Chapter 18

A FRAMEWORK FOR INVESTIGATING RAILROAD ACCIDENTS

Mark Hartong, Rajni Goel and Duminda Wijeskera

Abstract Positive train control (PTC) or communication-based control systems (CBTC) control trains using wireless network infrastructures. Consequently, investigations of accidents involving PTC- or CBTC-controlled trains require network forensic analysis. This paper describes a forensic analysis framework that leverages the communications capabilities of PTC systems. The framework incorporates a centralized database architecture that securely stores PTC-related and other digital data, and provides for efficient and flexible querying of the data during accident analysis.

Keywords: Positive train control, accident investigations, centralized logging

1. Introduction

The North American freight and passenger railroads are currently introducing wireless-network-based control systems collectively known as positive train control (PTC) or communications-based train control (CBTC) systems to enhance railroad safety and security [7]. PTC systems control the authority of trains to occupy specific track segments, enforce speed limits and other restrictions, maintain safe inter-train distances and provide protection for railroad maintenance employees. PTC commands run at the application layer of a wireless communications network. Accordingly, they have the same advantages and disadvantages as other applications based on wireless protocol stacks. A major disadvantage is the susceptibility to mal-actions at all layers, potentially resulting in undesirable incidents or railroad accidents.

When PTC systems are employed, investigations of railroad accidents and recreations of potential accident scenarios require forensic analysis of wireless-based communications networks in addition to the usual

Please use the following format when citing this chapter:

Hartong, M., Goel, R., Wijeskera, D., 2007, in IFIP International Federation for Information Processing, Volume 242, Advances in Digital Forensics III; eds. P. Craiger and S Shenoi;(Boston: Springer), pp. 255-265.

examination of physical equipment, human factors and environmental conditions. Unfortunately, current railway networks do not have mechanisms for the comprehensive, secure and centralized collection of forensic data. This hinders the resolution of accident investigations as well as the prompt implementation of corrective actions. For example, the investigation of the 2005 Graniteville (South Carolina) train collision [12] by the Federal Railroad Administration (FRA) and National Transportation Safety Board (NTSB) took eleven months.

Digital control and state data exchanged over wireless communications networks required for operating PTC systems can be augmented with additional digital forensic data to support accident investigations. This paper describes a forensic analysis framework that leverages the communications capabilities of PTC systems. The framework incorporates a centralized database architecture that securely stores PTC-related and other digital data, and provides for efficient and flexible querying of the data during accident analysis.

The next section describes related work in railroad accident investigations and network forensics. Section 3 introduces PTC systems, and describes the proposed forensic architecture and data items used for incident/accident analysis of PTC-controlled trains. Sections 4 and 5 show how data in the forensic repository can be used for accident recreation and post-accident analysis, respectively. The final section, Section 6, presents our conclusions.

2. Related Work

Safe railroad operation is considered to be a national priority by the United States Government, which invests significant resources for this effort through the FRA and NTSB. These organizations have regulatory mandates to investigate accidents and major railroad incidents [16, 18]. In doing so, they ask questions similar to those asked by forensic examiners: What happened? How did it happen? Why did it happen? How could it be prevented from happening again?

Network forensics involves the acquisition, presentation and analysis of network traffic and other digital evidence for legal proceedings [13]. Current forensic practices involve passive or active monitoring [6, 11], and techniques for evidence presentation and automated reasoning [19].

In contrast, accident investigations do not determine guilt and liability; rather, their goal is to quickly and efficiently improve system safety. The FRA has recognized that the immediate access and evaluation of accident data assists in implementing operational improvements, ideally before the track is put back into service [17]. Thus, the automated gath-

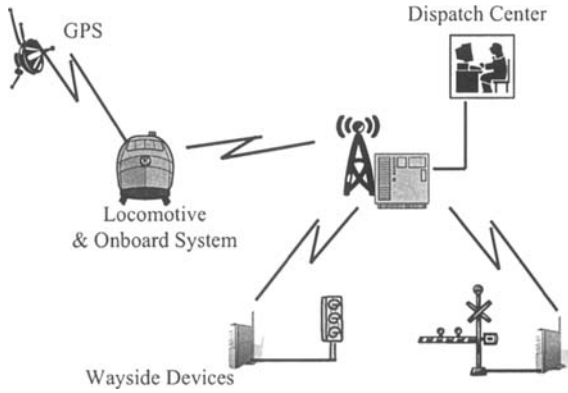

Figure 1. Generic PTC architecture.

ering of auditable evidence is very desirable for the purpose of railroad investigations.

