Transportation Research Part C 95 (2018) 679–697

Contents lists available at ScienceDirect

Transportation Research Part C

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

Data-driven perspectives for energy efficient operations in railway systems: Current practices and future opportunities

Valerio De Martinis, Francesco Corman*

ETH Zurich, IVT Institute for Transport Planning and Systems, Stefano Franscini platz 5, 8093 Zurich, Switzerland

A R T I C L E I N F O

Keywords:
Data driven
Model based
Railway energy efficiency
Calibration

A B S T R A C T

Railway systems must increase their performance and economic competitiveness to remain an effective and efficient transport mode. Energy efficiency goals are one of the main drivers for the future evolution of planning and operations of transport systems. An opportunity to improve energy efficiency together with reliability and feasibility of railway systems come from the huge amount of data being currently collected and available in the future. The hidden potential in large sets of data for improving energy efficiency can be fully exploited through novel, data-driven approaches. This paper discusses the relation of those future approaches with the current state of the art and challenges, highlighting natural advantages and possible weak points. We identify dimensions within the current literature describing the suitability of current approaches to embrace the data revolution, and the possible enhancements resulting from that. We refer to practical test cases based on real on-board monitoring of electric trains in Switzerland to identify current and future challenges in improving energy efficiency of train operations. We conclude with a discussion and a roadmap on the introduction of data-driven approaches for improving energy efficiency of railway systems.

1. Introduction

This paper analyzes how the potential of data-driven approaches could help railway systems address the required increase in energy efficiency. Data-driven approaches have so far only touched railway systems, with small episodic works on many fields (monitoring, maintenance, more fluid traffic, etc.) (Thaduri et al., 2015). Data is seen as the new revolution (Baur and Wee, 2017; Athey, 2017; Gallotti et al, 2016), plus the current Smartrail 4.0 initiative in Switzerland (Zischek, 2017). The widespread use of analytics on large data volumes poses large methodological, computational, ICT (Information and communications technology) challenges, but also gives clear opportunities for data-driven improvements of many sectors.

It is expected that a relatively large amount of data would bring improvements in more precise models, and more effective energy reduction approaches. But how? That is the topic of this paper. Our focus is on the energy efficiency of the train operations, i.e. we analyze multiple processes and sources of data, which reduce the final energy consumption associated to motion of trains. Within this context, we analyze and classify the most recent state of the art in terms of potential and requirements related to the use of data-driven approaches.

We setup a practical realistic test case based on data measured in real life, to understand the impact of the typical characteristics of big data and analytics towards usability and value, in the problem at hand. Typically, data-driven challenges derive from very large data sets, requiring advanced data storage and processing, and identified in the key characteristics on data: Volume, Velocity,

* Corresponding author.
E-mail addresses: valerio.demartinis@ivt.baug.ethz.ch (V. De Martinis), francesco.corman@ivt.baug.ethz.ch (F. Corman).

https://doi.org/10.1016/j.trc.2018.08.008
Received 12 October 2017; Received in revised form 21 August 2018; Accepted 21 August 2018
0968-090X/

Volatility, Variety, Veracity and Validity. **Volume** refers to the presence of large amounts of data, well beyond the RAM memory of most computers. This requires new solutions for data transmission, storage retrieval and elaboration. **Velocity** refers to the speed of data streaming into hosting/storage platforms; online control might require simultaneous transmission to systems for elaboration and automatic actions, regardless of the dynamics currently followed in analyzing the system. A high **Volatility** relates to information, which is considered valid, usable or available only for very short time. Most online control processes require a lot of data, but past states are not important per se; there is no need to store data, as the volatility of the system is too high to learn from past states. A high volatility requires transmission and elaboration before it is too old and unusable. **Variety** refers to the number of sources and types of data that can be collected and processed; for train operations this is typically structured data (quantities, small amount of categories). The term **Veracity** implies that a vast amount of data, collected by multiple sensors across a variety of operating conditions, unavoidably comes with biases, noises, errors, outliers, which might affect greatly the following data processing. A similar concept is associated to **Validity**, which pertains a general area of data correctness and accurateness for the intended use.

