form, the percent overhead (P_{OH}) can be calculated by the equation :

$$P_{OH} = (\frac{I}{I + OH})(100) \tag{5.1}$$

where

I = The number of information bytes exchanged

OH = The number of additional(overhead) bytes appended to the information bytes by the communications protocol

The information, I, represents a data pattern exchanged between a sender and a receiver that has meaning to both the parties involved. It conveys some fact or facts of interest to both of the communicating parties. The overhead, OH, also consists of data patterns. Unlike I, OH is directly related to the medium or media used to express it. The overhead can be determined once a understanding of the communications media and protocol has been obtained. The specific formulation of OH will vary widely depending the specfific protocols, security mechanisms, and transmission media. Abstractly, OH can be represented by the equation:

$$OH = B_{SenderAddress} + B_{ReceiverAddress} + P_{Information} + C_{Data} + C_{Padding} + S_{Data} + S_{Padding} \tag{5.2}$$

where

$B_{SenderAddress}$ is the number of bytes of information required to identify the sender

$B_{ReceiverAddress}$ is the number of bytes of information required to identify the receiver

$P_{Information}$ is the number of bytes of data required to properly format I for transmission across the media.

C_{Data} is the number of bytes of information required to control the transmission across the media.

$C_{Padding}$ is the number of bytes of data required to properly format C_{Data} for transmission across the media.

S_{Data} is the number of bytes required to convey any security information required to ensure the integrity, authenticity, and or confidentiality of I, $B_{ReceiverAddress}$, $B_{ReceiverAddress}$,

$P_{Information}$,

C_{Data}, or $C_{Padding}$

$S_{Padding}$ is the number of bytes of data required to properly format S_{Data} for transmission across the information.

The determination of what is I and what is OH is further complicated by the fact that most communication protocols are based on the concept of communicating peer layers between nodes. Each communicating peer layer utilizes a common shared protocol. However, communications between communicating peers is accomplished using the services of a lower level peer layer. The communications between a layer n, and a layer n-1, is by an interface. Just as peer-to-peer communications consist of information, with an associated overhead, interface communications between layers consist of information and overhead. What constitutes information and overhead data at one layer will differ from what constitutes information and overhead data at another layer. The exact numbers depend on the protocols involved. The various CBTC systems discussed in Chapter 3 have been developed using different underlying network architectures that have been selected and optimized based on each individual railroads' business and operating plans. The remainder of this chapter will utilize a specific protocol to illustrate the influence of time and space overhead on performance at the applications layer. A similar analysis would be required as part of the network design for each lower peer-to-peer layer, and for each specific CBTC system.

In addition to the communications protocol overhead, there are also the propogation and processing delays. With all forms of transmission medium there is a short time delay for the signal to travel (propagate) through the medium.

$$P_{PD} = (\frac{PS}{PV}) \tag{5.3}$$

where

PS = Physical Seperation

PV = Propagtion Velocity

and a processing delay P_C. P_C is the time it takes to process data at the tranmitter ans receiver.

5.4.1 Time Overhead due to Challenge-Response

The time overhead due to the challenge-response may potentially be significant yet can be managed. In a worst-case scenario, a PKI based schema may require re-authentication and re-establishment of a shared session key prior to every message exchange for services. This scenario could occur when there are extremely short crypto-periods for the shared session keys that preclude caching of the keys. It may also occur in situations where there is high speed, high-density traffic, and large numbers of wayside devices. In this case it is necessary to establish a large number of different session keys to support rapidly changing communications between the different trains, crews, and wayside devices.

In our example, the overhead communications for authentication and session key establishment between A and C requires 10 steps (Figure 5.3, 5.4 with an execution time = *timer* Table 5.1) exclusive of CA to CA communications. If the size of the data transfer between A and C is similar for each step, as the number of data messages exchange steps increases relative to the number of authentication and key session exchange steps, the overhead is increasingly amortized.

$$COH = \frac{A_{Exchange}}{N} \tag{5.4}$$

In the limit, as the number of data message exchange steps N approaches infinity, the overhead cost (COH), measured in overhead steps per data message steps, approaches 0. Conversely, if only 1 data message (4 operations) is exchanged before it is necessary to

Table 5.1: Communications Overhead Key Establishment

Step Number	Action	Execution Time
1	A:(Pub_C)	t_1
2	CA-A:(C,Pub_C):K_{CA-A}	t_2
3	A:(N_1,A):Pub_C	t_3
4	C:(A,Pub_A)	t_4
5	CA:(A,Pub_A):K_{CA-C}	t_5
6	C:(C,N_1,N_2):Pub_A	t_6
7	A:(A,N_2):Pub_C	t_7
8	A:(N_1,A,k_A,CRC):Pub_C	t_8
9	C:(N_1,N_2,C,k_C,CRC):Pri_C)Pub_A	t_9
10	A:(A,N_2):Pub_C	t_{10}

timer= $\sum t_i$, i = 1..10 See *timer* Section 6.4.4

re-authenticate and re-establish the session keys, the *COH* is 2.5. In practice more than 1 data message is exchanged before re-authentication and re-establishment of the session keys, so the *COH* is less than 2.5 and greater than 0.

5.4.2 Space Overhead

The additional network related protocol information to the information packets adds network protocol space overhead. A variety of proprietary communication networks are used in PTC. To illustrate the impact of space overhead, we use a commonly available, well-understood, non-proprietary protocol whose size increases are representative of those found in the proprietary protocols. For our representative example we will use IPSec [60]. This allows for an understanding of the impact of header size without revealing the underlying proprietary protocol.

The length of each IPSec header varies with the mode of operation, varying in length between from 24 bytes to 60 bytes. At one extreme IPSec in authentication transport mode

adds 24 bytes to the length of the data packet, while at the other extreme, IPSec Encapsulated Security Protocol (ESP) tunnel mode adds 60 bytes (a 20 byte tunneling header, an additional 8-byte ESP header, a 0 to 16-byte Initialization Vector (IV), and a 16-byte ESP Trailer). The resulting range of header overheads is shown in Table 5.2 and as a percentage of common data packet sizes using various Encryption and Signature Schemes in Table 5.4.

In addition to the IPSec Overhead, commonly used symmetrical encryption/decryption (3DES) [64], (AES) [65] and hashing (SHA) [66] algorithms use fixed size data blocks, and pad packets before encrypting or hashing them, further adding to the overhead. 3DES uses 64 bit blocks, and adds padding if the body of the packets is less than or not evenly divisible by 64. AES uses 128 bit blocks, and adds padding if the body of the packets is less than or not evenly divisible by 128. SHA-1 requires 512 bit blocks but has an effective length of 448 bits, and adds padding if the body of the packets is less than evenly divisible by 448 bits. SHA-1 also adds 8 bits for each 56 bit block of data.

The overall affect or overhead when encrypting packets with 3DES and sending them

Table 5.2: IP Sec Header Overhead

		AH Header-Trailer	**ESP Header-Trailer**
Mode	**IP Tunnel Header Size**	24	24-40 bytes
Tunnel	20 bytes	44 bytes	44-60 bytes
Transport	0 bytes	24 bytes	24-40 bytes

across an IPsec secured network link assuming SHA-1 hashing, ESP tunnel mode and an ESP IV of 8 Bytes can be calculated as follows. The TCP/IP headers are required regardless of whether IPsec is used or not, and can be ignored. To transmit 1 Byte of data;

- Add 7 Bytes for 3DES padding to reach the 8 Byte 3DES block size ($S_{Padding1}$)

- Add 48 Bytes for SHA-1 padding to reach the 56 Byte effective SHA-1 block size ($S_{Padding2}$)
- Add 8 Bytes for the SHA-1 message length information, reaching the 64 Byte actual SHA-1 block size ($S_{Padding3}$)
- Add 20 Bytes for the ESP tunnel mode header ($S_{Padding4}$)
- Add 8 Bytes for the ESP header ($S_{Padding5}$)
- Add 8 Bytes for the ESP IV ($S_{Padding6}$)
- Add 16 Byes for the ESP trailer ($S_{Padding7}$)
- Total packet size (minus TCP/IP headers) is: 116 Bytes

where

$$S_{Padding} = \sum S_{Padding-i}, i = 1..7 \tag{5.5}$$

Table 5.3 shows the padding overhead for different packet sizes due to encryption or hashing. When combined with the transport mode, the overhead percentage (to the nearest integer)

Table 5.3: Packet Size Increment due to Padding [67]

Packet Algorithm	**Block Size Bits**	**Small Packet 40 Bytes**	**Average Packet 350 Bytes**	**Large Packet 1500 Bytes**
3DES	64	0 bytes, 0%	2 bytes, 0.57%	4 bytes, 0.266%
AES	128	8 bytes, 20%	2 bytes, 0.57%	4 bytes, 0.26%
SHA1	512	16 + 8 bytes, 60%	42 + 56 bytes, 28%	12 + 216 bytes, 15.2%

is cited in Table 5.4. Each overhead is calculated as follows:

$$
\begin{aligned}
&OverheadPackets = \\
&\quad (IP_{TunnelHeaderSize}) + (AH_{HeaderTrailerSize}) \\
&\quad +(ESP_{HeaderTrailerSize}) \\
&\quad +(3DES_{BlockSizePadding}) + (AES_{BlockSizePadding}) \\
&\quad +(SHA1_{BlockSizePadding})
\end{aligned}
$$

$$Overhead_Percentage = (100)\frac{Overhead_Packets}{PacketSize} \tag{5.6}$$

As a worked example of the preceding, in the case of the configuration mode of Tunnel Mode ESP 3DES/SHA1 for a 40 byte data packet

$$OverheadPackets = (20) + (0) + (40) + (0) + (0) + (24) \tag{5.7}$$

$$\text{OverheadPackets} = 84$$

$$Overhead_Percentage = (100)\frac{84}{40} \tag{5.8}$$

Overhead$_{Percentage}$ = 210 Note the overhead calculations in Table 5.4 are associated with

Table 5.4: Total Header Overhead for Common Packet Sizes

	Packet Size- Bytes		
Configuration	**40**	**350**	**1500**
Transport Mode ESP 3DES/SHA1	160%	78%	18%
Transport Mode ESP AES/SHA1	155%	39%	17%
Transport Mode AH SHA1	180%	40%	17%
Tunnel Mode ESP 3DES/SHA1	210%	45%	19%
Tunnel Mode ESP AES/SHA1	225%	45%	19%
Tunnel Mode AH SHA1	170%	40%	18%

SHA-1. Although using SHA-1 is highly discouraged because recent efforts by security researchers have reduced the level of effort needed to successfully attack it from 2^{80} to 2^{69} operations [72], as a practical matter SHA-1 continues to be widely used because it still provides an adequate level of protection for most situations. If increased protection is required, the SHA-2 family of digital signatures (SHA-224, SHA-256, SHA-384, or SHA-512) is recommended pending completion of the NIST competition for SHA-3 [68]. A commonly used alternative to SHA-1 is MD-5 [59]. Like SHA-1, the use of MD-5 is discouraged as a consequence of the discovery of practical attacks against it 2004 by Wang et al [71].

Using our previous authentication and authorization example, the estimate of the number of packets that must be exchanged P_T, assuming reauthentication and reauthorization for each information data exchange and there are no failures in the authentication and authorization process, is given by

$$P_T = ((n_1)(A) + (n_2)(D))(OH_{PRO}) \tag{5.9}$$

n_1 is the number of authentication and authorization process steps for a single authentication and authorization

A is the size of the authentication and authorization packets

n_2 is the number of steps to exchange data after a single authentication has occurred

D is the size of the data packets

OH_{PRO} is the protocol overhead

If there are failures or repeats of any of the steps in the authentication process, the number of steps n_1 is increased accordingly. Similarly, if there are repeats or failures of any of the steps n_2 is increased accordingly. Table 5.5 illustrates the OTAR system performance when there are no failures or repeats in either n_1 or n_2 using the parameters as cited.

$n_1 = 10$

Table 5.5: Communications Overhead OTAR System- Tranmission Rates

Channel Transmission Rate (Bits)	Authentication and Authorization Delay (Seconds)
1.2 kbps	8.4
4.8 kps	2.1
9.6 kbps	1.05
19.2 kps	0.53

Table 5.6: Communications Overhead OTAR System- Enhanced Railroad Transmission Rates

Channel Tranmission Rate (bytes)	Authentication and Authorization Seconds
1200 kbps	1.05
4800 kbps	0.2625
9600 kbps	0.13
19200 kbps	0.0625

$A = 40$ bytes (Table 5.3)

$n_2 = 4$

$D = 40$ bytes (Table 5.3)

$OH_{PRO} = 225\%$ (assuming Tunnel Mode ESP AES/SHA1) (Table 5.3)

$$P_T = ((10)(40) + (4)(40))2.25 \tag{5.10}$$

$$\mathrm{P}_T = 1260 \text{ bytes}$$

for various communications transmission rates that are commonly available on the railroad. Table 5.5) and for higher speed links (Table 5.6) are calculated using:

$$Time = \frac{(P_T)}{(TR} \tag{5.11}$$

where P_T (in bytes) and TR (bits per second). Typical values for P_V in are 142,000 miles

Table 5.7: Performance Delays

Channel Transmission Rate (Baud)	Authentication and Authorization Delay (Seconds)
15	8.5
600	2.2
1200	1.11
4800	0.3625
9600	0.23
19200 kbps	0.1623

per second for radio through air, and 133,00 miles per second for coaxial cable. For transmissions between a wayside unit or an onboard computer system that is 2500 miles from the dispatcher (a worst case scenario in which a dispatcher on the east coast is communicating with a wayside or onboard unit on the west coast), P_{PD} is 0.017 seconds (17 milliseconds) for a unidirectional transmission, or 34 milliseconds for each bidirectional transmission. If one assumes that the tramsmiter and receiver take 30 milliiseconds each to process received data, the total propogation and delay time is roughly 100 milliseconds. When propagation delay is added to the delay associated with the channel transmission rate, the performance delays $\text{Performance}_{Delay}$ are shown in 5.7:

Chapter 6: INTEGRATING TRUST MANAGEMENT WITH SCHEDULING

Each railroad company is an independent commercial entity that interchanges crews, locomotives, and their associated consists with other railroads. These personnel and equipment exchanges occur at fixed geographical points where the tracks from one company are interconnected with tracks from another company. There are a limited number of these interchange points between any two companies, and they are geographically disperse. Because trains have a single degree of freedom with respect to their operations, that is, they can only operate along the tracks, any delay of a train at an interchange point as it crosses from the operating domain of one railroad to the operating domain of another has the potential to delay the movement of subsequent trains operating along the same line to the same interchange point. Figure 6.1 shows two railroad companies (Company A and Company B) that have a common interchange point. Train 1, Train 2, Train 3 and Train 4 are sequentially moving along the railroad operated by Company A to the interchange point. If Train 1 is delayed in moving from the domain of Company A to Company B, Trains 2 through 4 may potentially be delayed.

In the scenario illustrated in Figure 6.1, the consequences of the delay of Train 1 at the interchange point may be mitigated to some extent by the availability of a siding S. If the train dispatcher for Company A is aware sufficiently in advance of the arrival of Train 1 to the interchange point of a potential delay, the dispatcher could direct Train 1 into the siding S, allowing Train 2 to proceed along the main line to the interchange point. However, if the siding S is not available, or Train 1 has passed the point in which will allow the dispatcher to direct Train 1 into the siding S, Train 1 will block the following trains from reaching the interchange point. Even if the dispatcher was able to safely divert Train 1 into

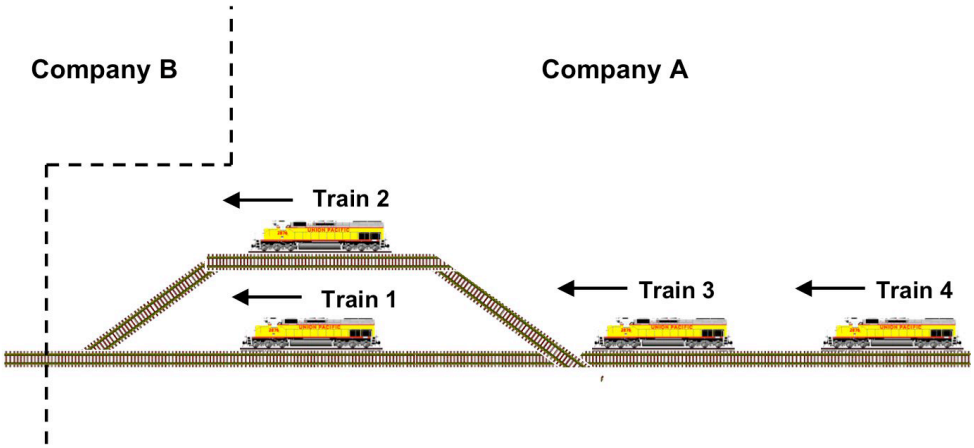

Figure 6.1: Interchange Point

the siding S, allowing Train 2 to proceed along the mainline to the interchange point, any delay encountered in the process of moving Train 2 at the interchange point will delay the following Trains 3 and 4.

As independent operating entities, each railroad operates its own trust management system within its own security domain. These trust management systems may or may not share a common security policy, or be fully interoperable. In order to ensure the safety and security of trains entering into its domain, each railroad must authenticate the trains and their crew entering its domain from the other domain. For example, Chapter 5 illustrates one possible trust management scheme. The inability to either correctly or promptly authenticate the crew and locomotive of Train 1 or Train 2 would delay Trains 3 and 4.

Because delays reduce the throughput on both railroad lines, there is a strong financial incentive on the part of both railroads to minimize delays. While delays can be mitigated with the addition of more sidings, it is expensive to do so. The Transportation Research Board of the National Academy of Sciences estimates the initial cost of adding a single additional siding with Centralized Traffic Control (CTC) on level grade to hold a single delayed train is on the order of $10 million [74]. The costs increase if additional work is required to establish a level grade for the siding. This estimate does not include additional recurring maintenance costs for track and roadbed, switch, and signal system maintenance. Other track improvements such as replacement of rails to support higher operating speeds or altering track alignment are equally, if not more, expensive.