3. Positive Train Control Systems

Locomotive crews in North America have traditionally communicated with office dispatchers and wayside devices using two-way radios, wayside signals or instructions written on paper that are handed over at various stations. The process did not require nor did it facilitate forensic data gathering capabilities. However, this began to change in the 1980's when Class I railroads in the United States and Canada developed the Advanced Railroad Electronic System (ARES) for integrating communications, command and control for railroad operations and business applications, and for enforcing positive train separation (PTS). The FRA subsequently expanded ARES to include the enforcement of speed restrictions and the protection of roadway workers within their authorities in addition to PTS. These three functions are now referred to as Level 1 Positive Train Control (PTC) [1, 2].

The generic PTC architecture presented in Figure 1 has three major functional subsystems:

- Wayside units, i.e., highway grade crossing signals, switches, interlocks, and maintenance of way workers.
- Mobile units, i.e., locomotives and other on-rail equipment with their onboard computers and location systems.
- Central office dispatch/control units.

In the PTC architecture, the dispatch office grants access requests for trains to occupy track segments. Trains enter and exit track segments when permitted by the track owner's dispatch office. Wayside devices monitor the track conditions and passing trains, and actively participate in communicating data between trains and dispatch offices.

For analysis purposes, we assume that PTC systems interoperate using trust management [8], where each railroad company maintains a certificate authority that issues and verifies the authenticity and validity of certificates presented by recognizable entities. Dispatch offices, trains and wayside devices on their own tracks are issued certificates and public/private key pairs to communicate with each other directly. To enable one railroad company's train that may be driven by a crew belonging to a second company to use a track segment belonging to a third company, the trust roots of all three railroad companies must cross certify each other. This helps ensure the authenticity and integrity of the collected data for accident investigations and reconstructions.

4. Network Forensics for Railway Accidents

The outcome of accident analysis is usually a description of one or more chains of interactions that produce multiple accident scenarios. The scenarios may occur due to human error, unexpected environmental conditions, unanticipated faults (e.g., equipment failure), and various communications-related issues (e.g., delayed or dropped packets carrying PTC information or deliberate attacks on network assets). Proper collection and analysis of accident data can be used to compute accident frequency and patterns. These can pinpoint locations needing special operational attention and safety improvements.

A preliminary logical design for collecting data from dispatch offices, trains and wayside devices, and maintaining it at a centralized repository is shown in Figure 2. Although NTSB and FRA investigators gather operational, environmental, human factors and maintenance data [3], due to space constraints, this paper considers only data related to operations and the environment.

A large amount of operational data is provided by mandatory locomotive event recorders [17]. Other data, such as track classifications are inferred from specific technical parameters that railroad companies are required to maintain to achieve specific levels of safe railroad operation. Track classifications are important because they regulate the maximum allowable speed of trains.

Environmental factors can significantly impact railroad operations. Precipitation and fog often reduce signal visibility, flash floods can wash

