Applied Ergonomics 53 (2016) 298–311

Contents lists available at ScienceDirect

Applied Ergonomics

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

Big data and ergonomics methods: A new paradigm for tackling strategic transport safety risks

Guy Walker *, Ailsa Strathie

Institute for Infrastructure and Environment, Heriot-Watt University, Edinburgh, EH14 8AS, UK

ARTICLE INFO

Article history:
Received 21 May 2015
Received in revised form 22 July 2015
Accepted 21 September 2015
Available online 6 November 2015

Keywords:
Big data
Leading indicators
Human performance
Methods

ABSTRACT

Big data collected from On-Train Data Recorders (OTDR) has the potential to address the most important strategic risks currently faced by rail operators and authorities worldwide. These risk issues are increasingly orientated around human performance and have proven resistant to existing approaches. This paper presents a number of proof of concept demonstrations to show that long standing ergonomics methods can be driven from big data, and succeed in providing insight into human performance in a novel way. Over 300 ergonomics methods were reviewed and a smaller sub-set selected for proof-of-concept development using real on-train recorder data. From this are derived nine candidate Human Factors Leading Indicators which map on to all of the psychological precursors of the identified risks. This approach has the potential to make use of a significantly underused source of data, and enable rail industry stakeholders to intervene sooner to address human performance issues that, via the methods presented in this paper, are clearly manifest in on-train data recordings. The intersection of psychological knowledge, ergonomics methods and big data creates an important new framework for driving new insights.
© 2015 Elsevier Ltd and The Ergonomics Society. All rights reserved.

1. Introduction

Data recording is the act of automatically logging information on system parameters over time. Data recording has become increasingly ubiquitous in rail transport operations and easily qualifies as a big data problem (e.g. Wu and Liu, 2014; Geisler et al., 2012). Big data is a rather broad term first used by NASA scientists to describe particularly thorny computer graphics problems which taxed the ability of computer hardware like memory and storage: "We call this the problem of big data. When data sets do not fit in main memory (in core), or when they do not fit even on local disk, the most common solution is to acquire more resources." (Cox, 1997, p. 235). In the absence of universal agreement on a definition (e.g. Press, 2014) it is possible to turn to Google™, where it is described more broadly as: "extremely large data sets that may be analysed computationally to reveal patterns, trends, and associations, especially relating to human behaviour and interactions." For this reason, big data looms large over the Ergonomics discipline and represents a potentially paradigm shifting trend going forward (see Drury, 2015). An example of the issues at stake can be found in the rail industry. Entire national train fleets are now required to carry On Train Data Recorders (OTDRs), or so called 'Black Boxes', for the purposes of post-accident investigation. Although the data is only accessed periodically (often rarely) the devices themselves are continuously recording how individual trains are being driven, at increasing rates, and across an increasing range of parameters. The outflow of data is therefore extensive and growing in terms of volume (the amount of data), velocity (the speed at which it is growing) and variety (the number of things being measured; Laney, 2001; Drury, 2015). Perhaps because of this, and the practical challenges involved in storing and mining such large quantities of data, it currently represents a significantly underused resource (Hart, 2003). The question, then, is what could it be used for? In this paper we aim to demonstrate that it has the potential to be coupled to existing ergonomics methods, and used to tackle the most important strategic risk issues currently faced by rail operators and authorities worldwide.

1.1. The black box paradox

The intersection of big data and the rail transport context embodies three paradoxes:

1.1.1. Paradox 1: the opportunities to use OTDRs for their original post-accident purposes are diminishing, because there are so few accidents

Improving safety trends are a very welcome outcome of improved safety practices, but because so few major rail accidents occur there are now relatively few opportunities to use OTDRs for their original purpose. With so few accidents there is also no longer enough data to construct reliable forward looking estimates in a way that might have been possible for certain historic engineering (e.g.

* Corresponding author.
E-mail address: G.H.Walker@hw.ac.uk (G. Walker).

http://dx.doi.org/10.1016/j.apergo.2015.09.008
0003-6870/© 2015 Elsevier Ltd and The Ergonomics Society. All rights reserved.

metal fatigue) or human performance issues (e.g. fatigue risk; Evans, 2011).

1.1.2. Paradox 2: the accidents 'left-over' from improving safety trends are proving resistant to existing methods

Safety data is levelling off, with a persistent class of human/system accident now elevated to the status of a key risk (RSSB, 2009; Stanton and Walker, 2011). When safety performance data reaches the level of that achieved in the rail sector it instead starts to become characterised by periodicities, cycles or discrete events. This is becoming evident in EU rail safety data, with a large scale rail accident occurring on average every six years (EU, 2003). This class of accident includes Signals Passed At Danger (SPADs) or overspeed events such as the Santiago de Compostela rail accident in Spain. Here it is possible to demonstrate numerous engineering and procedural safeguards, full compliance with necessary regulatory requirements, well trained individuals and a modern train fleet, yet a largely unforeseen accident still occurred.