6.1 Scope and Impacts of Delay

Delays add to a railroads cost of business and can have a significant impact on the US economy. In 1997, due to service delays on the Union Pacific (UP) railroad, the State of Texas alone encountered excess costs of over $1.0 billion. The impact of the delays was so wide spread that the US Surface Transportation Board found it necessary to invoke an

Emergency Order to address the situation [75]. While it is unlikely that delays resulting from the a failure to authenticate a train or its crew at the interchange point in a timely manner would have domain consequences such as those experienced in UPs 1997 meltdown, they would still have an adverse economic effect. For a railroad to take action to prevent, or at the least minimize, the adverse effects of authentication based delays proactivel, the time dependent behaviors of the trust management system must be integrated into the scheduling models used to control train operations.

In the US alone, major transit agencies such as the San Francisco Bay Area Rapid Transit (BART), New York City Transit (NYCT), Metropolitan Atlanta Rapid Transit Authority (METRA) and Washington Metropolitan Transit Authority (METRO) have implemented or are in the process of implementing CBTC systems. Other driverless CBTC systems can be found in people movers at or near major airports such as Tampa, Orlando, Atlanta, Washington (Dulles), Jacksonville, Las Vegas, San Francisco, Pittsburgh, Huston, Dallas/Ft Worth, and Detroit. While the underlying concepts and technology are similar for both the rail and non-rail systems, there is a fundamental difference between the two domains that eliminates delays associated with cross domain certification and authentication issue. That is unlike CBTC equipped trains in the general rail environment, where crews, and equipment are routinely exchanged between different railroad companies, CBTC equipped systems in the non-general rail environment are closed, inferring that they do not interchange equipment or crews with each other as in the general rail environment. As a consequence, the cross-domain certification issue becomes moot. While the manned transit systems still must consider the issue of authentication of the crews, the driverless systems totally eliminate them.

General rail systems outside the US that employ CBTC systems must also address the cross-domain certification issue. The European Union (EU) has been mitigated this issue through the use of a single, secure system. The European solution is called ERTMS

(European Rail Traffic Management System). ERTMS differs from US CBTC systems in that it uses a common mandated standard. The underlying communications technology for ERTMS is Global System for Mobile Communications- Rail (GSM-R) [76] GSM-R provides a secure platform for voice and data communication based on GSM. GSM authenticates the subscriber using a pre-shared key and challenge-response that provides confidentiality and authentication. The system design uses a common cryptographic key stored on a Subscriber Identity Card (SIM) that is integral to the system. Because the use of a SIM is mandatory in all GSM-R devices, the SIM is configured when initially installed. Consequently there is no need to exchange keying material over the communications channel when every user authenticates with a common network. A train fitted with complete ERTMS equipment can interoperate on any European Train Control System (ETCS) route without any technical restrictions

In non CBTC systems, the entire issue of authentication delays of crews or trains, either within a company, or between companies, is ignored. While this approach does eliminate one avenue of attack against the railroad (through its communications network), it still leaves them vulnerable to more traditional attacks against the physical infrastructure (track, signals, and wayside devices), equipment (locomotives and cars) or personnel, without any of the safety benefits.

6.2 Basic Model

The primary goal of efficient secure inter-domain rail operation is the minimization of rail Traffic Delays. Traffic Delays can be considered as a combination of two separate, but interrelated elements. First are delays resulting from the specific physical operating characteristics of the trust management, dispatching and communication systems. The physical operating characteristics include slack time built into the trains schedule, traffic congestion,

scheduled stops, authorized speeds, location of other trains, on track equipment, maintenance of way work zones, track physical condition, status of signals and communication bandwidth.

The second are delays that arise from the manner in which trains are scheduled to arrive and depart the interchange point. Two specific activities must occur before a train is authorized to pass from one railroad domain to the second domain. First, the train and the crew leaving one domain for the second domain must be authenticated before a movement authority can be granted to allow the cross-domain movement. Second, track space must be available in the second domain to allow the issuance of the movement authority. Delays in a train T_X moving from domain A to domain B can delay the subsequent scheduled movement of trains T_{X+1} through T_{X+N} resulting in increasing traffic delays. By minimizing, or eliminating authentication delays, delays in the granting and issuance of movement authorities is reduced, with resulting reductions in traffic delays.

Failures or induced delays resulting or associated with the trust management scheme have an impact on train routing. There is an extensive body of work on optimization of network wide routing in general, and railroad networks in particular. The problem, however, is greatly simplified because at any single interchange point, there is only a single line connecting the general rail networks of the two companies who have established the interchange point. This allows us to consider the track connecting the railroads through the interchange point as a single track railroad. Traffic movements on this single track railroad can then be optimized to minimize the impact of delays. There are numerous approaches already to accomplish this [77–86]. These approaches, however, do not consider authentication delays that could be induced by a trust management system.

Railroad routing is a highly constrained network optimization problem that has confounded traditional optimization methods. An extensive body of literature exists regarding different

methodologies for optimal dispatching and routing of trains in rail networks to minimize delays and their study remains the subject of extensive research by the Transportation Science and Logistics Society Railway Applications Section (RAS) of The Institute for Operations Research and the Management Sciences (INFORMS). No attempt will be made to either develop new, or improve upon existing dispatching and routing methodologies or consider more complex interchange configurations. The model assumes a single-track queue of trains and there are no other topological additions other than a single-track siding off the main line. There is no other merging or branching traffic. The model also assumes the availability of the details of specific mal-actor attacks against wireless communications based systems.

The other basic mechanism for the temporary storage of trains is the spur track. A spur track provides a location for the temporary storage of a train T_{X+l} following a train T_X. This decreases the cost of the creating temporary storage as compared to the cost of a siding, but with an increase in the complexity of train movements. For a spur opposing the direction of travel, a train TX_{X+1} following a train T_X would have to proceed past the entry point and back in to the spur wait for further movement authority. For a spur facing the direction of travel the following train T_{X+1} could proceed directly into the spur, but its exit would require holding the following train T_{X+2} at a position on the main so that the Train T_{X+1} could back out after receipt of a movement authority. (Figure 6.2),

The unidirectional model studied represents only one of a number of possible, increasingly complex, track configurations that could be created using sidings and spurs. Multiple parallel sidings at one location can be added, or sequential sidings at different locations and configurations (Figure 6.3) can be built off the main track. Multiple parallel facing or reverse spurs can be built (Figure 6.4), or combinations of facing and reverse spurs can be added (Figure 6.5) as well as more complex configurations of spurs and sidings (for example Figures 6.6, 6.7, 6.8, 6.9, 6.10, and 6.11). The number and configurations of spurs and sidings is limited only by the physical space available to lay the track, and the amount of

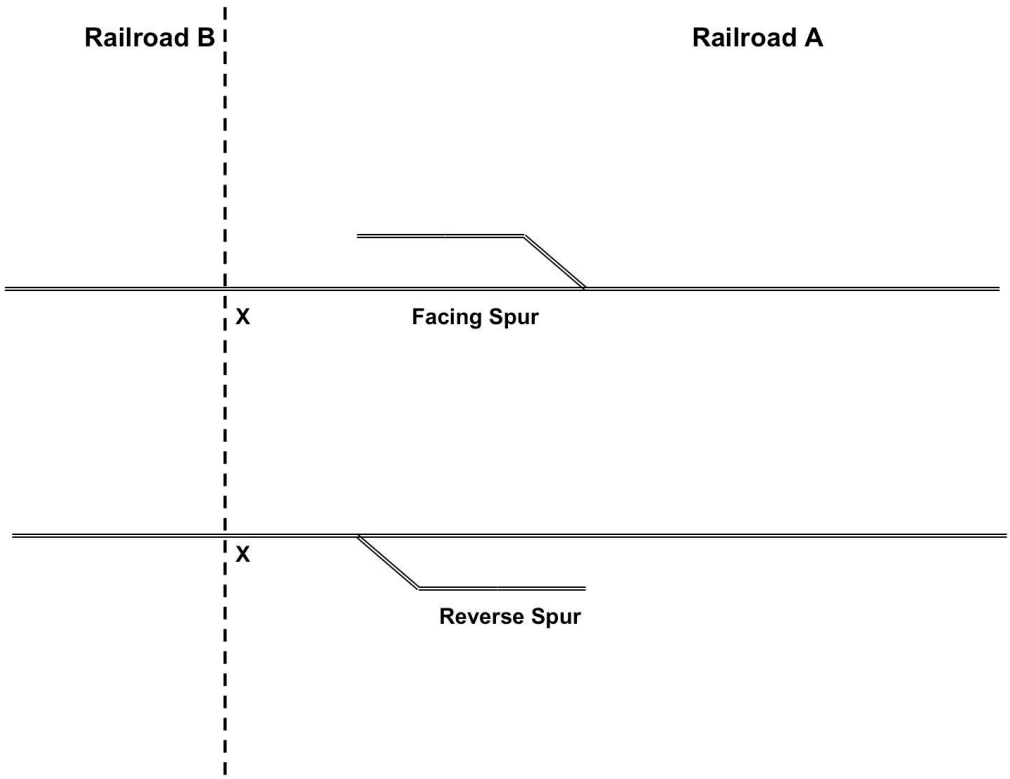

Figure 6.2: Reverse and Facing Spurs

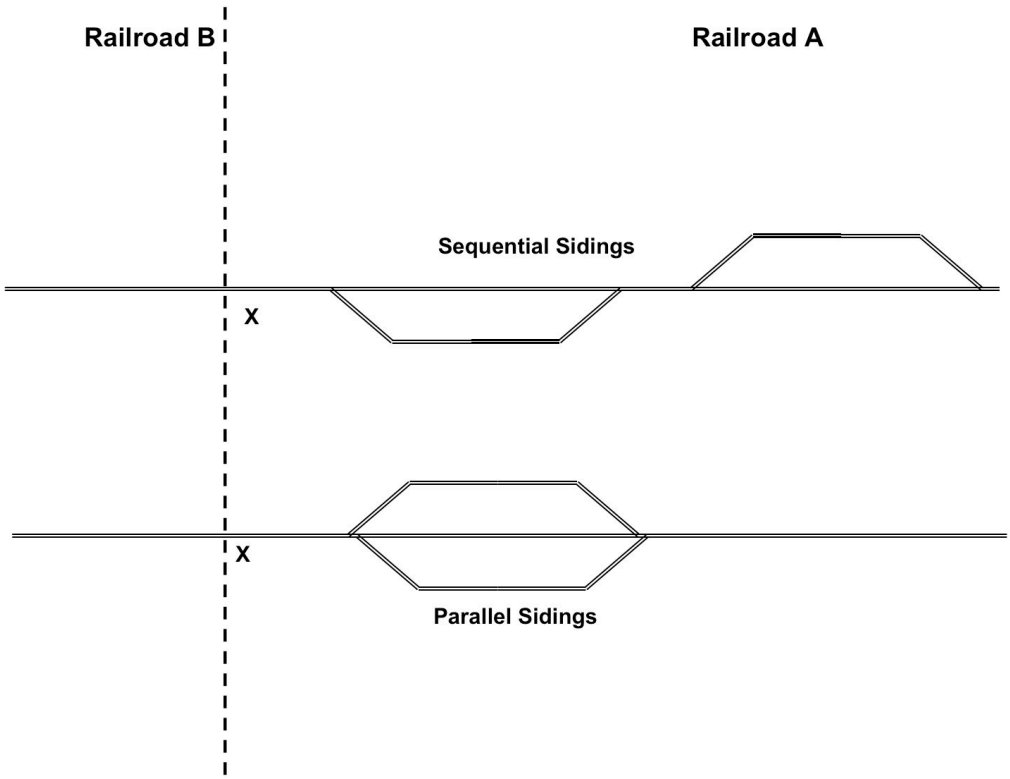

Figure 6.3: Parallel and Sequential Sidings

money available for construction. As the configurations become more complex, they allow the Domain A Dispatcher increasing latitude authorizing the movement of trains towards the interchange point to best support the railroads business and operational requirements.

A critical assumption that allows for the unidirectional analysis of the model is that trains moving from Domain B to domain A are doing so on a separate main track. This allows us to exclude the traffic from domain B to the analysis. Track configurations (and the associated analysis) in domain B are similar to those just discussed for domain A exist in domain B. The track configurations in domain A and domain B need not be the same, and are often different because the installed configuration is a business decisions of the railroads involved. At infrequently used interchange points, the presumption of a double main track is invalid and trains moving from domain A to domain B share the same track as trains moving from domain B to domain A. In this more complex situation, Dispatcher A and Dispatcher B must carefully coordinate the movement of trains in both directions so as to not only prevent gridlock (the inability of trains to move in either direction), or collisions (authorizing two opposing trains to occupy the same track at the same time). The complexity of the routing and scheduling problem, as well as the associated analysis, increases significantly.

6.3 Dispatcher & Train Interactions

While the preceeding discussion addressed alternative configurations that might exist, the interactions and performance issue studied for this work only considered the basic unidirectional configuration and the performance issues once authorization has been requested, and received, by a train waiting at the interchange point to cross domains. It does not address the larger issue of increased track configuration complexity, nor bidirectional unitrack sequencing of trains between dispatcher A and B. The movement of trains through out a rail system in domain A or in domain B is not necessarily optimized for behavior at an