The rest of the paper goes as follows. Section 2 describes the necessary background on railway systems and data-driven approaches, and typical applications for increasing energy efficiency in railway system operations. Section 3 reports a systematic categorization of the literature, Section 4 quantifies some practical challenges found in this field. Section 5 elaborates further on the trends seen in technology, academia and a roadmap towards data-driven approaches for energy efficient railway systems. Based on the review of the state of the art, and the challenges identified, we finally discuss future promising fields of application, challenges and threats. Section 6 concludes the paper with the most important remarks.

2. Background

2.1. Railway systems and their energy efficiency

Large investments in railway systems, from urban to national networks, confirm that railways are still one of the preferred solutions for mobility of people and goods. To remain competitive in a market with increased competition due to disruptive technologies (autonomous vehicles) and uncertainties (global economy), continuous improvement are required to be effective and attractive in terms of economics and performance (Meng et al., 2018). Energy efficiency is a key aspect of railway systems attractiveness and it plays a very important role for the future. Compared to smaller road-bound vehicles, the energy consumption per transported passenger (or unit of freight) of the railway system is much less than any road-bound transport (EEA, 2013). In addition, there is a global trend related to reduction of CO2 emission and energy waste. From a systemic perspective, railway systems need to contribute on these aspects while performing its services with unchanged or improved quality; i.e. from using alternative cleaner energy sources (Turner, 2015) to optimizing processes and procedures for decreasing consumptions. Energy consumption in railway systems mostly comes from tractive efforts applied to keep vehicles moving, while auxiliary systems have lower impacts on the overall energy needs (González-Gil et al., 2014). From now on, we restrict our focus to the operations (i.e. tractive efforts required to make trains move) of both passenger and freight trains. Moreover, we focus on electric trains, although some proposed content can be extended to diesel powertrains and hybrid systems (with onboard energy storage).

We report in Fig. 1 a schematic and generic description of subsystems and their elements that play a role in the determination of energy for the motion of electric trains. Every box represents the generic concept of an aspect for a subsystem; the shaded blue box in the middle refers to aspects where we can identify in detail relations, when we analyze the related literature. Arrows are generic representations of the interdependencies that we find more relevant to the scope of this paper; there are many more out of the scope. We describe Fig. 1 from left to right. For a complete description of the subsystems and their interaction, a comprehensive discussion can be found in Hansen and Pachl (2014) and Iwnicki (2006).

The mobility needs of people are typically represented as **customer demand**. Trains are operated to fulfill this demand. If

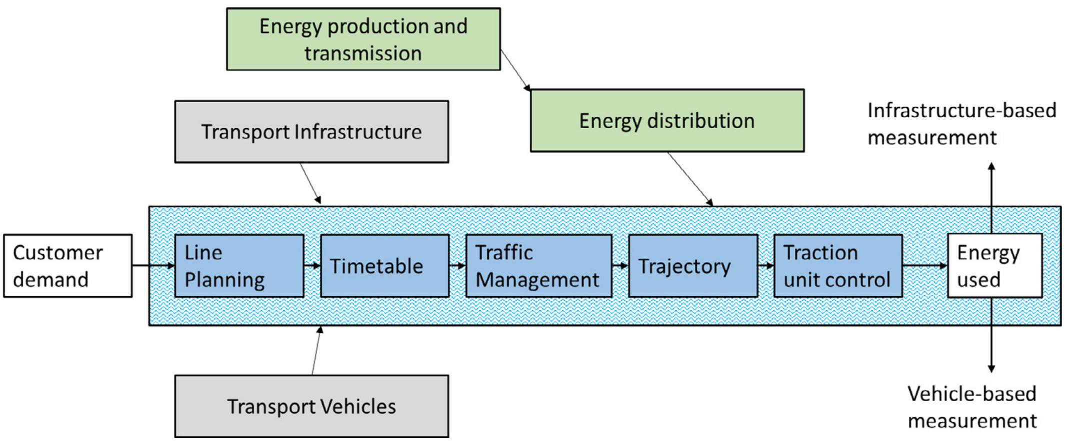

Fig. 1. Schematic representation of interdependencies between aspects subsystems involved in energy consumption.