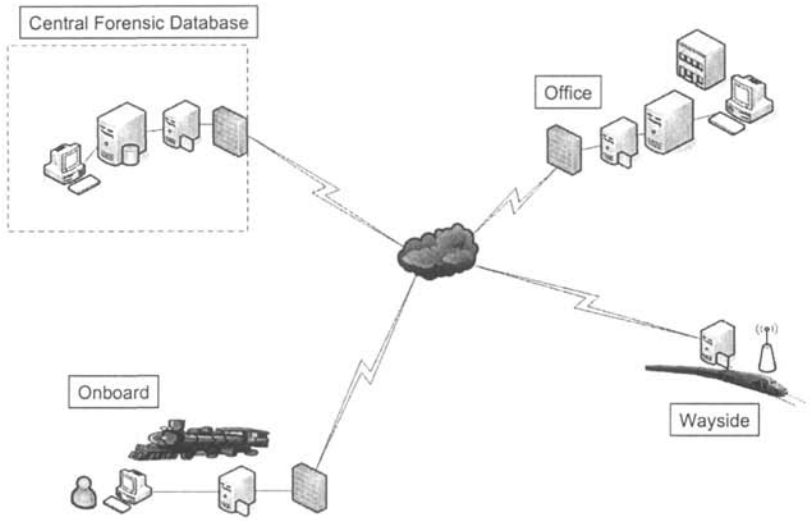

Figure 2. Centralized forensic data collection framework.

out tracks, excessive heat may warp tracks, crosswinds reduce stability and may even blow railcars off the tracks. Ice and snow may cause regional delays or shutdowns, and pre-existing accumulations of rain, ice and snow not associated with current weather conditions may also cause serious problems. Environmental and operational data collection and monitoring systems include EMD's Functionally Integrated Railroad Electronics (FIRE) and GE Transportation Systems' Expert On-Alert [10]. However, due to the absence of regulatory requirements, economic considerations mainly determine whether or not railroad companies deploy automated data collection and monitoring systems.

Tables 1–3 summarize the three types of operational and environmental data that may be obtained from central offices, onboard systems and wayside systems, respectively.

4.1 Forensic Database Architecture

We propose a centralized database to store, manage and query data items used for incident/accident analysis of PTC-controlled trains. Currently, this data is widely scattered, requiring significant efforts to collect and organize it before any analysis can be performed. For example, regulations mandate only a 48-hour retention period for locomotive event recorder data. Furthermore, while the recorders are tamper resistant, they are neither tamper proof nor crash hardened.

Table 1. Office data.

Data Class	Data Type	Attribute
Operational	Static	- Date - Time - Communication System Status - Track Characteristics: (i) Track Name & Number, (ii) Track Type, (iii) Track Class & Geometry, (iv) Track Database - Train Information: (i) Train ID, (ii) Train Type, (iii) Crew, (iv) Consist Data, (v) Location
	Command	- Date - Time - Office ID - Message Information: (i) Office Authorities & Special Instructions Issued, (ii) Onboard ID, (iii) Onboard Authorities & Special Instructions Acknowledged, (iv) Wayside ID, (v) Wayside Authorities & Special Instructions Acknowledged
Environment		- Date - Time - Location - Temperature - Dew Point - Wind Speed - Precipitation

The proposed centralized forensic database overcomes these limitations by logically storing data collected from locomotives and other devices. The centralized database must implement strict access controls to ensure that the integrity of the forensic data is maintained.

The forensic database comprises several relational tables shown in Figure 3. Tables are created for each PTC system entity that submits or receives operational and environmental data (mostly in the form of network packets). The schema and database design depend on the types of data collected by each entity, while the frequency of transmission and the communications bandwidth determine the data collection rates. Queries issued to the forensic database can reveal the accuracy and integrity of PTC commands sent or received by the various entities.

5. Post Accident Analysis

A promising method to identify the causal factors and the resulting accident scenarios with their evidence is to pre-analyze possible misuse

Table 2. Onboard data.

Data Class	Data Type	Attribute
Operational	Static	- Date - Time - Communication System Status - Train Information: (i) Train ID, (ii) Train Type, (iii) Consist Data, (iv) Crew Data, (v) Track Database, (vi) Location, (vii) Train Control System Status, (viii) Trailing Tons, (ix) Brake Pressure, (x) Throttle Position, (xi) Alerter Status, (xii) Horn & Bell Status, (xiii) Generator, (xiv) Light, (xv) Distance Traveled
	Command	- Date - Time - Onboard ID - Message Information: (i) Onboard Authorities & Special Instructions Issued, (ii) Office ID, (iii) Office Authorities & Special Instructions Acknowledged, (iv) Wayside ID, (v) Wayside Authorities & Special Instructions Acknowledged
Environment		- Date - Time - Horizontal Visibility Range

Table 3. Wayside data.