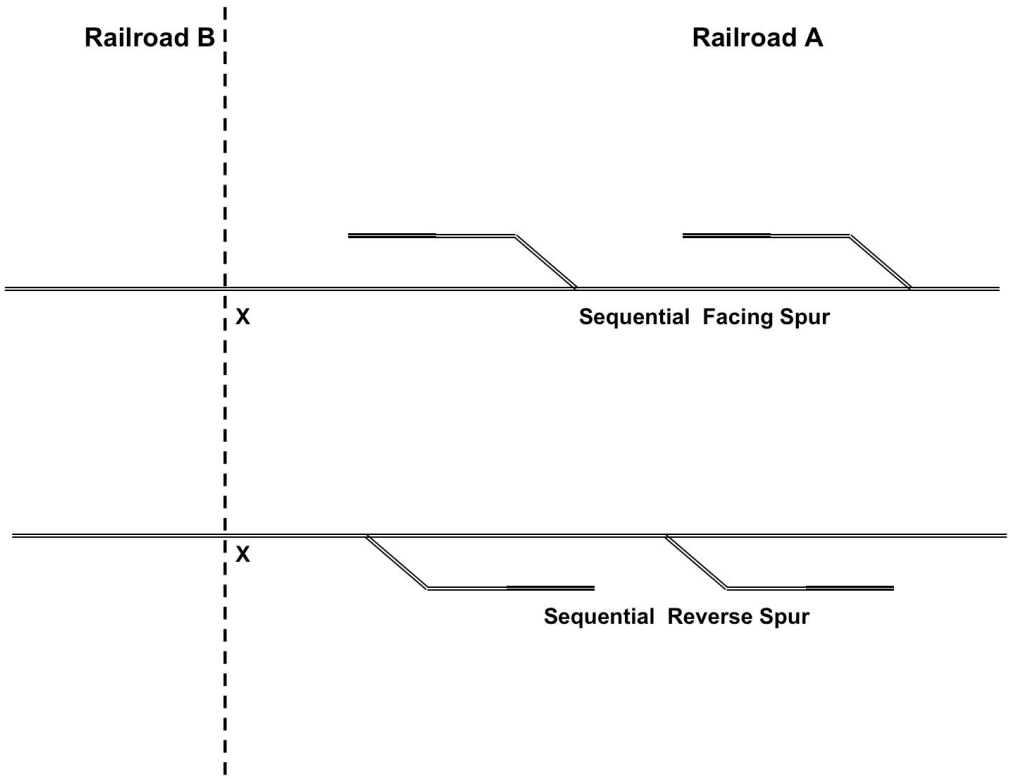

Figure 6.4: Multiple Spurs

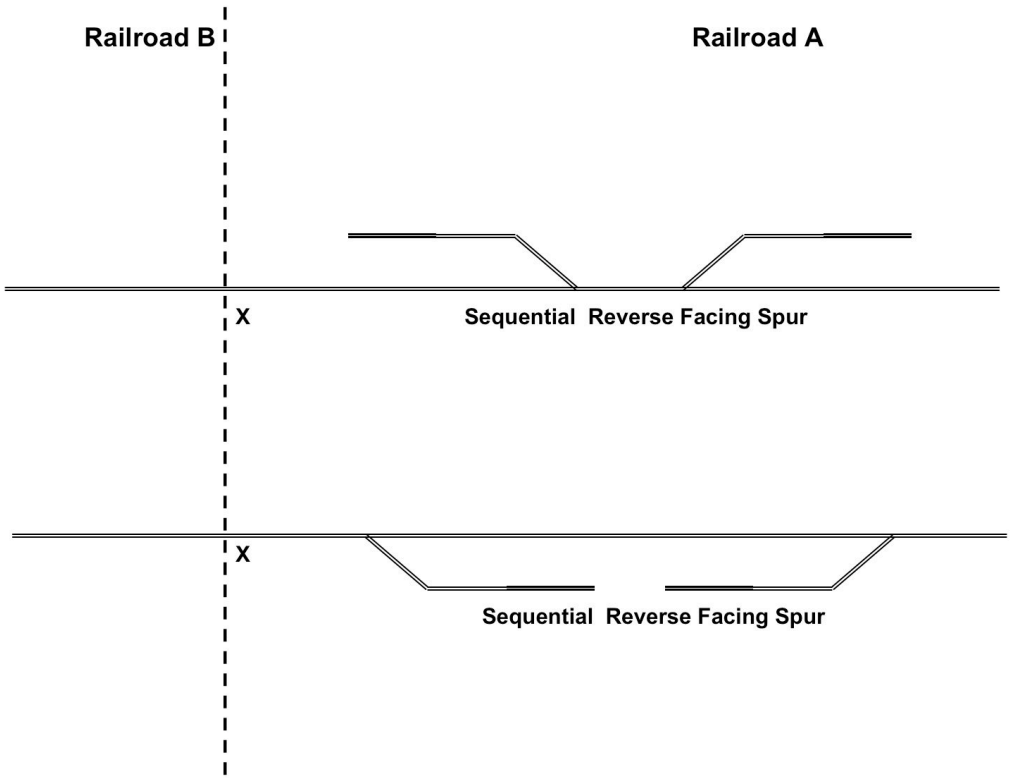

Figure 6.5: Sequential Spurs

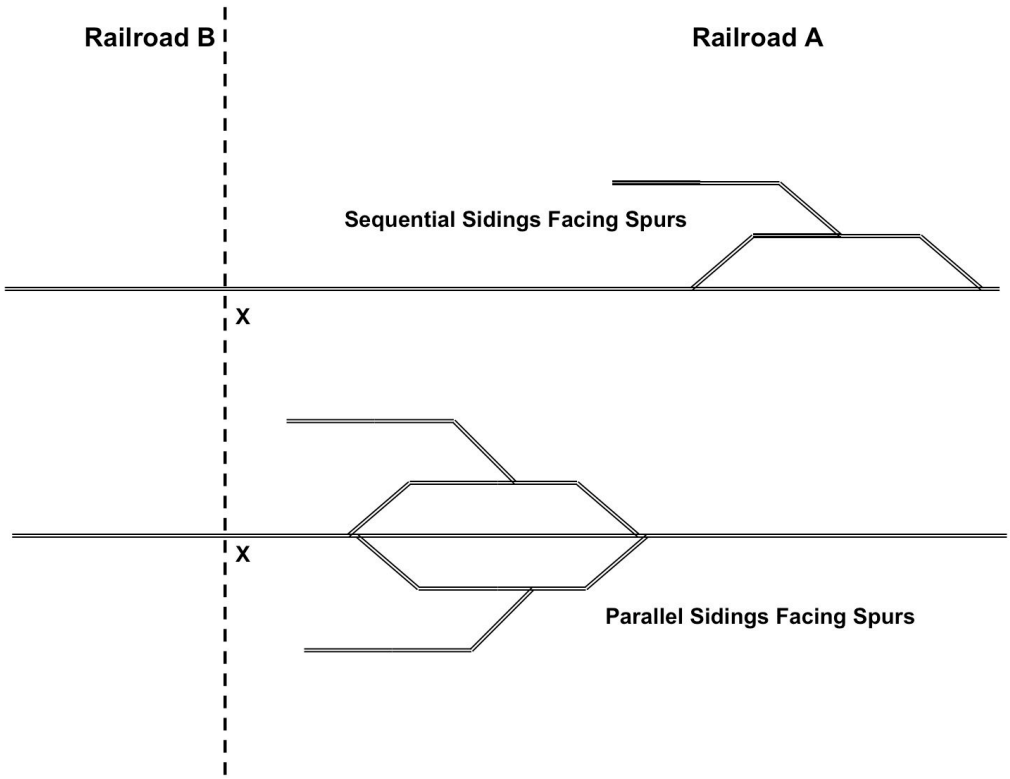

Figure 6.6: Sequential and Parallel Sidings and Facing Spurs

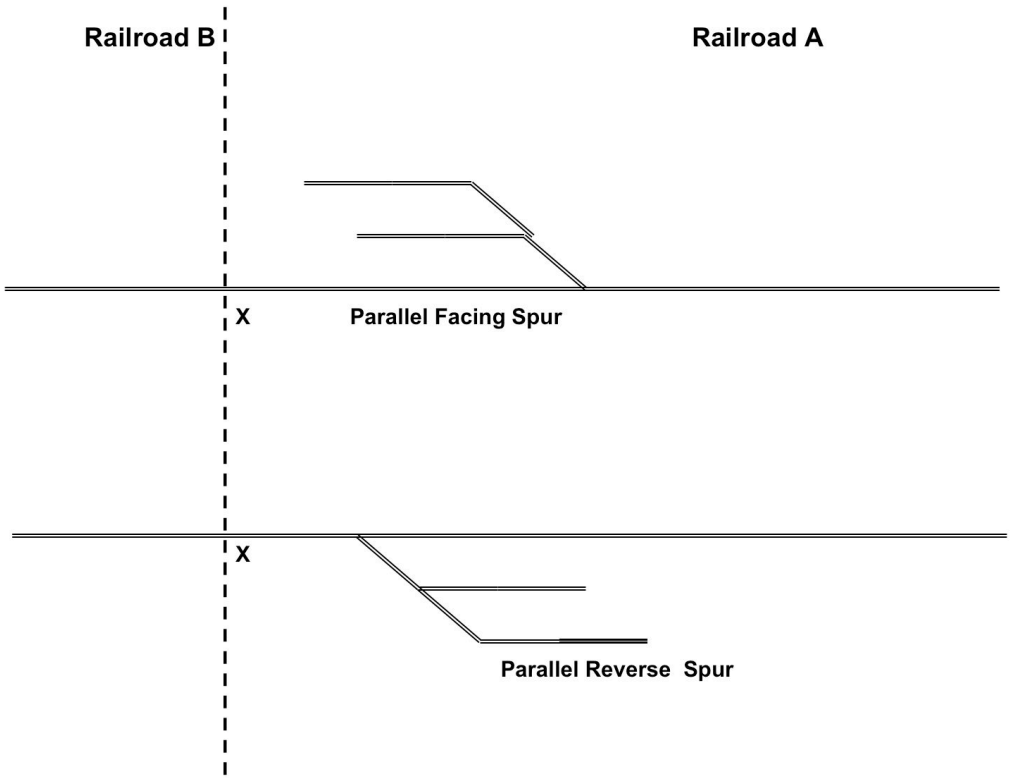

Figure 6.7: Parallel Facing and Reverse Spurs

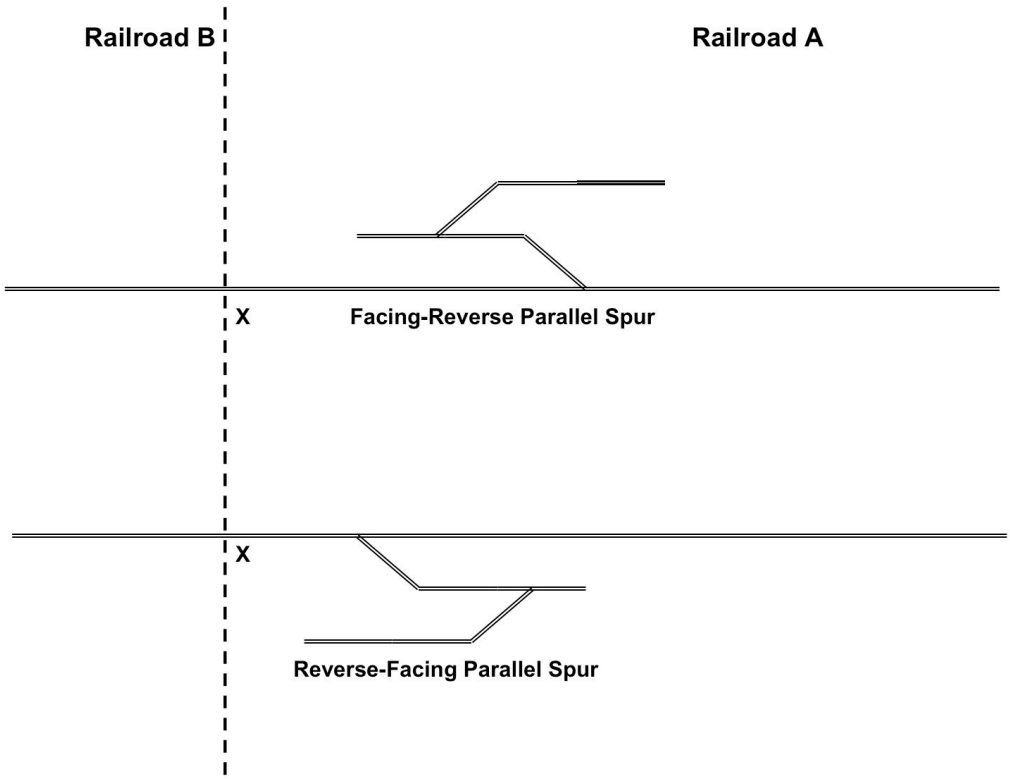

Figure 6.8: Parallel Facing and Reverse Spurs

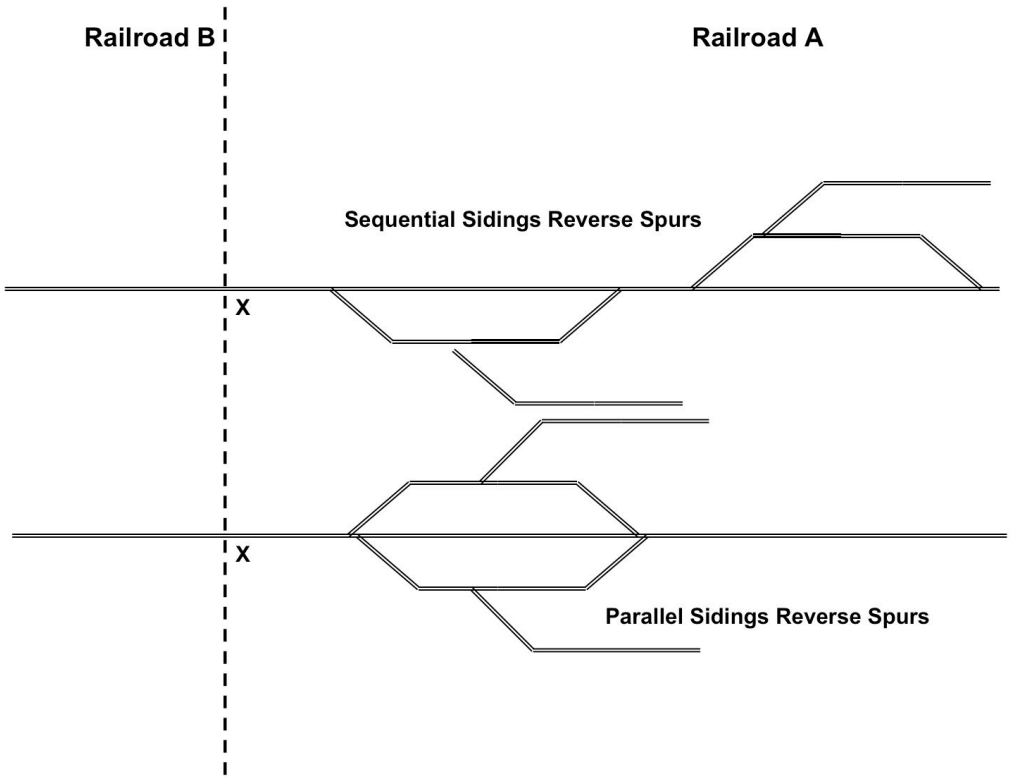

Figure 6.9: Sequential and Parallel Sidings and Reverse Spurs

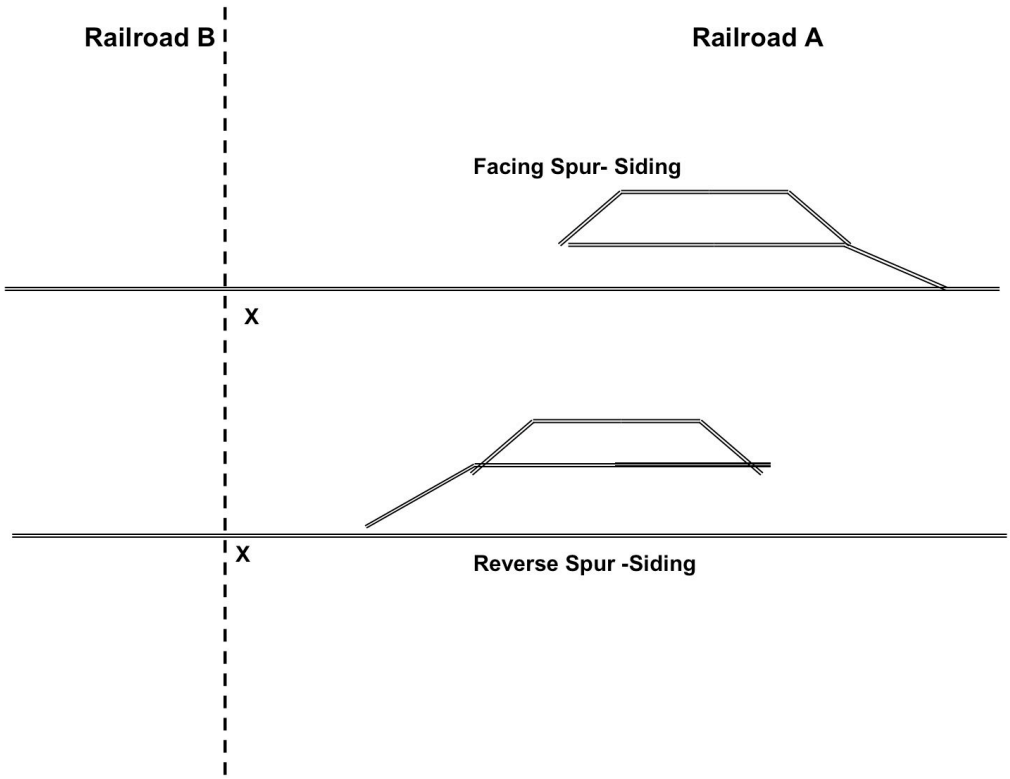

Figure 6.10: Facing and Reverse Spurs and Sidings

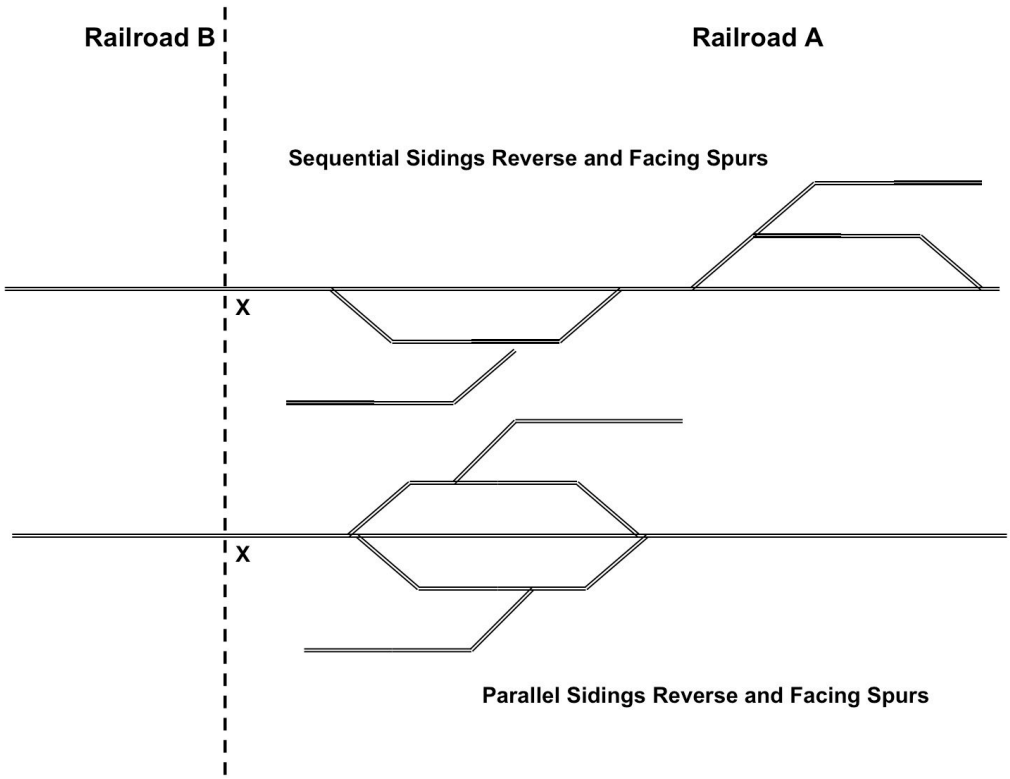

Figure 6.11: Sequential and Parallel Sidings with Facing and Reverse Spurs

interchange point, but rather it is optimized to support the most efficient use of rail assets (cars, locomotives, and track) within each individual railroad's operating domain. This is a more general operations research problem, and has been the subject of a significant study. Current research includes [100–107].

Movement authorization of trains is only a small part of the general railroad planning process. Not only must the railroad planning process address which particular rail lines are used (line planning), but must also address customer service requirements (demand analysis), consist management (allocation of train cars and locomotives), and crew management (distribution and allocation of the trains crew). Each of these have different, and often competing goals. Computing an optimal system wide (strategic) solution requires the ability to schedule the right trains frequently enough to be service-responsive to customers, long enough to be cost effective, and spaced so as to minimize transfer time in yards and congestion over the right of way. Solutions to this larger planning and scheduling problem is outside the scope of this work. We however provide an algorithmic description of the possible tactical behaviors of Dispatcher A and Dispatcher B regarding the movement of trains from domain A to domain B.

Our description of the tactical behaviors of Dispatcher A and Dispatcher B relies on the following assumptions:

1. There is a main track and a single siding in domain A and a single main track in domain B.
2. All trains in domain A are of the same length, but may have different priorities for movement.
3. Train movements are from Domain A to Domain B.

4. Dispatcher A (DS_A) and B (DS_B) have exchanged a session key between each other. Dispatcher A (DS_A) has authenticated locomotive L_X and the associated engineer E_X prior to receiving movement requests.

5. Dispatcher A (DS_A) controls the signal whose aspect controls the movement of a train from domain A while Dispatcher B (DS_B) controls the signal whose aspect controls the movement of a train into domain B.

6. For a train to leave domain A and enter domain B, both the Dispatcher A and Dispatcher B have to authorize movement, coordinating the signal aspects.