1.1.3. Paradox 3: improving safety trends mean that OTDR data needs to be used less frequently, yet we are collecting more of it than ever before

The third and final paradox is that the opportunities to use On-Train Data Recorders (OTDRs) for their original purpose (i.e. post-accident analysis) are diminishing at the same time as the technical capabilities of data recorders are increasing (Geisler et al., 2012; Morcom, 1970). What this means is that Exabyte's of non-accident data are being collected day in and day out, but not currently used. This is a significant irony given the "widespread concern within the industry that the background indicators – rather than the headline grabbing ones – have remained worryingly stable" (Wolmar, 2012).

The opportunity embedded in these paradoxes is to use big data from transport recordings to detect accident precursors, specifically those accidents which have proven resistant to current approaches and which are responsible for the current plateauing of accident trends.

1.2. Human Factors Leading Indicators

Leading indicators are measurable precursors to major events such as an accident. The indication of a precursor leads, or comes before, the actual event itself. Lagging indicators are the opposite. These are so called loss metrics that can only become apparent after an event (Rogers et al., 2009). Leading indicators are said to be proactive because they enable steps to be taken to avoid seriously adverse consequences. Lagging indicators are said to be reactive in that a seriously adverse event needs to occur before it can be learnt from. For this reason, leading indicators are also sometimes referred to as positive performance indicators and lagging indicators as negative performance indicators. The concept of leading and lagging indicators originally derives from the field of economics and the need to understand when one phase of a cyclical economic process this will change to another (Mitchell and Burns, 1938). The terms have been appropriated more recently by the safety and risk field, particularly in view of developments in Safety Management Systems (SMS) since the 1990s. The notion of leading indicators is also beginning to emerge from the ergonomics literature in the form of research on behavioural markers (e.g. Nixon et al., 2015), the use of novel human performance proxy measures (e.g. McIntire et al., 2014), and a more long standing interest in the modelling and prediction of human performance (Hamilton and Clarke, 2005). Leading indicators, in a safety management and ergonomics context, can be defined as "something that provides information that helps the [the organisation] respond to changing circumstances and take actions to achieve desired outcomes or avoid unwanted outcomes" (Step Change, 2003, p. 3). That 'something' can be defined according to the risk factors underlying the troublesome class of operational accident (or near accident) that is the focus of this paper.

The reason for the emphasis on human performance can be seen in the 'broad causes' attributed to recent rail accidents. Out of seven broad causes attributable to European rail accident data (Evans, 2011), four out of seven, including the top three, involve a prominent human performance dimension. Expressed in descriptive terms, these broad causes involve a task in which a combination of the following happens: 'getting out of sequence', 'losing situational awareness', 'allocating attention incorrectly and/or 'allowing prior experience to override the correct action', or combinations of all four. Table 1 summarises a comprehensive review of the literature (reported elsewhere, Walker and Strathie, 2012) linking these descriptive terms to a set of specific risk factors. If we are able to detect when these risk factors are present, using big data as the input, then we should also be able to make progress on these key strategic risks.

The basic research problem can be stated thus: despite considerable improvements in safety performance in the rail sector, a persistent class of accident/near accident continues to occur. These incidents reside at the interface of people and systems. What is required is a means to detect the presence or emergence of such problems before they manifest themselves as a serious operational incident. This paper describes how big data from OTDRs can be used to 'drive' established ergonomics methods to provide leading indicators of specific risk factors. Four proof-of-concept demonstrators are presented to show how these approaches work, selected on the basis of their potential scalability when used with big data.

2. Methodology

2.1. Data file and parameters

On-Train Data Recording (OTDR) devices are the equivalent of aircraft 'black boxes' and continuously record a range of parameters referring to the individual train unit and how it is being driven. The study uses a real-world OTDR sourced from a UK train operating company. The UK rail industry has an excellent safety record, but like rail operators the world over, the accidents which do occur tend to foreground the role of human factors. Likewise, the range of parameters measured is similar to those measured in all other regions. The OTDR data file itself is a continuous download from a single traction unit. The recording started at 05:34:57 on the 6th July 2012 and ceased at 21:36.32 on the 11th July 2012. This is a period of 136 h, 1 min and 35 s during which the train made 107 journeys and travelled 1638 miles. The raw data takes the form of a Comma Separated (CSV) file containing a data matrix 191,021 time samples (rows) deep by 72 parameters (columns) wide: a total of 13.8 million data points. The logger itself scans the parameters for changes at a rate of 20 mS but, in the present system, to economise on memory requirements data are only logged when one of the 72 parameters changes (up to the maximum scanning/sampling rate). In the present case the mean sampling rate was 2.56 s. The OTDR device itself was a UK Railway Group Standards compliant Arrowvale unit which recorded 72 parameters, 25 of which are in addition to those mandated. In terms of data classification four of the parameters; time, distance and two speed signals derived from a driven and non-driven axle, are continuous ratio data. The remaining 68 are nominal/binomial (i.e. on or off). Explanations for the channels relevant to the present analyses are contained in the proof of concept descriptions.

Download English Version:

https://daneshyari.com/en/article/548311

Download Persian Version:

https://daneshyari.com/article/548311

Daneshyari.com