Data Class	Data Type	Attribute
Operational	Static	- Device Type - Date - Time - Communication System Status - Device Status & Health - Device Information: (i) Device ID, (ii) Location
	Command	- Date - Time - Message Information: (i) Device ID, (ii) Wayside Device Authorities & Special Instructions Issued, (iii) Office Authorities & Special Instructions Acknowledged, (iv) Onboard Authorities & Special Instructions Acknowledged
Environment		- Date - Time - Device Measured Variable

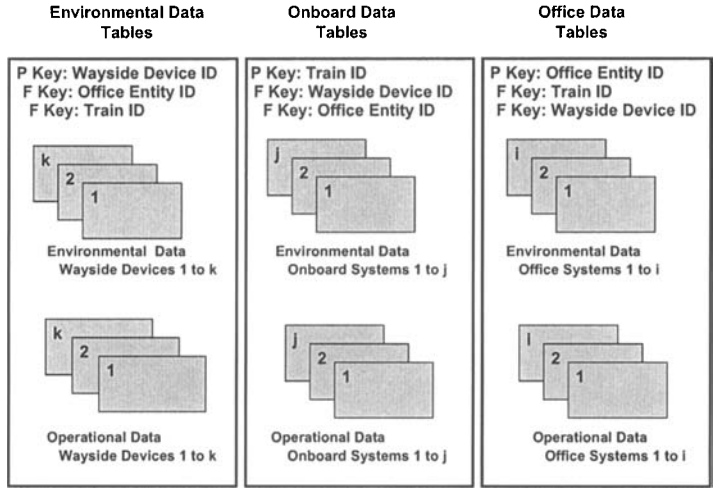

Figure 3. Forensic database tables and keys.

cases for PTC systems [9]. Use cases specify functional requirements provided by the system to its actors [14]. On the other hand, misuse cases [15] specify the foreseeable interactions between potential malactors and the system. Database queries can be crafted to search for evidence of misuse. For example, the following SQL query against the database defines an overspeed accident that results in a derailment (if the query evaluates to TRUE).

```
(SELECT Train Information.Throttle Position
   FROM Onboard Data.Operational.Static Train Information
   WHERE Train Information.Throttle Position = 8)
  AND
(SELECT Train Information.Location
   FROM Onboard Data.Operational:Static
   WHERE Train Information.Location = Curve 1)
  AND
(SELECT Track Characteristics.Track Type
   FROM Office Data.Operation.Static.Track Characteristics
   WHERE Track Characteristics.Track Type = Class 3)
```

6. Conclusions

The main objective of an accident investigation is to formulate recommendations that prevent future accidents. Ideally, the investigation is conducted by individuals who are experienced in accident causation and investigative techniques, and are very knowledgeable about the opera-

tional environment. However, collecting evidence is arduous and time-consuming; data may be minimal, missing or difficult to access. The evidence is also subject to omission, contamination or obliteration; therefore, it should be recorded immediately and preserved carefully, and its chain of custody should be maintained. The proposed methodology for application layer network forensics provides a basis for all these tasks.

There are, however, several implementation issues that must be addressed. In a operational environment where rail traffic is heavy and closely spaced, the volume of operational and environmental data that must be transmitted may exceed the communications bandwidth. Even if the communications infrastructure can handle the network traffic, the database transaction processing capability may be exceeded. The required capabilities can only be determined in the context of railroad operating environments and specific implementations.

Human factors issues are extremely important in accident investigations. FRA studies have revealed that certain kinds of human errors (e.g., improperly lining switches, failing to latch and lock switches, improperly conducting shoving movements) account for an inordinate number of accidents. FRA's 2003 study [4] reports that 133 (91%) of the 146 head-on, rear-end and side collisions were attributed to human causes. Likewise, 2004 accident data [5] reveals that 184 (91%) of the 202 collisions (56 more than in 2003) were due to human factors.