7. The siding can contain only one train, the main track parallel to the siding may also contain one train.

8. There are up to N trains in the queue awaiting authorization to enter domain B.

9. Requests for authorizations from A to B are in order of increasing distance of trains from interchange point.

6.4 Algorithmic Behavior

The behavior of each of the dispatchers and trains can be expressed algorithmically as follows:

6.4.1 Data Structures

```
Track = array [0..maxlocation] of Map
Map = record
        Train_ID = text;
        Location = int;
        Location_Status = (occupied, unoccupied, unknown);
       end_(record;

Authority_request = Authority_type;
Authority = Authority_type;
```

```
Authority_type = record
           Train_ID = text;
           Current_location = int;
           Next_location = int;
           Movement_authorized = (yes, no);
           Certifcate_authenticate = (yes, no);
           Certificate = authentication_info;
           Timer = int;
           end_record;

Authentication_info = record
              Engineer_cert = certificate;
              Train_cert = certificate;
              end_record

event = (initialize_map,
        receive_authority_request (Domain, Authority_request),
        receive_authority (Domain, Authority),
        receive_authority_request_timeout (Domain,Authority_request);
        receive_authority_timeout (Domain,Authority)

Domain = (A, B, Train(i).Train_ID) of Status;
Status = (operational, nonoperational);

Trains = array [1..maxtrain] of Train_info;
Train_info = record
          Train_ID = text;
          Current_location = int;
          Certificate = authentication_info;
          end

Train_symbol = train_info;

i,k = int;
maxtrain = int;
maxlocation = int;
maxtimer = int;
timer = int;
countreqA = int;
countreqB = int;
countID = int;
maxcountBreq = int;
maxcountAreq = int;
maxcountIDreq = int;
countA = int;
```

```
countID = int;
```

6.4.2 Dispatchers and Trains Main Programs

```
/* Dispatcher Domain A                    */
/* Authority Request received from        */
/* Train.Train_ID                         */
/* or Authority is receivd from Domain    */
/* B                                      */

Dispatcher_A
begin
While (A.Status = operational)
   receive (event);
     begin case
       case event =
         initialize_map;
         handle initialize_map;
       case event =
         receive_authority_request (A, Authority_request);
         handle receive_authority_request (A, Authority_request);
       case event =
         receive_authority_request_timeout (A,Authority_request);
         handle receive_authority_request_timeout(A,Authority_request);
       case event =
         receive_authority (A, Authority);
         handle receive_authority (A, Authority);
       case event =
         receive_authority_timeout (A, Authority);
         handle receive_authority_timeout (A, Authority);
       else THROW_ERROR;
     end case;
end.

/* Dispatcher Domain B                    */
/* Authority Request received from A      */
/* asynchronously                         */

Dispatcher_B
begin
While (B.status = operational)
  receive(event);
     begin case
      case event =
```

```
            receive_authority_request (B, Authority_request);
            handle receive_authority_request (B, Authority_request);
      case event =
            receive_authority_request_timeout (B,Authority_request);
            handle receive_authority_request_timeout (B,Authority_request);
      else THROW_ERROR;
     end case;
end.


/* Train Domain */

Train
begin
   While (Train(i).Status = operational)

/* Train must transmit a request for a new authority   */
/* to Dispatcher A. The Authority_request fields       */
/* are assumed to be prepared externally and           */
/* and sent asynchronously                             */
/* authority requests are of the form                  */
/* reeceive_authority (A, Authority)                   */


receive (event);
  begin case
    case event =
     receive_authority_request (A, Authority_Request);
     send (event);
    case event =
     receive_authority_request_timeout (A,Authority_request);
     handle receive_authority_request_timeout (A,Authority_request);
    case event =
     receive_authority (Authority.Train_ID, Authority);
     handle receive_authority (Authority.Train_ID, Authority);
    case event =
     receive_authority_timeout (Authority.Train_ID, Authority);
     handle receive_authority_timeout (Authority.Train_ID, Authority);
    else THROW_ERROR;
     end case;
end.
```

6.4.3 Initialize

```
/*  Event Code- Event initialize_map        */
/*  Domain B  Track(0) is main              */
/*  Domain A  Track(1) is main              */
/*  Domain A  Track(2) is siding            */
/*  Domain A  Track(3..maxlocation) is main */

Intialize;
begin
   i: = 0, k:= 0;
   repeat
       receive(Track(i).Location_status);
          begin case
            case (Track(i).Location_status = unoccupied)
                 or (Track(i).Location_status = unknown)
                   Track(i).Train_ID := nil;
            case (Track(i).Location_status = occupied)
                receive (Train_Symbol)
                Track(i).Train_ID := Train_Symbol;
                Train(k) := Train_Symbol;
                k := k + 1;
            else THROW_ERROR;
          end case;
        Track(i).Location := i;
        i := i +1;
    until i = maxlocation;
end. /* intialize map */
```

6.4.4 Request Authority

```
/* Event Code receive_authority_request (Domain, Authority_request) */
/* Authority Requests can be received by A and B                    */

receive_authority_request (Domain, Authority_request);
begin
    reset_start (timer);
    case Domain = A
       /* A is responsible for certifying the trains   */
       /* Both engineer and locomotive cert must be    */
       /* valid before A will approve an authority     */
       /* request and either return it to orginator    */
       /* or forward it to B                           */
      begin
         if (Authority_request.Certifcate.Engineer_cert is valid) and
```

```
       (Authority_request.Certifcate.Train_cert is valid) and
       (timer < maxtime)
/* timer is given by Table 5.1  */
  then  Authority_request.Certifcate_autheticate := yes
  else  Authority_request.Certifcate_autheticate := no

  if (Authority_request.Certifcate_autheticate = no)
            /* A denies authority request  */
   then
      Authority.Train_ID := Authority_request.Train_ID;
      Authority.Current_location :=
                 Authority_request.Current_location;
      Authority.Next_location :=
                 Authority_request.Next_location;
      Authority.Movement_authorized := no;
        Authority.Certificate := Authority_request.certificate;
        event := receive_authority (Authority.Train_ID, Authority);
        send (event);
    else
 /* crew and locomotiove are authenticated and dipatcher A  */
 /* must determine if authority request can be handled by A */
 /* or must be passed to B                                  */

    begin case
      case ((Authority_request.Current_location = 1) or
            (Authority_request.Current_location = 2))
             and (Authority_request.Next_location = 0)
              /* A must prepare authority request for */
              /* transmission to B                    */
           then event
           := receive_authority_request (B,Authority_request);
              send(event);

              /* No interaction with domain B        */
              /* Train has possible routing into the */
              /* main or the siding                  */
       case (Authority_request.Current_location = 3) and
            ((Authority_request.Next_location = 2) or
            (Authority_request.Next_location = 1)
            then
            if Track(2).Location_status = unoccupied then
              /* train will go in siding */
               Authority.Train_ID
                  := Authority_request.Train_ID
               Authority.current_location
                  := Authority_request.Current_location;
```

```
    Authority.Next_location :=  2;
    Authority.Movement_authorized := yes;
       Authority.Certificate
         := Authority_request.Certificate;
    event:=
     receive_authority (Authority.Train_ID, Authority);
     send (event);

   else if Track(1).Location_status = unoccupied then
   /*train will stay on main*/
     Authority.Train_ID := Authority_request.Train_ID;
     Authority.Current_location
        := Authority_request.Current_location;
     Authority.Next_location :=  1;
     Authority.Movement_authorized := yes;
     Authority.Certificate
        := Authority_request.Certificate;
     event :=
       receive_authority (Train.Train_ID, Authority);
       send(event);

     else
   /* main and siding track both blocked */
   /* do not advance train               */
      Authority.Train_ID
       := Authority_request.Train_ID;
      Authority.Current_location
       := Authority_request.Current_location;
      Authority.Next_location
       := Authority_request.Next_location;
      Authority.Movement_authorized := no;
      Authority.Certificate
       := Authority_request.Certificate;
      event :=
       receive_authority (Authority.Train_ID, Authority);
      send(event);

case (authority_request.current_location > 3)
/* train is further up the track than */
/* the main/siding branch             */
/* check to see if next block is fre  */
if
 Track(Authority_request.Current_location -1).Location_status
                      = unoccupied
/* move the train */
 then
```

```
                Authority.Train_ID
                  := Authority_request.Train_ID;
                Authority.Current_location
                  := Authority_request.Current_location;
                Authority.Next_location
                  := Authority_request.Current_location -1;
                Authority.Movement_authorized := yes;
                Authority.Certificate
                  := Authority_request.Certificate;
                event :=
                  receive_authority (Authority.Train_ID, Authority);
                  send (event);

                else then
            /* next block is not free  */
            /* do not move the train */
                Authority.Train_ID
                  := Authority_request.Train_ID;
                Authority.Current_location
                  := Authority_request.Current_location;
                Authority.Next_location
                  :=  Authority_request.Current_location;
                Authority.Movement_authorized := no;
                Authority.Certificate
                     := Authority_request.Certificate;
                event :=
                 receive_authority (Authority.Train_ID, Authority);
                 send(event);
              end if;
          end case ;
   case (timer > maxtimer)
         event :=
           receive_authority_request_timeout(A,Authority_request);
         send( event);
   end case;
  end if;

case Domain = B then
   /* A has passed decision making to B          */
   /* Check if track and available to move into  */
  begin case
     receive (Track(0).Location_status);
     case (Track(0).Location_status = unoccupied) and
         (Authority_request.Next_location = 0) and
         (timer < maxtimer) then
```

```
            /* advance train */
                Authority.Train_ID := Authority_request.Train_ID;
                Authority.Current_location
                      := Authority_request.Current_location;
                Authority.Next_location :=  0;
                Authority.Movement_authorized := yes;
                Authority.Certificate
                      := Authority_request.Certificate;
                event := receive_authority (A, Authority);
                send(event);

          case (Track(0).Location_status = occupied) and
               (timer < maxtimer) then
            /* do not advance train */
                Authority.Train_ID := Authority_request.Train_ID;
                Authority.Current_location
                       := Authority_request.Current_location;
                Authority.Next_location
                       :=  Authority_request.Next_location;
                Authority.Movement_authorized := no;
                Authority.Certificate
                       := Authority_request.Certificate;
                event := receive_authority (A, Authority);
                send(event);

           case  (timer > maxtimer)then
              event :=
               receive_authority_request_timeout(B,Authority_request);
              send(event);
       end case;
      end if;
end
```

6.4.5 Receive Authority

```
/* Event Code receive_authority (Domain, Authority)      */
/* A can receive an authority from B and Pass to         */
/* train, or train can receive authority from A directly */

receive_authority(Domain, Authority);
begin
  restart (timer);
  case (Domain = A) and (timer < maxtimer)
```

```
     then
      /* repackage and forward to train  */
      event := receive_authority (Authority.Train_ID, Authority);
      send(event);

case (Domain = A) and (timer > maxtimer)
      then
        event := receive_authority_timeout(A,Authority);
        send(event);

     /*  train is receiving authority    */
       /* advance the train */
case (Domain = Authority.Train_ID)
   then
   if (Authority.Movement_authorized = yes)
      and (timer < maxtimer)
      then
        Track(Authority.Next_location).Train.ID
                             := Authority.Train_ID;
        Track(Authority.Next_location).Location_status
                             := occupied;
        Track(Authority.Current_location)
                             := unoccupied;
     else
      if (Authority.Movement_authorized = yes)
                      and (timer > maxtimer)
      then
        event:=
          receive_authority_timeout(Authority.Train_ID,Authority);
          send(event);
      end if;
   else       /* do not advance the train   */
    if (authority.movement_authorized = no)
                    and (timer < maxtimer)
      then
        Track(Authority.Current_location)Train.ID
                     := Authority.Train_ID;
        Track(Authority.Next_location).Location_status
                     := unoccupied;
        Track(Authority.Current_location)
                     := occupied;
      else
        if (Authority.Movement_authorized = yes)
                     and (timer > maxtimer)
       then
         event :=
```

```
            receive_authority_timeout(Authority.Train_ID,Authority);
           send (event);
         end if;
  end if ;
end;
```

6.4.6 Timeout

```
/* Event Code receive_authority_timeout (Domain, Authority)         */
/* Event Code receive_authority_request_timeout (Domain, Authority) */
/* On timeout, reset timer and retry  again until max retries       */
/* accomplished, then THROW_ERROR                                   */

receive_authority_request_timeout(Domain,Authority);
begin
   receive(event);
   reset(timer);
   case event =
       receive_authority_request_timeout(A,Authority_request);
       countreqA  := countreqA +1;
       if countreqA <  maxcountAreq then
         event := receive_authority_request(A,Authority_request);
       else THROW_ERROR;
   case event =
        receive_authority_request_timeout(B,Authority_request);
          countreqB  := countreqB +1;
          if countreqB < maxcountBreq then
              countreqB := 0;
              event := receive_authority_request(B,Authority_request);
          else THROW_ERROR
   send(event);
end

receive_authority_timeout(Domain,Authority);
begin
   receive(event);
   reset(timer);
   case event =
      receive_authority_timeout(A,Authority);
          countA  := countA +1
          if countA < maxcountA then
             countA := 0;
             event :=
```

```
                receive_authority(A,Authority_request);
          else THROW_ERROR;
   case event =
          receive_authority_timeout(Authority.Train_ID,Authority);
          countID  := countID +1;
          if countID < maxcountID then
             event :=
               receive_authority(Authority.Train_ID,Authority_request);
          else THROW_ERROR;
   send(event);
end
```

6.4.7 Some Usage Scenarios

The basic usage of the algorithms is shown in Figure 6.12 and is composed of the following elements:

- A train T_X with a PTC sytem PTC_X comprised of:

  E_X, the engineers certificate,

  L_X the locomotive certificate,

  PTC_X the installed PTC system,

  V_X, the initial train velocity, and

  DB_X the safe stopping (braking) distance.

- CA_A is the certificate authority of domain A.

- CA_B is the certificate authority of domain B.

- DS_A is the dispatcher of domain A.

- DS_B is the dispathcer of domain B.

- STM_A is the trust management infrastructure of domain A.
- STM_B is the trust management infrastructure of domain B.
- MA_X is a movement authority.
- S_X is a siding of length L.

A train T_X that intends to move from domain A to domain B submits the engineers certificate E_X and the locomotive certificate, L_X, to the dispatcher DS_A who forwards it into the trust management infrastructure STM_A of domain A. Upon receipt the Certificate Authority CA_A examines the received certificates E_X and L_X to determine if it can authenticate them.

If CA_A can authenticate the received E_X and L_X, CA_A notifies the dispatcher of domain A that the crew and locomotive have been authenticated. The crew aboard the locomotive can then request a movement authority MA_X from DS_A to proceed from domain A to domain B. If the CA_A cannot authenticate the engineers certificate E_X and/or the locomotive certificate L_X, the CA_A makes a secure query to other CAs in STM_A, or STM_B in an attempt to validate the certificates. In the situation that the engineers certificate E_X and/or the locomotive certificate L_X cannot be authenticated within STM_A or STM_B, this result is passed back to dispatcher DS_A who then denies any authorization request for train T_X to enter into domain B. Likewise, if authentication is successful, this result is also passed back to DS_A who may accept a request for A to pass into domain B as in Figure 6.13.

A train T_X that has requested entry from one domain to another is prohibited from proceeding into the new domain until the movement authority MA_A has been approved by the dispatcher of the new domain. In the event that T_X does not receive a response to a request, or the response to a request is delayed, T_X proceeds to the limit of its currently granted authority and stops. If T_X is already at the limits of the authority, then T_X remains halted

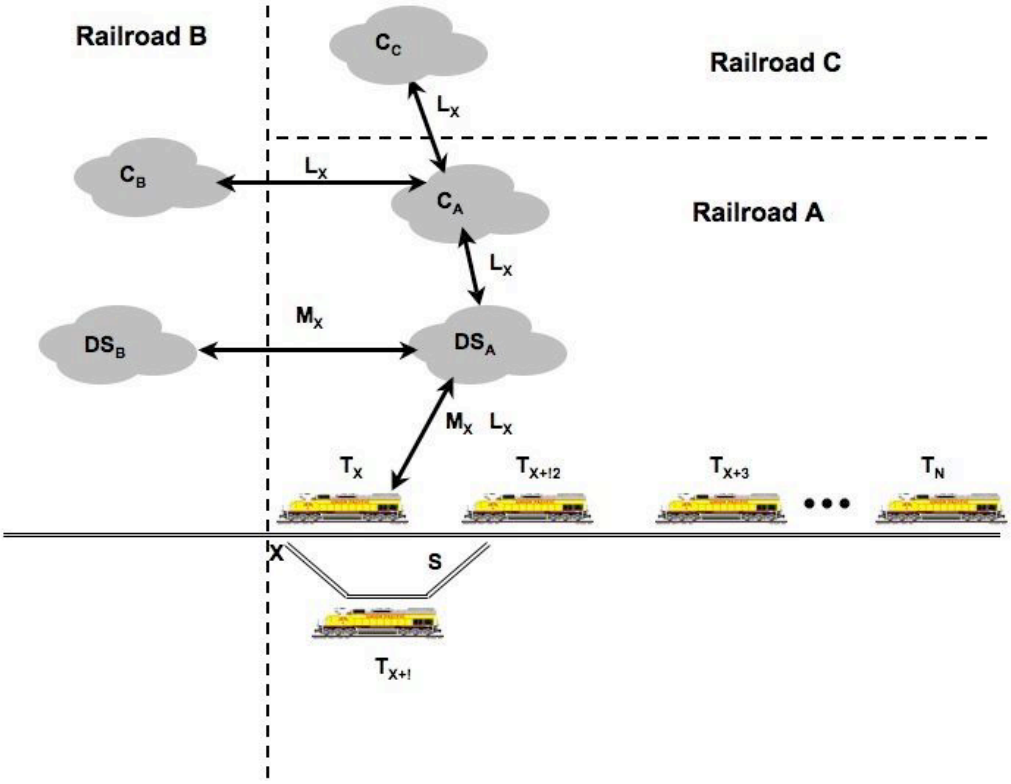

Figure 6.12: Basic Model

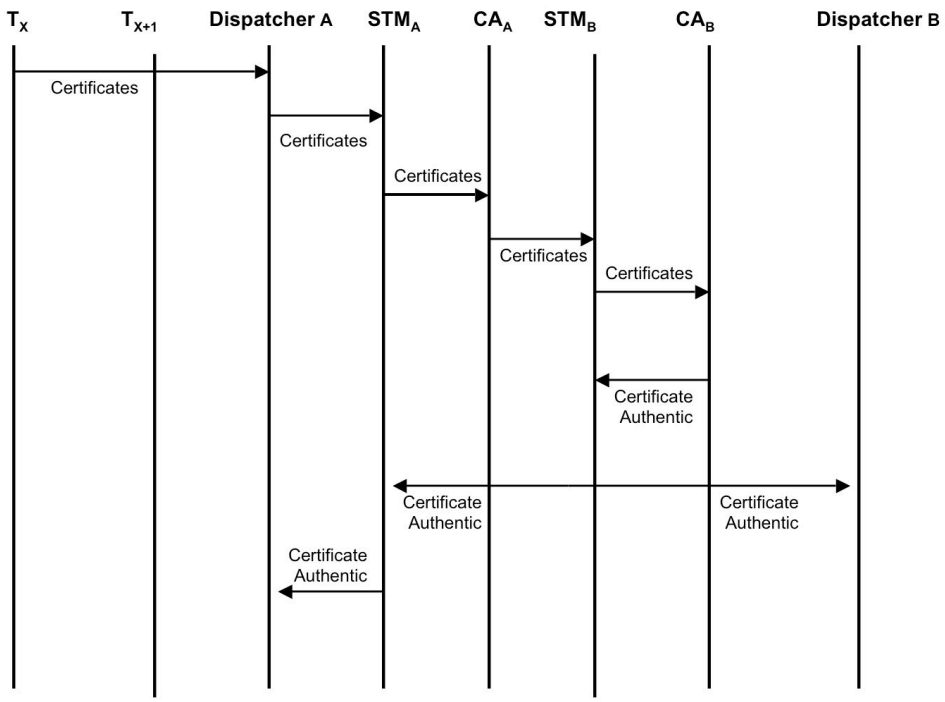

Figure 6.13: Successful Authentication

until the authority to proceed is received. The movement of subsequent trains, such as T_i for $i \geq X+1$ and $i \leq N$, are rescheduled by the dispatcher in the current domain by modifying the movement authorities to preclude collisions and overrun of authority limits as necessary.

Figure 6.14 represents the scenario where dispatcher DS_B approves MA_X for T_X (track in domain B is available), dispatcher DS_A relays the approved MA_X to T_X, and T_X transitions from domain A to domain B. Dispatcher DS_A may then reschedule T_{X+1} to advance to the block vacated by T_X, and advance subsequent trains T_i for $i \geq X+2$. Figure 6.15 represents the scenario where DS_B denies the movement authority request. This denial may be the result of an inability to authenticate T_X or the unavailabillty of track in domain B. The denial is relayed to T_X by DS_A and the engineer, or the PTC System if the engineer fails to take action, stops T_X. Dispatcher DS_A must reschedule T_{X+1} and subsequent trains. Figure 6.16 represents the scenario where T_X passes the security check but Dispatcher A denies authority MA_X to train T_X even though DS_B has approved the movement. In this case the PTC System stops T_X from entering domain B. Dispatcher DS_A then reschedules T_{X+1} and subsequent trains.

There are three possible situations that may be encountered by a train T_{X+1} that is following train T_X in Domain A

1. If the main line and siding are clear, T_{X+1} may take the main or siding and proceed to the interchange point without delay.

2. If the main is clear and the siding is blocked or the main is blocked and the siding is cleared, T_{X+1} may take the clear track and proceed to the interchange point without delay.

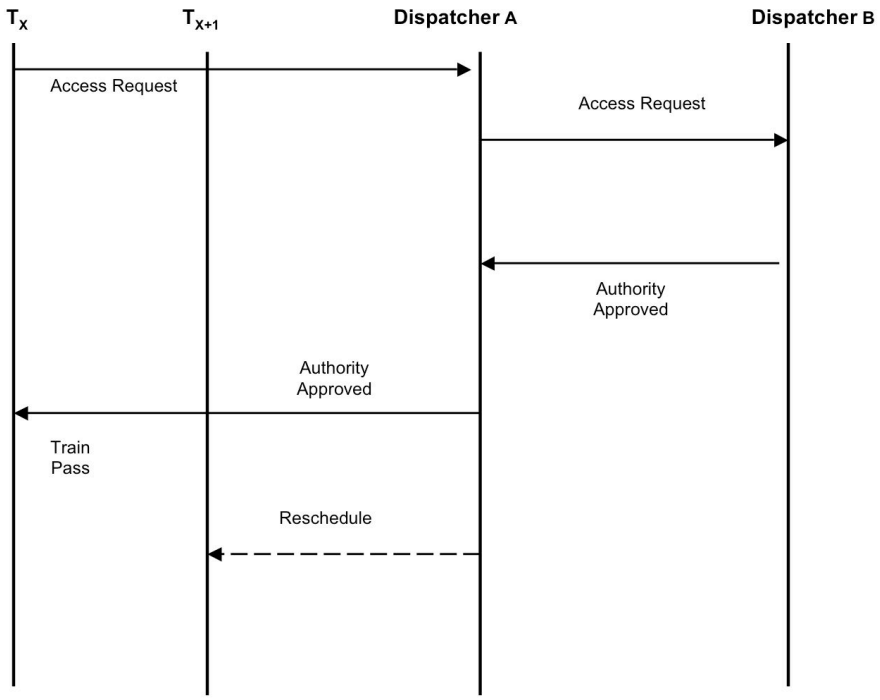

Figure 6.14: Lead Train Passes Security Check-Track Available

3. If the main and siding are blocked T_{X+1} may have to wait until the main or siding is clear in order to proceed to the interchange point.

In the later situation T_{X+1} can continue movement to the interchange point if the length of time it takes for T_X to receive their authority MA_A and move beyond the interchange point is less than the time it takes to stop T_{X+1} from T_{X+1}'s current velocity.

In the simplest case, where there is a single mainline running between domains A and B,

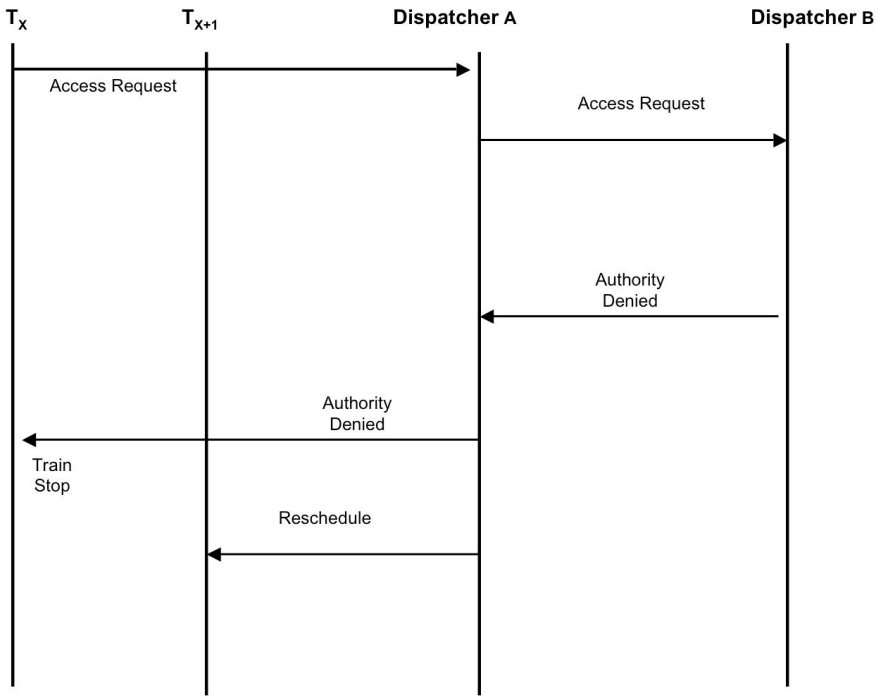

Figure 6.15: Lead Train Passes Security Check-Domain B Track Blocked

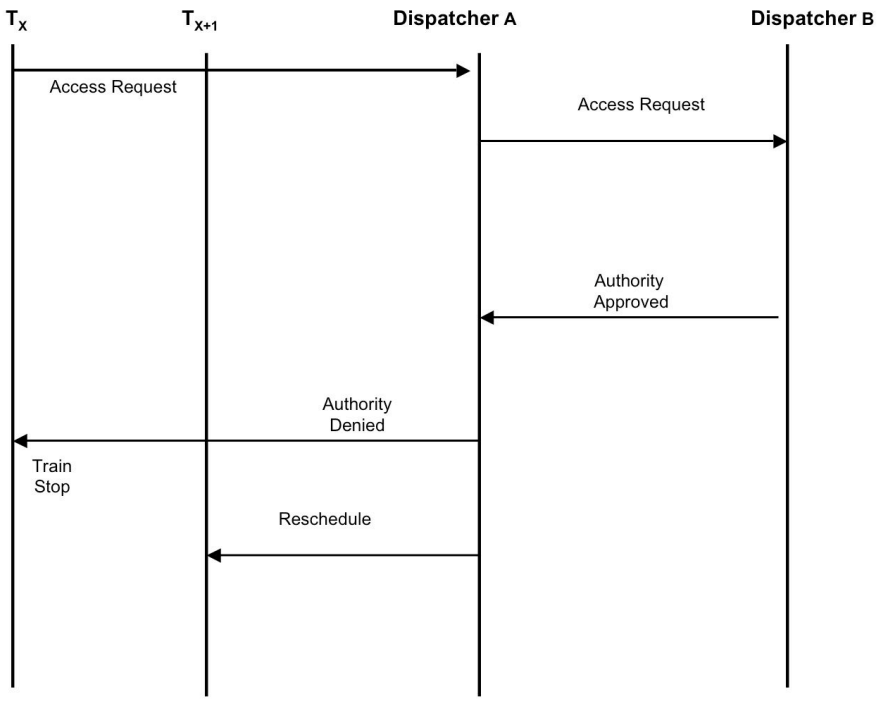

Figure 6.16: Lead Train Passes Security Check- Domain A Track Blocked

denial of entry of Train T_X will require rescheduling of the movement of subsequent trains T_{X+1},T_{X+2}....T_N. In order to preclude a train-to-train collision between the end of Train T_X with the head of train T_{X+1}, train T_{X+1} must receive notification of the requirement to stop before it proceeds beyond the safe stopping distance BD_{X+1}. If the movement of train T_{X+1} is not rescheduled, and train T_{X+1} does not stop before reaching the location of T_X, T_{X+1}, and T_X may collide. Also if the stopped train T_X is released to proceed into the next domain before the train T_{X+1}, reaches the safe stopping distance, a collision can be avoided.

The potential for a collision between train T_{X+1} and train T_X will be affected by the velocity of train T_{X+1}, the time of release of a stopped train T_X, the communication delays associated with information exchangs between CA_A, and CA_B, the dispatcher processing delays DS_A, and DS_B, as well as the PTC system processing times PTC_A, and PTC_B. The velocity V_{X+1} of train T_{X+1} directly affects the safe stopping distance BD_{X+1}. As V_{X+1} increases, the safe stopping distance BD_{X+1} increases, requiring greater separation of trains T_X and T_{X+1} to preclude a collision.

In general, delays of trains proceeding from A to B are prevented when the total delay time associated with certificate authentication and movement authorization is less than the time required to stop the train. If the former is less than the later, then the dispatcher is able to pass the appropriate authorizations to an on-coming train sufficiently in advance of the required safe stopping distance to enable the oncoming train to pass at speed.

Chapter 7: SAFETY OF THE INTEGRATED MODEL

The consequences of the delay imposed by the overhead can be demonstrated in terms of their effect on train movements. Train movements are authorized in terms of blocks. A block on a railroad is a predefined segment of track. Entry into a block may be authorized either by the signal aspect or the verbal authority of the dispatcher. In order to prevent delays, either the siding or the main track must be cleared prior to the arrival of a following train. Prevention of a collision, assuming moving blocks, requires that the delays for a train occupying either a siding or mainline block and the clearance time for the train to clear the block must be less or equal to the time it takes for a following train to brake to a zero velocity. Their moving block varies with the position of the train, and their length BV_L is varies as a function of the train's length TL_L and braking distance BD_V associated with the trains velocity at any given point in time.

$$BV_L = TL_L + BD_V \tag{7.1}$$

As BD_V goes to zero, the block length BC_L equals the length of the train TL_L. This compared to fixed block systems. Fixed blocks are segments of tracks whose length BF_L is a constant value between two geographically fixed boundary points BF_{UPPER} and BF_{LOWER}.

$$BF_L = BF_{UPPER} - BF_{LOWER} \tag{7.2}$$

There is a critical difference between block length (BF_L or BV_L) and limits of an authority. The limit of authority is a fixed geographical position to which the dispatcher has authorized a train to proceed. In the case of moving blocks, this position is a particular milepost

or other fixed geographic marker. In the case of fixed blocks, this position may the entry point to the block, the exit point to the block, or any location between the block entry and exit point. The limit of authority is specified by the dispatcher, and allows the dispatcher to control the location of the trains under their control.

Stopping distances and times for train have been extensively studied (for example [87–89]). Commercial tools to calculate this information using more complex models are known to exist, most notably the RailSim Train Performance Calculator (TPC) by Systra Consulting, and the Train Operation and Energy Simulator (TOES) by the Association of American Railroads. These estimators reflect a railroad's operating philosophy, the type of train (for example passenger or freight), the mass and its distribution of the train, the gradient of the territory the train is operating on at the time of braking, the crews reaction time, and the type of braking (full service, dynamic, or emergency) and the associated deceleration rate induced by the brakes. The inclusion of these types of additional factors in the braking calculations represents valuable intellectual property for the railroad, limiting access to these tools. The work I present here is only a first order approximation and excludes variations in the types of cars (i.e. tank, box, railrider, etc), variations in the methods and type of braking (emergency or dynamic, conventional air or electronic pneumatic), track profile (grades and curves), behavior of the locomotive power based on track conditions, details of consist loading and position in the consist of power (head end, middle, or pushing).

7.1 Characteristics of Railroads A and B

The two railroads A and B described algorithmically previously when modeled have the following characteristics:

1. Different public key based trust management infrastructures, STM_A and STM_B. The impact of this assumption is that keys shared between a locomotive of one company Cs train with company Ds crew with STM_A cannot be used when seeking entry to region owned by B

2. Trains may poses different braking distances (BD_X). The time required to obtain the necessary wayside, office, or onboard information, processes it, and promulgate the results to the required entities may differ, and may exceed the time between transmission of data sets. Stopping distances for train have been extensively studied (for example [87–89]). Commercial tools to calculate safe braking distances are available, most notably the RailSim Train Performance Calculator (TPC) by Systra Consulting, and the Train Operation and Energy Simulator (TOES) by the Association of American Railroads. These tools are proprietary and have extremely expensive licensing fees, well beyond the funding available for this research effort. As a consequence, worst case estimates using available formula are used.

3. Separate dispatch and scheduling systems exist in each of their respective domains A and B. Depending upon the delays encountered by the train that seeks entry to Bs territory, As scheduler needs to be informed of the impending delay. In response, A's scheduler needs to inform the trains approaching the interchange point to readjust their speed and position.

4. Separate communications infrastructures exist in each of their respective domains A and B. Cross-protocol communication is possible. The availability of a communication infrastructures using TCP/IP is assumed and is a realistic assumption because currently most railroad wireless and wire-line communication is migrating to TCP/IP.

5. Each train T_X has a velocity V_X and a braking distance BD_X that is as a function of velocity.

6. A single T_X occupies length (BLX_A and BLX_B) in their respective territories. We will assume the length of the blocks will vary based on train speeds and safe braking distance. This behavior is known as "moving blocks", and yields the most effective use of limited track resources. In a moving block, trains do not have to be separated by fixed block distances.

7. Each T_X requires a unique movement authority (MA_X) prior to entering into a block BLX_A or BLX_B. Any T_X that does not receive a valid MA_X may continue to reaching the safe BD_X and then must stop at the last milepost boundary specified in the last valid MA_X. The train may not proceed past the milepost until receipt of a valid MA_X. This is an axiom of safe railroad operation.

8. Any T_X that does not receive a valid MA_X and is a distance less than BD_X must immediately take action to stop forward movement of the train.

9. Each train T_X has an associated E_X, the engineers certificate, L_X the locomotive certificate.

7.2 Physics of Braking and Accelerating Trains

The approximate estimate for time to stop assumes constant deceleration in ideal track conditions (i.e. straight (no curvature), level (no up or down grade) track, and dry. It reflects the same variables (train length, train mass, braking efficiency, target speed, gradient, and distance to target) in [90] to predict braking distances for the European Train Control System (ETCS) system and the predictive braking curves based on the International Union of Railways (UIC) 546 standard [91]. A similar standard is under development by the IEEE [92]. Additional work on braking curves can be found in [93–99]. These estimates also assume that all cars in a particular consist are identical and have similar braking characteristics. Likewise, the time to clear a block assumes constant acceleration in ideal track

conditions, with identical lococmotives.

7.2.1 Time to Clear T_X

Assuming constant acceleration from an initial velocity of 0, the time to clear TC_X the interchange point (in seconds) is

$$TC_X = \sqrt{\frac{(2)(L_X)(M_X)}{(\frac{(375)(F_X)}{V_X}) - (M_X)(R_A)}} \tag{7.3}$$

R_A is estimated using the Davis equation. First developed in the mid 1920's, and modified in the late 1970's, it provides an estimate of the rolling resistance in pounds per ton [92].

$$R_A = (0.6 + \frac{20}{w_X} + (0.01)(V_X) + \frac{(K_a)(V_X)^2)}{(Car_X)(w_X)(n_X)}) \tag{7.4}$$

where

M_X is the weight of the train T_X (tons)

L_X is the length of the T_X (Ft)

V_X is the final velocity of T_X (mph)

F_X is the tractive force of T_X locomotives (HP)

R_A is the drag of the consist when accelerating (lb/ton)

w_X is the weight per axle per consist car in T_X (tons)

n_X is the number of axles per consist car in T_X

Car_X is the number of cars in the consist in T_X

K_a is the acceleration drag coefficent. $\text{K}_a = 0.07$

The tractive force F_X is given by

$$F_X = (N_{Loco})(HP)(E) \tag{7.5}$$

where

N_{Loco} is the number of locomotives in T_X

HP is the Horsepower per locomotive in T_X

E is the locomotive efficiency %

7.2.2 Time to Stop T_{X+1}

Assuming constant deceleration, the time to stop TS_{X+1} (i.e. final velocity $\text{V}_{X+1} = 0$)in seconds is

$$TS_{X+1} = \frac{(0.04583)(M_{X+1})(V_{X+1})}{F_{X+1} + R_D} \tag{7.6}$$

and the drag R_D of T_{X+1} is given by

$$R_D = (M_{X+1})(0.6 + \frac{20}{w_{X+1}} + (0.01)(V_{X+1}) + \frac{(K_b)(V_{X+1})^2)}{(Car_{X+1})(w_{X+1})(n_{X+1})}) \tag{7.7}$$

where

M_{X+1} is the mass of the train T_{X+1} (tons)

V_{X+1} is the initial velocity of T_{X+1} (mph)

F_{X+1} is the braking force of consist T_{X+1}

R_D is the drag of the consist T_{X+1} when decelerating

w_{X+1} is weight per axle per consist car in T_{X+1}

n_{X+1} is the number of axles per consist car in T_{X+1}

K_{b2} is the braking drag coefficent. $K_b = 1.4667$

Car_{X+1} is the number of cars in the consist in T_{X+1}

The braking force F_{X+1} is given by

$$F_{X+1} = (Car_{X+1})(CarWeight_{X+1})(BF)(Brake_{Avail})(2000) \tag{7.8}$$

where

Car_{X+1} is the the number of cars in the consist T_{X+1}

$CarWeight_{X+1}$ is the weight of a car in the consist T_{X+1} (tons)

BF is the brake ratio (5%)