For the database solution to be viable, it is important that queries be created that accurately model accidents. Because safety flaws that are identified by accident investigations are quickly rectified, it is difficult to discern the complex interactions of the safety problems that remain. This factor along with the rarity of accidents makes the task of accurately modeling accidents a challenging endeavor.

Note that the views and opinions expressed in this paper are those of the authors. They do not reflect any official policy or position of the Federal Railroad Administration, U.S. Department of Transportation or the U.S. Government, and shall not be used for advertising or product endorsement purposes.

References

[1] Federal Railroad Administration, Railroad Communications and Train Control, Technical Report, Department of Transportation, Washington, DC, 1994.

[2] Federal Railroad Administration, Implementation of Positive Train Control Systems, Technical Report, Department of Transportation, Washington, DC, 1999.

[3] Federal Railroad Administration, *FRA Guide for Preparing Accident/Incident Reports*, Department of Transportation, Washington, DC, 2003.

[4] Federal Railroad Administration, *Railroad Safety Statistics - 2003 Annual Report*, Department of Transportation, Washington, DC, 2003.

[5] Federal Railroad Administration, *Railroad Safety Statistics - 2004 Annual Report*, Department of Transportation, Washington, DC, 2004.

[6] S. Garfinkel and E. Spafford, *Web Security, Privacy & Commerce*, O'Reilly, Sebastopol, California, 2002.

[7] M. Hartong, R. Goel and D. Wijesekera, Communications-based positive train control systems architecture in the USA, *Proceedings of the Sixty-Third IEEE Vehicular Technology Conference*, vol. 6, pp. 2987–2991, 2006.

[8] M. Hartong, R. Goel and D. Wijesekera, Key management requirements for positive train control communications security, *Proceedings of the IEEE/ASME Joint Rail Conference*, pp. 253–262, 2006.

[9] M. Hartong, R. Goel and D. Wijesekera, Use-misuse case driven analysis of positive train control, in *Advances in Digital Forensics II*, M. Olivier and S. Shenoi (Eds.), Springer, New York, pp. 141–155, 2006.

[10] T. Judge, How healthy are your locomotives? *Railway Age*, April 2001.

[11] S. Mukkamala and A. Sung, Identifying significant features for network forensic analysis using artificial intelligence techniques, *International Journal of Digital Evidence*, vol. 1(4), pp. 1–17, 2003.

[12] National Transportation Safety Board, Collision of Norfolk Southern Freight Train 192 with Standing Norfolk Southern Local Train P22 with Subsequent Hazardous Materials Release at Graniteville, South Carolina, January 6, 2005, Railroad Accident Report NTSB/RAR-05/04, Washington, DC, 2005.

[13] M. Ranum, K. Landfield, M. Stolarchuk, M. Sienkiewicz, A. Lambeth and E. Wal, Implementing a generalized tool for network monitoring, *Proceedings of the Eleventh USENIX Systems Administration Conference*, 1997.

[14] J. Rumbaugh, Getting started: Using use cases to capture requirements, *Journal of Object-Oriented Programming*, vol. 7(5), pp. 8–12, 1994.

[15] G. Sindre and A. Opdahl, Templates for misuse case description, *Proceedings of the Seventh International Workshop on Requirements Engineering: Foundations of Software Quality* (www.nik.no/2001 /21-sindre.pdf), 2001.

[16] U.S. Government, Investigations, Section 225.31, Federal Railroad Administration, *Title 49, Code of Federal Regulations*, Washington, DC, pp. 367–368, 2006.

[17] U.S. Government, Event Recorders, Section 229.135, Federal Railroad Administration, *Title 49, Code of Federal Regulations*, Washington, DC, pp. 409–413, 2006.

[18] U.S. Government, Functions, Section 800.3, National Transportation Safety Board, *Title 49, Code of Federal Regulations*, Washington, DC, p. 121, 2006.

[19] W. Wang and T. Daniels, Network forensic analysis with evidence graphs, *Proceedings of the Digital Forensics Research Workshop*, 2005.