$Brake_{Avail}$ is the % operable brakes

7.3 Consist Delay and Safety

Safe operation of the railroad requires that any Train T_{X+1} not run into the preceeding Train T_X. For this safety criterion to occur the consist delay between Train T_X and T_{X+1}

must satisfy the equation.

$$ConsistDelay + TC_X \leq TS_{X+1} \tag{7.9}$$

Substituting the equations 7.3 and 7.6 into 7.9 delay equation and solving for the delay yields the maximum delay that between two trains T_X and T_{X+1}.

$$ConsistDelay < \frac{(0.04583)(M_{X+1})(V_{X+1})}{F_{X+1} + R_D} - \sqrt{\frac{(2)(L_X)(M_X)}{(\frac{(375)(F_X)}{V_X}) - (M_X)(R_A)}} \tag{7.10}$$

where

$$R_A = (0.6 + \frac{20}{w_X} + (0.01)(V_X) + \frac{(K_a)(V_X)^2)}{(Car_X)(w_X)(n_X)}) \tag{7.11}$$

$$R_D = (M_{X+1})(0.6 + \frac{20}{w_{X+1}} + (0.01)(V_{X+1}) + \frac{(K_b)(V_{X+1})^2)}{(Car_{X+1})(w_{X+1})(n_{X+1})}) \tag{7.12}$$

On a track that is operating at maximum capacity the relationship ConsistDelay $\leq T_{STOP}$ - T_{CLEAR} continues to hold between sequentially ordered trains. At maximum capacity, the movement of a train from one location to the next requires that the lead train clear the location it is occupying before the trailing train can stop in the location just cleared. This is no different than the case of advancing through the interchange point, the interchange point is simply a special case of a block boundary. Instead of being the boundary between two adjacent blocks in the same domain, it is simply the boundary between two adjacent blocks, one of which is one domain, the other of which is a second domain. If trains T_X and T_{X+1} occupy the main and siding, subsequent trains T_{X+2} through T_{X+N} are blocked from advancing since the trains are restricted to a single degree of motion along the track. The results of this section, derived from the physics of train movement, and the results of Chapter 5, based solely on the trust management system can be combined into a single

equation. The right hand of the inequality is equation 7.10, while the left hand side is the time delay due to padding, propogation, and processing delays. (Equation 5.2) divided by the communications transmissionrate (TR) plus the system response time ($\mathrm{SYS}_{ResponseTime}$) and the operators response time ($\mathrm{OP}_{ResponseTime}$).

$$\frac{B_{SenderAddress} + B_{ReceiverAddress} + P_{Information} + C_{Data}}{TR} +$$

$$\frac{C_{Padding} + S_{Data} + S_{Padding}}{TR} +$$

$$\frac{SYS_{ResponseTime} + OP_{ResponseTime} + SYS_{Propagation}}{TR} <$$

$$\frac{(0.04583)(M_{X+1})(V_{X+1})}{F_{X+1} + R_D} -$$

$$\sqrt{\frac{(2)(L_X)(M_X)}{(\frac{(375)(F_X)}{V_X}) - (M_X)(R_A)}} \tag{7.13}$$

where $\mathrm{B}_{SenderAddres}$, $\mathrm{B}_{ReceiverAddress}$, $\mathrm{P}_{Information}$, C_{Data}, $\mathrm{C}_{Padding}$,S_{Data}, $\mathrm{S}_{Padding}$ defined in equation 5.2), M_{X+1}, M_X, V_X, V_{X+1}, L_X, L_{X+1}, L_X, L_{X+1}, R_D, R_A, F_X, and F_{X+1} defined from equations 7.6 and 7.3, and

TR is the communication tranmission rate

$\mathrm{SYS}_{ResponseTime}$ is the length of time it takes for the system to process the data once recieved and change it into information

$\mathrm{OP}_{ResponseTime}$ is the length of time it takes for the operator to respond to a command once received

$\mathrm{SYS}_{Propagation}$ is the propagation delay for the communications medium

$SYS_{ResponseTime}$ is a function of the performance characteristics of the office subsystem, wayside subsystem, and the onboard subsystem involved in a particular message exchange. $OP_{ResponseTime}$ is a function of human factors behavior in receiveing, processing, and executing a received command. The advantage of establishing this single safety equation relating all elements is that it allows for the designer to develop risk based performance budgets for the various elements in their design. As long as the overall equation remains true, the designer is free to experiment with various options to achieve the required performance at a particular cost point.

7.4 An Illustrative Example

The behavioral characteristics of the railroad vary greatly depending upon the operating parameters of the trains operating along the railroad. Finding the optimal combination of train parameters that minimizes ConsistDelay is a complex problem in operations research. The following example, however, illustrates the use of these equations. For the purposes of this example we will assume T_X and T_{X+1} are identical with properties as follows:

- Number of Locomotives = 3
- Length of locomotive = 100 feet
- Horsepower per locomotive = 4500 HP
- Weight per locomotive = 200 tons
- Locomotive Efficiency = 95%
- Number of Cars = 100
- Weight of a Car = 60 tons

Table 7.1: Time for T_X to Accelerate and Clear Track

Velocity (mph)	Time (seconds)
10	17.05
20	24.48
30	30.53
40	36.00
50	41.28
60	48,57

Table 7.2: Following Train Stop Time

Velocity (mph)	Time (seconds)
10	11.27
20	33.51
30	33.69
40	44.81
50	55.86
60	66.82

- Length of a Car = 100 feet
- Braking Efficiency = 5%
- Axles per Car = 2
- Percent of Brakes Operable = 85% (Minimum operating brakes allowed by Federal Regulations)
- Train Length = 10300 Feet
- Communications Bandwidth = 4800 bps

All braking is provided by consist cars, locomotive dynmanic braking is not considered. The estimated delay that can be accepted is listed in Table 7.3 Negative numbers indicate that a collision can occur and that a Signal Passed at Danger (SPAD) event has occurred. Train T_X will not have cleared the interchange point before Train T_{X+1} arrives. As can

Table 7.3: Allowable Delay: $V_X = V_{X+1}$

Velocity T_X mph	Velocity $T_X + 1$ mph	Clearance Time T_X secs	Stop Time T_{X+1} secs	Max Delay Time
10	10	17.05	11.27	-5.78
20	20	24.48	22.51	-1.97
30	30	30.53	33.69	3.17
40	40	36.00	44.81	8.81
50	50	41.28	55.86	14.58
60	60	46.57	66.82	20.25

be seen in Figure 7.1, at speeds below roughly 21 miles pr hour, a following train T_{X+1} will always overtake a leading train T_X. Above that speed, leading trains T_X can clear the track they occupy before the arrival of following train T_{X+1}.

An alternative way to view combinations of leading train clearance time, and following train stopping time is with a radar chart (7.2). In this chart, the spokes represent locomotive speeds, the rings represent clearance times in seconds. As can be seen, for the example configuration, in almost all cases, the time for a leading train to clear the block is less than the time it takes to stop the following train and some delay can occur without adversely impacting subsequent train movements. As the velocity V_X of T_X increases relative to V_{X+1} of train T_{X+1}, Train T_{X+1} takes less time to decelerate to a stop than it takes train T_X to accelerate to clear the interchange point. (Table 7.4). Because the available horsepower of T_X remains unchanged, the length of time it takes to reach the final velocity V_X remains unchanged. However, as the velocity V_{X+1} of train T_{X+1} decreases, the time it takes to bring train T_{X+1} to stop decreases. Similarly, as the velocity V_X of T_X decreases relative to V_{X+1} of train T_{X+1}, it takes less time to accelerate T_X, and T_X clears the interchange point faster than it takes to bring T_{X+1} to stop. This is shown in Table 7.5.

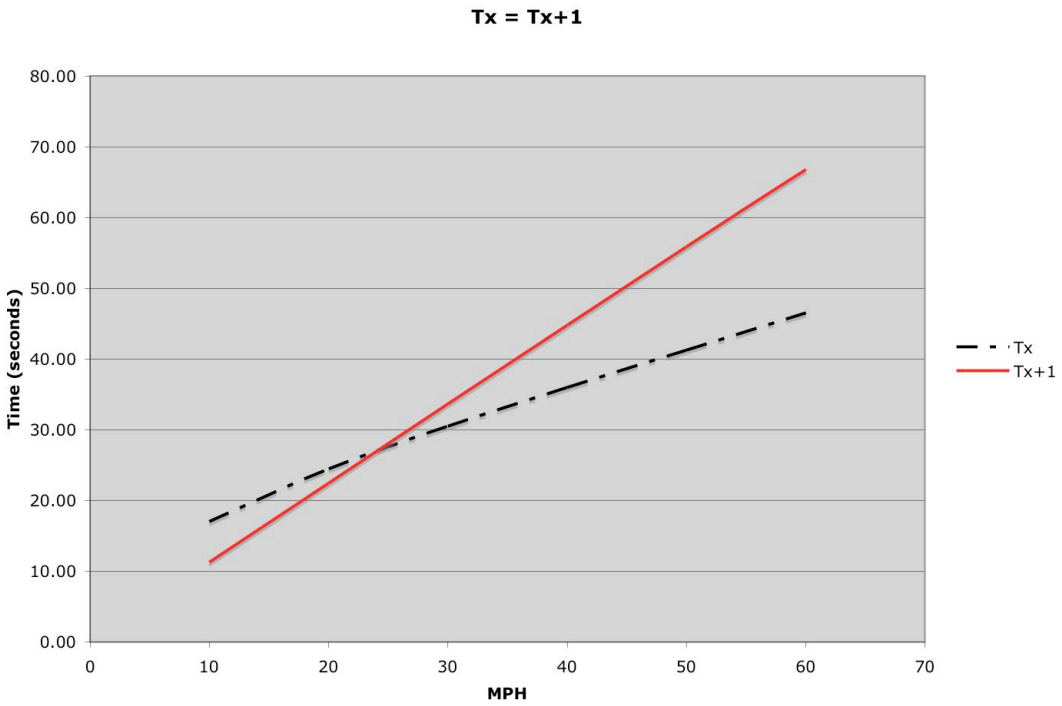

Figure 7.1: Clearance & Stopping Time: $V_X = V_{X+1}$

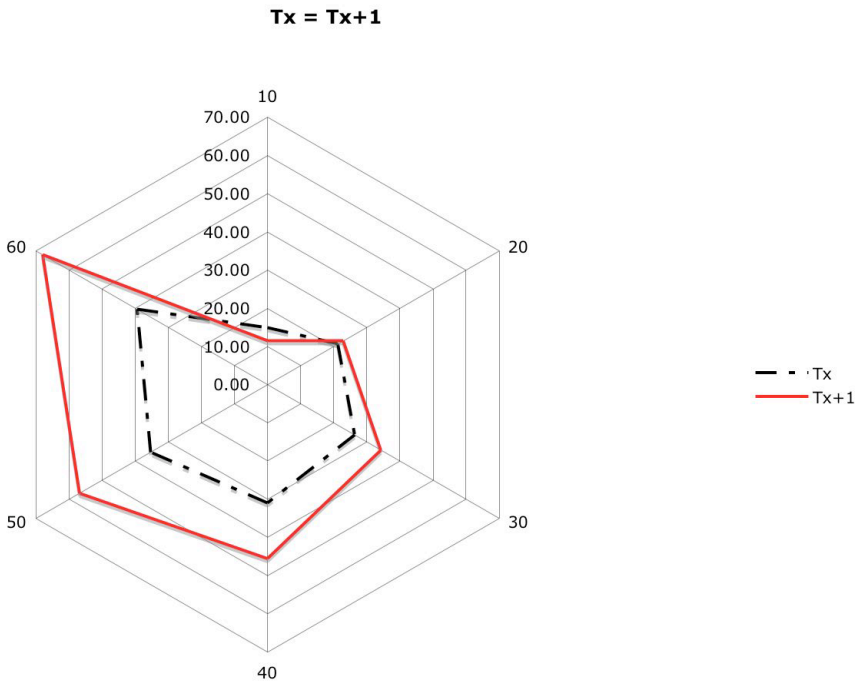

Figure 7.2: Clearance and Stopping Coverage: $V_X = V_{X+1}$

Table 7.4: Allowable Delay: $V_X > V_{X+1}$

Velocity T_X	Velocity $T_X + 1$	Clearance Time T_X	Stop Time Time T_{X+1}	Max Delay Time
10	5	17.05	5.64	-11.41
20	15	24.48	16.90	-6.93
30	25	30.53	28.11	-2.42
40	35	36.00	39.26	3.26
50	45	41.28	50.35	9.07
60	55	46.57	61.35	14.78

Table 7.5: Allowable Delay: $V_X < V_{X+1}$

Velocity T_X	Velocity $T_X + 1$	Clearance Time T_X	Stop Time Time T_{X+1}	Max Delay Time
5	10	11.98	11.27	-0.71
15	20	21.04	22.51	1.47
25	30	27.61	33.69	6.09
35	40	33.31	44.81	11.51
45	50	38.65	55.86	17.21
55	60	43.91	66.82	22.91

By decreasing the tractive effort to accelerate T_X relative to the braking force of T_{X+1}, the clearance time of T_X increases relative to the braking time of T_{X+1}. In Table 7.6 the tractive effort available is provided by a single 4500 HP locomotive (as opposed to 3 4500 HP locomotives). As a consequence, the time it takes to accelerate T_X to V_X increases, and the maximum allowable delay time decreases.

Train length also affects the allowable delay time, because the shorter trains or the trains with less braking capability take less time to clear the interchange point. For example, if the length of T_X and T_{X+1} decreases (for example from 100 cars to 75 cars) and the number and tractive effort of the locomotives stays the same (3 locomotives @ 4500 HP each) both

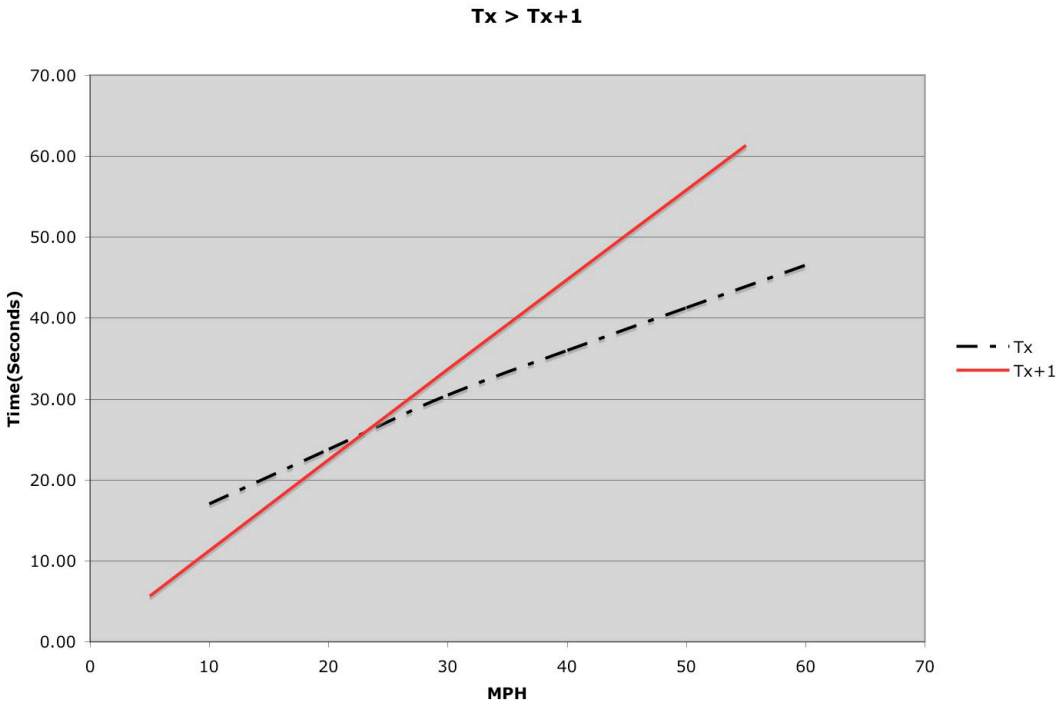

Figure 7.3: Clearance & Stopping Time: $V_X > V_{X+1}$

Table 7.6: Allowable Delay: HP T_X < HP T_{X+1}

Velocity T_X	Velocity $T_X + 1$	Clearance Time T_X	Stop Time Time T_{X+1}	Max Delay Time
10	10	29.39	11.27	-18.12
20	20	43.55	22.51	-21.04
30	30	56.66	33.69	-22.97
40	40	70.81	44.81	-25.99
50	50	88.12	55.86	-32.26
60	60	112.68	66.82	-45.86

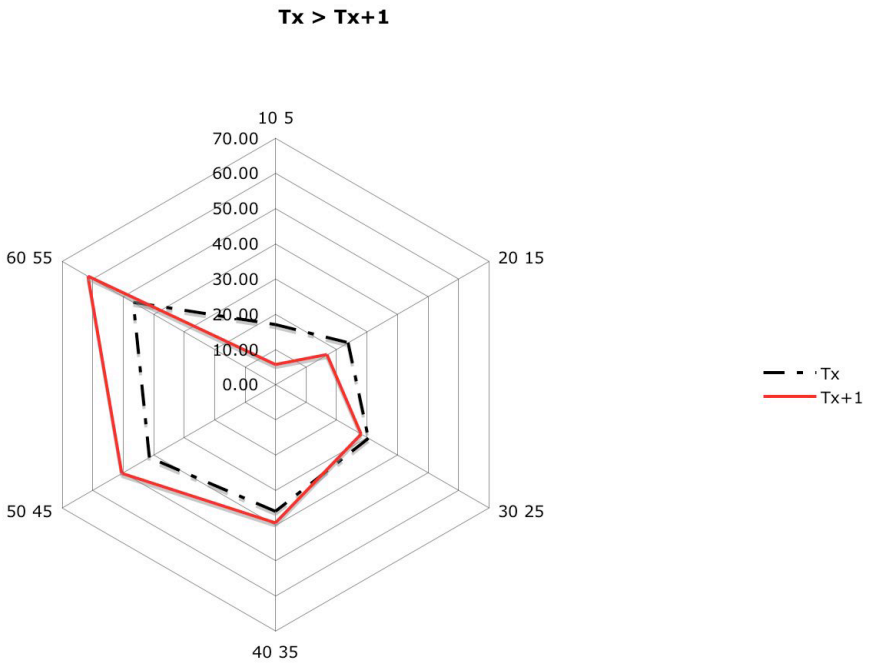

Figure 7.4: Clearance and Stopping Coverage: $V_X > V_{X+1}$

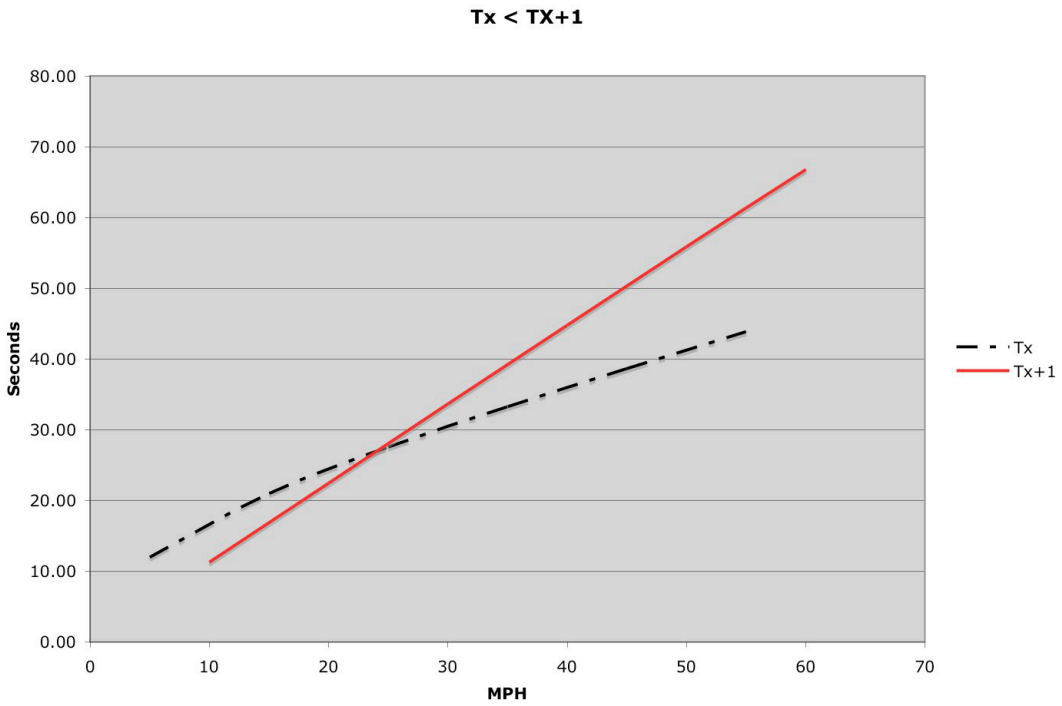

Figure 7.5: Clearance & Stopping Time: $V_X < V_{X+1}$

Figure 7.6: Clearance and Stopping Coverage: $V_X < V_{X+1}$

Table 7.7: Allowable Delay: 75 Car Consist

Velocity T_X	Velocity $T_X + 1$	Clearance Time T_X	Stop Time Time T_{X+1}	Max Delay Time
10	10	14.86	11.48	-3.38
20	20	21.26	22.92	1.68
30	30	26.39	34.30	7.91
40	40	30.96	45.60	14.63
50	50	35.27	56.80	21.54
60	60	39.48	67.90	28.43

Table 7.8: Allowable Delay: 125 Car Consist

Velocity T_X	Velocity $T_X + 1$	Clearance Time T_X	Stop Time Time T_{X+1}	Max Delay Time
10	10	19.01	11.15	-7.86
20	20	27.40	22.26	-5.13
30	30	34.32	33.33	-0.99
40	40	40.71	44.34	3.63
50	50	47.00	55.29	8.29
60	60	53.50	66.17	12.67

the time it takes T_X to clear the interchange point decreases, and the time to brake T_{X+1} decreases as shown in Table 7.7. Likewise if the length of T_X and T_{X+1} increase (for example from 100 cars to 125 cars) with the same number of locomotives (3 @ 4500 HP each), the length of time it takes T_X to clear the interchange point increases and the braking force increases, with the net result illustrated in Table 7.8.

7.5 The Impact of Aggregate Communications Overhead and Delay

The allowable delays previously calculated are based on the physical characteristics of the locomotive and its' consist as well as the communications bandwidth (4800 bps) available to

exchange data . However, as shown in Chapter 5 the trust management system introduces additional delays for the authentication process and which vary based on the tranmission rate, propagation, and processing time. The worst-case scenario occurs as a consequence of initial authentication of the actors and the first message exchanged. To obtain the total time for a consist to clear, or a consist to stop, the communications overhead times T_{OH} 5.5, must added to the time to clear of T_X and time to stop T_{X+1}. Provided T_X and T_{X+1} require the same length of time to authenticate (i.e. T_{OH} is a constant for train T_X or train T_{X+1}), the delay T_{OH} cancels out and the delay between individual trains (T_X and T_{X+1}) remains the same as previously calculated.

The assumption that there are no authentication or communications delays is, however, unrealistic. Even in a benign environment, communications disruptions may occur as a consequence of phenomena such as normal atmospheric interference, electromagnetic interference by the AC or DC generators onboard the locomotive, or physical items such as buildings or foliage. To ensure that collisions between a leading train T_X and a following train T_{X+1} do not occur, the authentication and the communications delays $T_{COMMDELAY_TX}$ associated with train T_X must be less than the communications delays $T_{COMMDELAY_TX+1}$ associated with train T_{X+1}. If the difference in communications delays is greater than the allowable delay between T_X and T_{X+1}, then the potential exists for the trains to collide. The situation where the communications delay is greater than the alllowable delay between T_X and T_{X+1} results in a condition referred to as SPAD (Signal Passed at Danger). SPAD occurs when a following train fails to stop at a red signal that is providing protection for the train in the preceeding track segment.

PTC system designs assume that communications disruptions are likely to occur. To mitigate against this eventuality, not only are the commands retransmitted several times to

Table 7.9: Delay and Approximate Separation: $V_X = V_{X+1}$

Delay	Separation Distance
-5.78	Collision
-1.97	Collision
3.17	139 feet
8.81	487 feet
14.58	1062 feet
20.25	1782 feet

ensure receipt and acknowledgement, each transmitting and receiving device is equipped with a timer. In the event of a communications disruption that precludes receipt of a valid message, a timer on the device will expire, forcing the device to its most restrictive safe state. This timer increases the time delay by 9 to 15 seconds, it ensures the safety of following trains, albeit with a decrease in system throughput.

To better place the impact of delay on rail operations, we will estimate the separation distance between a train T_X and T_{X+1} operating at the same velocity. An approximation of the distance between trains is shown in Table 7.9. With trains T_X and T_{X+1} operating with under condition of nearly simultaneous movement authorities (a method of operation known as moving block and a capability made possible with both trains being equipped with PTC), the required train separation is significantly less than if train movements were not simultaneously. With the moving block method of operations, the separation between trains moving at 60 mph can be as low as roughly 3/10th of a mile. When contrasted to the roughly 1.1 miles required by fixed blocks, the traffic density can be increase by roughly a factor of three. This makes significantly better use of the available track resources, and increases system throughput.

Chapter 8: SUMMARY

The objectives of this dissertation were to prodice:

1. a safety and security protocol to schedule PTC equiped trains passing through an interchange between two domains,
2. minimize traffic delays and maximize system velocity.
3. amidst communication based attacks on different dispatch and scheduling systems with different communications systems in the presence of malactors

8.1 Attainment of First Objective

An algorithm for the safe and secure scheduling of trains through the interchange point between two domains has been defined in Chapter 6. The algorithm supports prevention of train-to-train collisions under a worst-case traffic density scenario. This is a scenario that is representative of rail operations found on the Powder River subdivisions of both the Union Pacific and BNSF Railways. These railroads operate their main tracks at maximum density around the clock. The algorithms presented also minimizes the chance of a signal passed at danger, with it's potential for tail end- to head end collisions bewtween two trains operating on the same line.

8.2 Attainment of Second Objective

The OTAR based trust management system demonstrated in Chapter 5, and its' associated performance characteristics, provides an example of the establishment and use of a secure trust management system that supports attainment of system safety and security. After discussing the basic architectures of Communication Based Train Control (CBTC) Supervisory Control and Data Acquisition (SCADA) Systems (Chapter 3 and their relationship to current train control methods (Chapter 2), the dissertation has highlighted the key security weaknesses associated with wireless CBTC/PTC systems. In the process of highlighting these weaknesses, we have specified requirements that CBTC/PTC systems must posses for safe and secure operations under different types of malactor attacks (Chapter 4). By illustrating the similarity of different CBTC/PTC systems (Chapter 3), we have shown that a common set of security management system requirements exists (Chapter5). This common set of requirements allows individual railroads to select different trust management systems best suited to their business needs, as well as provide a basis for systems interoperability. Interoperability supports the free exchange of rail equipment between railroads, improving asset utilization. The improved asset utilization increases system throughput by eliminating the need for time consuming exchanges of crews and locomotives as consists move between railroads. Since system velocity is directly related to the number of trains able to pass through the system in a specific amount of time, increasing throughput improves the system velocity.

8.3 Attainment of Third Objective

The dissertation has also demonstrated that interoperable PTC systems can be communications and dispatch systems independent (Chapter 6). Any communications system with

acceptable bandwidth and latency, appropriate to the railroads operational characteristics, is acceptable. As long as the CBTC/PTC system and associated trust management support the required intra-domain security and traffic-scheduling constraints, and sufficient track space is available to allow cross-domain traffic movement, rail operations can continue regardless of the specific dispatching systems utilized. In the example provided in the dissertation, two entirely different dispatch systems, with different operating and performance characteristics, have illustrated the capability from domain A to domain B delays in the authentication process cannot only delay the granting of the movement authority by the respective domain dispatchers but the subsequent authority releases to the trains involved. This in turn can delay the subsequent scheduled movement of trains, resulting in increasing traffic delays. This dissertation has defined two cooperating dispatchers whose behavior will support the movement of trains, and prevent without Signals Passed at Danger, when the track is saturated (Chapter 6).

8.4 General Applicability to Current Systems

The work I have presented is generally applicable to most of the PTC systems currently under development or implemented in the US. Table 8.1 lists each of the PTC systems currently under development, or deployed, in the US. An "X" in the column means that system, as currently designed or implemented, provides that PTC functionality. A "Y" in the column indicates

- That the trust methodology and scheduling methodology presented in this work can be used to address the tactical scenario of unidirectional interchange traffic with a single siding for the listed system.

Table 8.1: Method Applicability

System Name	PTC Level 1	PTC Level 2	PTC Level 3	PTC Level 4
ACSES	XY			
ITCS	XY	XY	XY	
ETMS 1	XY	XY		
ETMS 2	XY	XY	XY	
VTMS	XY	XY	XY	
OTC	XY	XY	XY	
Train Sentinel	XY	XY		
CBTM	XY	XY		
Chicago Metra	Cross Over Only			
CAS	XY	XY	XY	
NAJPTC	XY	XY	XY	XY

- That an extension of our work to include more complicated track geometries and consist makeup's would be applicable for evaluating scheduling and trust management system performance for the listed system.

8.5 Future Work

Like Intelligent Transportation System (ITS) equipped automobiles operating on limited access highways communicating wirelessly with different Traffic Control Centers operating disperse trust management systems, CBTC equipped trains are restricted to a single degree of freedom operations communicating wirelessly with different Dispatching Centers. Any delay of a train at an interchange point as it crosses from the operating domain of one railroad to the operating domain of another has the potential to delay the movement of subsequent trains operating along the same line to the same interchange point, minimizing traffic delays.

Developing a model for a more general secure cross-domain authentication and authorization scenario supports the interoperation of multiple railroads. As shown in Chapters 6 and 7, the model provides for computing potential delays. This provides a basis for evaluating the suitability of different implementations and the associated communications systems. This work has been limited to the tactical scenario of a unidirectional movement between two domains using a single authentication and authorization approach (OTAR), assuming traffic saturation on a single mainline track with siding to the interchange point. This work needs to be expanded to account for bidirectional train movements, multiple mainline tracks and sidings, and strategic, as well as tactical scheduling and routing.

Through Use-Misuse Case Analysis, we have identified ways in which a mal-actor can prevent the functional objectives of a PTC system from being enforced as designed. These identified impacts on the PTC Use Cases can be used to enhance their design so that enhanced systems are resilient to Misuse Cases. However a shortcoming of Use-Misuse Case analysis is a lack of support for quantitative analysis, a critical issue in ongoing studies of PTC systems. Combining Use-Misuse Case analysis to identify ways in which a mal-actor can prevent the functional objectives of a system from being enforced, as well as quantitatively and qualitatively evaluate system failure modes, is an area open for additional study and formalization. As shown, failure in timing delays can result in significant adverse consequences to safe railroad operations. Although we have used Use/Misuse Cases as the basis for showing that loss of Quality of Service (QoS) (in our case timing) can result in adverse safety consequences, we have not related these consequences or mitigations to establish levels of risk. Establishing these relationships is essential to determine the optimum use of limited resources for best improving communications and PTC subsystem performance continues to remain an open research area. A closed form solution for determining the optimal combination of resources is unlikely, making statistical evaluation of open form

solutions necessary and is a subject of future research.

By supporting various implementations of a standard PTC architecture with differing levels of functionality, individual railroads, and the railroad industry can develop effective, economical, and interoperable train control technology that can serve the interests of safety and other intelligent transportation systems. However, in order to ensure the maximum effect of these enhancements, additional work needs to be undertaken to address the economic impacts of various security requirements of PTC associated with other possible alternative techniques and technologies. The entire area of security network management for PTC systems requires further study, both in terms of policy and technology. Potential interoperability policy and technical issues must be resolved. There are also a number of implementation related issues that have not been fully addressed in this work. In a operational environment where rail trafic is heavy and close together, the volume of operational and environmental data that must be transmitted may exceed the communications bandwidth. The required capabilities can only be determined in the context of the railroads operating environment and the particular implementation mechanisms.

We have proposed a rudimentary system that uses distributed trust management to ensure distributed authentication and authorization and OTAR for online key exchanges. They result in timing and processing overheads that need to be considered during the design stage of such a system. Designing an effective security solution to PTC requires analyzing its strength, performance and cost against potential risk. If appropriately chosen, and when considered in light of organizational and environmental factors, a combination of managerial, operational, and technical controls can synergistically work together to ensure safe and secure interoperable PTC systems. Ongoing research in this, and related PTC security requirement specification, will provide us with sufficient data for a detailed system design

and cost evaluation.

Basic PTC systems, while they are economically unviable in terms of their safety case alone, may, when combined with other advanced technologies such as train pacing systems, electronically controlled pneumatic brakes, locomotive health monitoring, or integrated Intellegent Transportation System Vehicular Adhoc Network (VANET) highway grade crossing activation and warning systems, offer significant business and societal benefits that make PTC system installation more economically viable [108]. This is critical, because the installation of PTC systems has been recently mandated [109] on all passenger rail lines, all mixed passenger/freight rail lines, and all rail lines carrying Toxic by Inhalation (TIH) materials. The ability to generate additional railroad and societal benefits is required to eliminate any adverse impacts of this regulatory mandate, Success, however, will depend upon the ability to rely on the transmitted information, which will be a function of the security that can be provided. If successful, secure CBTC system installation will represent a revolutionary change from almost ninety years of train and signal control that will allow the complete replacement of traditional signaling systems.

Biblography

Bibliography

[1] M. Hartong, R. Goel & D. Wijesekera, *Communications Based Positive Train Control Systems Architecture in the USA*,Proceedings of the 63rd Vehicular Technology Conference, VTC 2006-Spring Melbourne, Australia, May 2006.

[2] Federal Railroad Administration,*Railroad Communications and Train Control*,Report to Congress, Washington DC, July 1994.

[3] Railroad Safety Advisory Committee, *Implementation of Positive Train Control Systems*, Report of the Railroad Safety Advisory Committee to the Federal Railroad Administrator, August 1999.

[4] National Transportation Safety Board, *Positive Train Control Systems Symposium* March 2-3, 2005 Ashburn, Virginia.

[5] National Archives and Records Administration, *49 CFR Parts 209, 234, and 236 Standards for the Development and Use of Processor Based Signal and Train Control Systems; Final Rule*, Federal Register, 7 March 2005.

[6] Government Printing Office, *Public Law 110-432, Federal Railroad Safety Improvement Act of 2008.*

[7] Association of American Railroads, Policy & Economics Department, *Freight Railroad Statistics 2007*, Washington, DC, 2007.

[8] National Transportation Safety Board, *Collision of Norfolk Southern Freight Train 192 with Standing Norfolk Southern Local Train P22 with Subsequent Hazardous Material Release at Graniteville, SC January 6, 2005* NTSB # RAR-0504, November 2005.

[9] American Railway Engineering and Maintenance of Way Association, *Communications & Signaling Manual of Recommended Practices, Volume 4, Part 16.4.50,* American Railway Engineering and Maintenance of Way Association, Washington, DC, 2005.

[10] U.S. Surface Transportation Board, Office Of Economics, Environmental Analysis And Administration, *Statistics Of Class I Freight Railroads In The United States 2007*, Washington DC, 2007.

[11] National Transportation Atlas Databases (NTAD) 2003,*Federal Railroad Administration (FRA) National Rail Network 1:100,000 (line) 2003* ed., Bureau of Transportation Statistics (BTS) Washington DC.

[12] A. Billonnet, *Using Integer Programing to Solve the Train Platforming Problem* Transportation Science, Volume 37, Issue 2, May 2003.

[13] E. Petersen, *Over the Road Transit Time for a Single Track Railway*, Transportation Science, vol. 8, 1974.

[14] T. Carinc, J. Ferland, & J. Rousseau, *A Tactical Planning Model for Rail Freight Transportation* Transportation Science vol. 18 1984

[15] W. Sutewong, Doctoral Disertation *Algorithims for Solving the Train Dispatching Problems for General Networks* University of Southern California, August 2006.

[16] National Transportation Safety Board, *NTSB Most Wanted Transportation Safety Improvements 2006-2007*, Washington,DC , November 2007.

[17] Docket FRA-2003-15432,*Burlington Northern Waiver Petition*, US Department of Transportation Document Management System, http://regulations.gov/.

[18] R. Lederer, *Electronic Train Management System*, BNSF Railways Presentation at National Transportation Safety Board 2005 Symposium on Positive Train Control, Ashburn, VA http://www.ntsb.gov/events/symp_ptc/presentations/14_Lederer.pdf.

[19] R. Haag, *Electronic Train Management Systems* WABTEC Railway Electronics Presentation at National Transportation Safety Board 2005 Symposium on Positive Train Control, Ashburn, VA, http://www.ntsb.gov/events/symp_ptc/presentations/16_Haag.pdf

[20] *Waiver for Petition of Compliance*, Docket FRA 2002-11533, US Department of Transportation Document Management System, //regulations.gov/.

[21] R. Kollmar, *Michigan Positive Train Control Project Incremental Train Control System*, National Passenger Rail Corporation (AMTRAK) Presentation at National Transportation Safety Board 2005 Symposium on Positive Train Control, Ashburn, VA, http://www.ntsb.gov/events/symp_ptc/presentations/07_Kollmar.pdf.

[22] J. Baker, *ITCS Incremental Train Control System*, GE Global Signaling Presentation at National Transportation Safety Board 2005 Symposium on Positive Train Control, Ashburn, VA, http://www.ntsb.gov/events/symp_ptc/presentations/11_Baker.pdf.

[23] Public Transportation Association, *2007 Public Transportation Fact Book*, American Public Transportation Association, Washington, DC, May 2007.

[24] W. Banks & R. Barclay, *An Analysis of a Strategic Rail Corridor Network (STRACNET) for National Defense*, Military Traffic management Command, Washington, DC, Nov 1976.

[25] B. Weinstein & T. Clower, *The Impact of the Union Pacific Service Disruptions on the Texas and National Economies: An Unfinished Story*, Railroad Commission of Texas, February 1998.

[26] Congressional Research Service of the Library of Congress, *Cyber Attacks and Cyber Terrorism- Vulnerabilities and Policy Issues for Congress* Report RL32114, Washington DC, October 17, 2003.

[27] Policy and Economics Department, Association of American Railroads *Mandatory Hazmat Rerouting*, Association of American Railroads, Washington, DC, April 2007.

[28] Title 49 US Code of Federal Regulations Part 171.8.

[29] Office of Freight Management and Operations, Federal Highway Administration *Freight Facts and Figures 2006*, US Department of Transportation, Washington DC.

[30] M. Orr, *Public Health Risks of Railroad Hazardous Substance Emergency Events*, Journal of Occupational and Environmental Medicine, Vol. 43, 2001.

[31] National Capital Planning Commission, *Rail Realignment Feasibility Study Securing Freight Transportation in the National Capital Region*, NCPC Washington, DC, April 2007.

[32] Department of the Army, *Treatment of Chemical Agent Causalities and Conventional Military Chemical Agents*, Field Manual Army FM-285, Department of the Army, Washington, DC, Feb 1990.

[33] *Cyber Security of Freight Information Systems*, Transportation Research Board of the National Academy of Sciences, Washington, DC 2003.

[34] C. Chittester & Y. Haines,*Risks of Terrorism to Information Technology and to Critical Interdependent Infrastructure*, Journal of Homeland Security and Emergency Management, Volume 1 Issue 4, 2004 Berkley Electronic Press.

[35] Title 49 Code of Federal Regulations Part 172 Subpart F.

[36] Federal Railroad Administration, Office of Safety Analysis, *AccidentIncident Overview 2007*, US Department of Transportation, Washington DC.

[37] A. Carlson, D. Frincke,& M. Laude, *Railway Security Issues: A Survey of Developing Railway Technology* Proceedings of the International Conference on Computer, Communications, & Control Technology, International Institute of Informatics and Systemics, 2003.

[38] P. Craven, *A Brief Look at Railroad Communication Vulnerabilities*, Proceedings 2004 IEEE Intelligent Transportation Systems Conference Washington, D.C

[39] The Presidents National Security Telecommunications Advisory Committee Wireless Task Force Report *Wireless Security* January 2003

[40] United States General Accounting Office, GAO Testimony Before the Subcommittee on Technology Information Policy, Intergovernmental Relations and the Census, House Committee on Government Reform, *Critical Infrastructure Protection Challenges and Efforts to Secure Control Systems*, Tuesday, March 30, 2004.

[41] *Diversification Of Cyber Threats* Institute For Security Technology Studies At Dartmouth College, Investigative Research For Infrastructure Assurance Group, May 2002.

[42] *Information Assurance Technical Framework (IATF), Release 3.1*, Information Assurance Solutions, US National Security Agency Fort Meade, MD, September 2002.

[43] *Federal Information Security Management Act of 2002* (Public Law 107-347) December 2002.

[44] A. Menezes, P. van Oorschot, & S. Vanstone, *Handbook of Applied Cryptography 5ed*, CRC Press, August 2001.

[45] M. Blaze, J. Feigenbaum, & J. Lacy, *Decentralized Trust Management* Proceedings of the IEEE Symposium on Security and Privacy, May 6-8, 1996, Oakland, CA, USA.

[46] Object Management Group, *UML Language 2.0 Superstructure Version 2.0 and Infrastructure Version 2.0 (formal /05-07-04)*, May 2004.

[47] G. Sindre & A. Opdahl, *Eliciting Security Requirements by Misuse Cases*, Proceedings of TOOLS Pacific 2000, 20-23 Nov 2000.

[48] I. Alexander, *Initial Industrial Experience of Misuse Cases in Trade-Off Analysis*, Proceedings of the IEEE Joint International Conference on Requirements Engineering (RE02), Essen, DE 2002.

[49] J. McDermott & C. Fox, *Using Abuse Case Models for Security Requirements*, 15th Annual IEEE Computer Security Applications Conference, 1999.

[50] P. Coad, E. LeFbebreve, J. Delucaca, *Java Modelling in Colour with UML: Enterprise Components and Processes*, Prentice Hall, 1999.

[51] A. Cockburn, *Writing Effective Use Cases*, Addison-Wesley, 2000.

[52] P. Krutchen, *The Rational Unified Process- An Introduction*, Addison-Wesley, 2003.

[53] D. Kulliack & E. Guiney,*Use Cases, Requirements in Context*, ACM PRESS, 1990.

[54] S.Robertson & J. Robertson,*Mastering the Requirements Process*, Addison Wesley ,1999.

[55] C. Larman, *Applying UML and Patterns, An Introduction to Object Oriented Analysis and Design*, Prentice-Hall, Upper Saddle River, NJ, 1998.

[56] L. Benoit, S. Ovemyer, & O. Rambod, *Proceedings of 23rd International Conference on Software Engineering (ICSE 2001)*, Toronto, Canada.

[57] M. Blaze, J. Feigenbaum, J. Ioannidis, & A. Keromytis, *The Role of Trust Management in Distributed Systems Security*, Chapter in Secure Internet Programming: Security Issues for Mobile and Distributed Objects, (Vitek and Jensen, eds.) Springer-Verlag, 1999.

[58] Joint Staff *Manual for Employing Joint Tactical Communications* CJCSM 6231.05A, November 1998.

[59] Internet Engineering Task Force, *RFC 1321, The MD5 Message-Digest Algorithm*, April 1992.

[60] Internet Engineering Task Force, *RFC 2401 Security Architecture for the Internet Protocol (IPSEC)*, November 1998.

[61] International Organization for Standards, *"ITU-T X.509 ISO/IEC9594-8: 2001 Information Technology-Open Systems Interconnection,The Directory: Public Key and Attribute Framework*

[62] B. Lampson, M. Abadi, M. Burrows, & E. Wobbe, *Authetication in Distrbuted Systems: Theory and Practice*, ACM Transactions on Computer Science Vol. 10(4), 1992.

[63] G. Lowe, *Breaking and Fixing the Needham-Schroeder Public-Key Protocol Using FDR* Lecture Notes in Computer Science, Proceedings of the Second International Workshop on Tools and Algorithms for Construction and Analysis of Systems, March 27-28, 1996 Passau, Germany.

[64] National Institute of Standards *NIST Pub 46-3, Data Encryption of Standard* October, 1999.

[65] National Institute of Standards *NIST Pub 197, Advanced Encryption Standard*, November 2001.

[66] National Institute of Standards *NIST Pub 182-2 Digital Signature Standard* August 2002.

[67] R. Savarda & M. Karash, *Explaining the Gap between Specification and Actual Performance of IPsec VPN Systems*, TSIC Insight, Volume 3, Issue 9, 4 May 2001.

[68] Department of Commerce, NIST, *Announcing Request for Candidate Algorithm Nominations for a New Cryptographic Hash Algorithm (SHA-3) Family*, Federal Register, Vol. 72, No 212,Friday, November 2, 2007.

[69] K. Pahlavan, P. Krishnamurthy, A. Hatami, M. Ylianttila, J. Makela, R. Pichna, & J. Vallstron, *Handoff in Hybrid Mobile Data Networks*, IEEE Personal Communications, April 2000.

[70] Telecommunications Industry Association, *TIA/EIA/TSB-102.AACA "Project 25 Digital Over the Air Rekeying Protocol*, April 2001.

[71] X. Wang, D. Feng, et al., *Collisions for Hash Functions MD4, MD5, Haval-128, and RIPEMD*, CRYPTO 04, revised August 17,2004.

[72] X. Wang, Y. Yin, & H. Yu, *"Finding Collisions in the Full SHA-1"*, Advances in Cryptography, Proceedings of the 25th Annual international Cryptology Conference, Crypto'05, 14-18 August 2005, Santa Barbara, CA.

[73] M. Hartong, R. Goel, D. Wijeskera, *Key Management Requirements for Positive Train Control Communications Security*, Proceedings of the 2006 IEEE/ASME Joint Rail Conference, 2006, Atlanta, Georgia,4-6 April 2006.

[74] Transportation Research Board of the National Academies, *US Railroad Efficiency: A Brief Economic Overview*, Proceedings of the Workshop on Research to Enhance Rail Network Performance, April 5-6, 2006 Transportation Research Board of the National Academies, Washington, DC, 2007.

[75] Surface Transportation Board, *Joint Petition for Service Order, STB Service Order No. 1518*, 31 October 1997.

[76] International Union of Railways, *GSM-R Procurement Guide*, February 2007.

[77] M. Lubbecke, U. Zimmermann, *Engine Routing and Scheduling at Industrial In-Plant Railroads*, Transportation Science, Volume 37, Issue 2, May 2003.

[78] A. Higgins, E. Kozan, *Modeling Train Delays in Urban Networks* Transportation Science, Volume 32, Issue 4, April 1998

[79] Q. Lu, M. Dessouky, & R. Leachman, *Modeling Train Movements through Complex Rail Networks*, ACM Transactions on Modelling and Computer Simulations (TOMACS) Volume 14, Issue 1, January 2004.

[80] D. Parkes & L. Ungar, *An Auction Based Method For Decentralized Train Scheduling* Proceedings of the Fifth International Conference on Autonomous Agents Montreal, Quebec, 2001.

[81] J. Lee, K. Sheng & J. Guo, *Fast and Reliable Algorithm For Railway Train Routing*, Proceedings of the IEEE Region 10 Conference on Computers, Communications, Control Engineering, Beijing, China 19-21 October 1993.

[82] A. D'Ariano, M Pranzo, & I. Hansen,*Conflict Resolution and Train Speed Coordination for Solving Time Table Pertubations*, IEEE Transactions on Intellegent Transportation Systems, June 2007.

[83] T. Ho, J. Norton ,& C. Goodman, *Optimal Traffic Control At Railway Junctions*, IEE Proceedings - Electric Power Applications – March 1997 – Volume 144, Issue 2.

[84] M. Lewellen & K. Tumay, *Network Simulation of a Major Railroad*, Proceedings of the 30th Winter Simulations Conference, , Washington, DC, 13-16 December 1998.

[85] S. Graff & P. Shenkin, *A Computer Simulation Of A Multiple Track Rail Network*, Sixth International Conference on Mathematical Modeling, 4-7 August 1987, St. Louis, MO, USA.

[86] T. Ho & T. Yeung, *Railway Junction Traffic Control By Heuristic Methods*, IEE Proceedings - Electric Power Applications – January 2001 – Volume 148, Issue 1.

[87] D. Barney, D. Haley, & G. Nikandros, *Calculating Train Braking Distances* Proceedings of the Sixth Australian Workshop on Safety Critical Systems and Software , Brisbane, Australia, June 2001.

[88] IEEE Std. 1474.1-2004 *IEEE Standard for Communications-Based Train Control, Appendix D.*

[89] M. Malvezzi, P. Presciani, B. Allotta, & P. Toni, *Probabilistic Analysis Of Braking Performance In Railways*, Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, Professional Engineering Publishing, Volume 217, Number 3, 2003.

[90] B. Vincze & G. Tarmai, *Development and Analysis of Train Brake Curve Caculation Methods with Complex Simulation*, Proceedings of International Exhibition of Electrical Equipment for Power Engineering, Electrical Engineering, Electronics, Energy and Resource-saving Technologies, Household Electric Appliances(ELECTR 2006) May 23, 24, 2006 Zilina. Slovika..

[91] BS 05/19984709 DC (UIC-546) EN 15179. *Railway Applications. Braking. Requirements for the Brake System of Passenger Coaches* British Standards Institution , 17 March 2005.

[92] *Draft Guide for the Calculation of Braking Distances for Rail Transit Vehilcles* IEEE P1698/D1.3, Februrary 2008.

[93] F. Yan & T. Tang, *Formal Modeling and Verification of Real-time Concurrent Systems*, Proceedings of the IEEE Internationl Conference on Vehicular Electronics and Safety, ICVES 2007. 13-15 December 2007, Beijing, China.

[94] E. Khmelnitsky *On an Optimal Control Problem of Train Operation*, IEEE Transactions on Automatic Control, Vol 45, Issue 7, July 2000.

[95] Z. Li-yan, L. Ping, J.Li-min,& Y Feng-yan, *Study on the Simulation for Train Operation Adjustment under Moving Block* Proceedings of the 2005 Intelegent Transportation Syztems, September 13-16, 2005 Vienna, Austria.

[96] H. Takeuchi, C. Goodman & S. Sone, *Moving Block Signalling Dynamics: Performance Measures and Re-starting queued electric trains* IEE Proceedings Electric Power Applications, 8 July 2003 Volume: 150, Issue: 4.

[97] H. Krueger, E. Vaillancourt, A. Drummie, S. Vucko, & J. Bekavac, *Simulation in the Railroad Environment* Procedings of the 2000 Winter Simulation Conference,10-13 December 2000, Orlando, FL, USA.

[98] B. Friman *An Algorithm for Braking Curve Calculations in ERTMS Train Protection Systems* COMPRAIL 2006 Tenth International Conference on Computer System Design and Operation in the Railway and Other Transit Systems 10-12 July 2006 Prague, Czech Republic.

[99] W. Rudderham, *Longitudinal Control System of the Intermediate Capacity Transit System* Proceedings of the 33rd IEEE Vehicular Technology Conference, 25-27 May 1983 Toronto, Ontario, Canada.

[100] M. Dessouky, Q. Lu, J. Zhao, & R. Leachman, *An Exact Solution Procedure to Determine the Optimal Dispatching Times for Complex Rail Networks* IEE Transactions, Vol 32 Issue 2, February 2006.

[101] M. Khan, D. Zhang, M. Jun, & J.Zhu *An Intelegent Search Technique to Train Sceduling Problem based on Gentic Algorithms* Proceedings of the 2006 International Conference on Emerging Technogies, Perhwar, Pakistan, 13-14 November 2006

[102] A. Tazoniero, R. Gonclaves, & F. Gomide *Decision Making Strategies for Real Time Train Dispatch and Control* Analysis and Design of Intellent Systems Using Soft Computing Techniques, Springer Berlin/Heildelberg Vol 41, 2007

[103] F. Li, Z. Gao, K/ Li, & L. Yang, *Efficient Scheduling of Railway Traffic Based on Global Information of Train*, Transportation Research, Part B: Methodological, Evseviner, 2008

[104] M. Penicka, *Formal Approach to Railway Applications*, Formal Methods and Hybrid Real Time Systems, Lecture Notes in Computer Science Volume 4700/2007 Springer, 2007

[105] M. Carey & I. Crawford, *Scheduling Trains on a Network of Busy Complex Stations*, Transportation Research, Part B: Methodological, Vol 41, Issue 2, February 2007, Evseviner.

[106] J. Tornquist, *Computer-based Decision Support for Railway Traffic Scheduling and Dispatching, A Review of Models* Proceedings of the 5th Workshop on Algorithmic Methods an Models for Optimazation of Railways, September 14, 2005 Palma de Mallorca, Spain.

[107] L. Anderegg, I Stephan, E. Gantenbein, I Stature, *Train Routing Algorithms: Concepts, Design Choices, and Practical Considerations* Proceedings of the 5th Workshop on Algorithm Engineering and Experiments, Baltimore, Maryland, January 2003.

[108] *Benefits and Costs of Positive Train Control, Report in Response to Request of House and Senate Appropriations Committees*, August 2004, Federal Railroad Administration.

[109] Public Law 110-432 *Rail Safety Improvement Act of 2008* U.S. Government Printing Office, October 2008.

Curriculum Vitae

Mark Hartong is a Senior Electronics Engineer in the Office of Railroad Safety, Federal Railroad Administration, U.S. Department of Transportation, Washington, DC. The Federal Railroad Administration is the regulatory and enforcement agency responsible for promoting safe and successful railroad transportation within the United States, and advancing the executive branch policies regarding freight and passenger rail. As an interdisciplinary electronics engineer he serves as the senior technical authority for safety and security critical electronics and software used in railroad systems.

Prior to joining the Federal Railroad Administration in 2003, he was a Staff Systems Engineer with Lockheed Martin Corporation where he provided systems engineering support for a wide range of classified communications hardware and software development programs. As a naval officer in the US Navy he qualified both as a Submarine Warfare Officer and a Naval Engineering Duty Officer. Submarine Warfare is the community within the Navy that involves the use of submarines towards the missions of sea control, projection of power ashore, and strategic deterrence. Engineering Duty is the Navy community responsible for research and development, design, acquisition, construction, modernization, and life cycle management of ships and ship's systems.

Mr. Hartong received a M.S. in Software Engineering and Certificate in Information Systems Security from George Mason University, Fairfax, Virginia and a M.S. in Computer Science from the U.S. Naval Postgraduate School, Monterey, California. While at the Naval Postgraduate School, in addition to his thesis research, he also served as a research associate on a separate classified Naval Research Laboratory communications project. Mr. Hartong also holds a B.S. with honors in Mechanical Engineering from Iowa State University, Ames, Iowa. He is a Registered Professional Engineer.