Safety Science xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Safety Science

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

Towards increased reliability by objectification of Hazard Analysis and Risk Assessment (HARA) of automated automotive systems

Siddartha Khastgir [a,*], Stewart Birrell [a], Gunwant Dhadyalla [a], Håkan Sivencrona [b], Paul Jennings [a]

[a] WMG, University of Warwick, UK
[b] Qamcom Research And Technology AB, Gothenburg, Sweden

ARTICLE INFO

Article history:
Received 21 November 2016
Received in revised form 15 March 2017
Accepted 27 March 2017
Available online xxxx

Keywords:
Hazard
HARA
ISO 26262
Functional safety
Reliability

ABSTRACT

Hazard Analysis and Risk Assessment (HARA) in various domains like automotive, aviation, and process industry suffers from the issues of validity and reliability. While there has been an increasing appreciation of this subject, there have been limited approaches to overcome these issues. In the automotive domain, HARA is influenced by the ISO 26262 international standard which details functional safety of road vehicles. While ISO 26262 was a major step towards analysing hazards and risks, like other domains, it is also plagued by the issues of reliability. In this paper, the authors discuss the automotive HARA process. While exposing the reliability challenges of the HARA process detailed by the standard, the authors present an approach to overcome the reliability issues. The approach is obtained by creating a rule-set for automotive HARA to determine the Automotive Safety Integrity Level (ASIL) by parametrizing the individual components of an automotive HARA, i.e., severity, exposure and controllability. The initial rule-set was put to test by conducting a workshop involving international functional safety experts as participants in an experiment where rules were provided for severity and controllability ratings. Based on the qualitative results of the experiments, the rule-set was re-calibrated. The proposed HARA approach by the creation of a rule-set demonstrated reduction in variation. However, the caveat lies in the fact that the rule-set needs to be exhaustive or sufficiently explained in order to avoid any degree of subjective interpretation which is a source of variation and unreliability.

1. Introduction

Over 90% of the on-road accidents occur due to human error (Singh, 2015). Therefore, an ability to assist or replace the human driver in the driving task has a potential to reduce the number of accidents. The introduction of Advanced Driver Assistance Systems (ADAS) and Automated Driving (AD) systems has been driven by the fact that these systems will be able to improve road traffic safety. This is due to the higher ability of an automated system to react to a possible hazardous situation as compared to the most alert manual driver (Carbaugh et al., 1998). Apart from safety benefits, AD systems and ADAS also offer the potential for increased operational efficiency by increasing road through-put by reducing the proximity between vehicles (Bishop, 2000; Kesting et al., 2008; van Arem et al., 2005).

In 1996, Sweden adopted a "Vision Zero" policy which states that *"eventually no one will be killed or seriously injured within the road transport system"* (Johansson, 2009). It brought together multiple stakeholders like vehicle manufacturers, road designers, state, city councils, municipalities and individuals, in order to achieve the mission of zero on-road fatalities. According to Vision Zero's viewpoint, a holistic approach needs to be adopted. While changes in vehicles is a major aspect of the solution (with the introduction of passive safety, active safety and automated features), other aspects include changes in roads, streets, knowledge/awareness of individuals and legislations (Tingvall, 1998). While the principles of Vision Zero concept is valid for every country, the identification of changes and their implementation differs from country to country and the cultural aspect of the country needs to be taken into consideration in the strategic analysis plan (Johansson, 2009).

While ADAS and AD systems are an important part of achieving a Vision Zero concept, both ADAS and AD systems offer new challenges for testing and the safety analysis of the systems (Khastgir et al., 2015). Variety of ADAS and AD systems exist or are in development, each of them offers a different kind of a challenge. As we move towards higher levels of automation in the SAE's six levels of automation (level 0–5) (SAE International, 2016), testing and risk analysis becomes harder as it needs to include larger number of variables and their interactions in the analysis. The authors discuss

* Corresponding author.
E-mail address: S.Khastgir@warwick.ac.uk (S. Khastgir).

http://dx.doi.org/10.1016/j.ssci.2017.03.024
0925-7535/

risk analysis within the scope of this paper. Section 1.1 discusses risk analysis in a general setting, Section 1.2 briefly discusses reliability through objectification of the risk analysis process and Section 1.3 discusses automotive risk analysis.

1.1. Reliability and validity of risk analysis

Safety analysis is a two-step process. In the first step one needs to identify the hazards for which the Hazard Analysis and Risk Assessment (HARA) is to be performed. There are various methods for identifying hazards like System Theoretic Process Analysis (STPA)/Systems Theoretic Accident Model & Processes (STAMP) (Leveson, 2004, 2011a, 2011b), JANUS (Hoffman et al., 2004), Accimaps (Salmon et al., 2012), HFACS (Baysari et al., 2009; Chen et al., 2013; Wiegmann and Shappell, 2001b), Fault-tree analysis (Lee et al., 1985; Reay and Andrews, 2002), bow-tie analysis (Abimbola et al., 2016; Khakzad et al., 2012), FMEA (Stamatis, 2003), etc. Some of these methods were developed for simpler systems and fall short in their ability to meet the requirements for the analysis of modern systems which have multiple interactions between the system and software components and the human operator (Fleming et al., 2013). Another source of identifying hazards is from experience of previous accidents and their accident investigations. However, being retrospective in nature, they cannot be taken as the only source of possible hazards, but should influence future hazard identification process and safety management process (Stoop and Dekker, 2012). While accident investigations provide new knowledge about the possible avenues of system failures, they are never exhaustive. This is evident by the deja-vu experience of similar accidents repeating themselves in a 20–30 year cycle (Le Coze, 2013). Identifying hazards has its challenges and is a research question in its own right. While it is possible to identify hazards based on the "known knowns" and accommodate for the "known unknowns", it is extremely difficulty to foresee the unknown knowns and even more so for the "unknown unknowns" which form the *"Black Swan"* category for hazards (Aven, 2013). Previous accidents, however, provide an insight to the occurrence of "*Black Swan*" type of accidents by increasing experts' knowledge of possible factors for risk analysis (Khakzad et al., 2014). While the authors appreciate that hazard identification is an important area for research with on-going activities, it remains out of scope of this paper. Identification of hazards will be discussed by the authors in future publications.

The second step of the safety analysis process involves the analysis of the hazard and the corresponding risk assessment for the hazard. Risk in general has been suggested to be a construct and not an attribute of the system (Goerlandt and Montewka, 2015), due to the subjective nature of risk (Aven, 2010a; Tchiehe and Gauthier, 2017). However, in the automotive domain, a decomposition of risk provides a different insight. An Automotive Safety Integrity Level (ASIL) rating in automotive HARA comprises of a severity, exposure and a controllability rating. Controllability and Severity of any system are system attributes. However, exposure for a system remains a construct and is open to subjective variation as it is influenced by the expert's knowledge which governs the probability rating (Aven, 2010b; Aven and Reniers, 2013). Automotive HARA and ASIL will be discussed in detail in Section 2–6. This paper deals with the classification of hazards (once they have been identified) and their subsequent risk assessment.

While HARA governs the risk management, i.e., the mitigation steps and the rigour required in the application of the steps; it is plagued by some fundamental challenges of its validity and reliability (Aven and Zio, 2014). One of the fundamental issues with risk assessment is the biases or assumptions made by stakeholders performing the assessment due to subjective interpretation of the underlying process or lack of knowledge of the underlying uncertainties or lack of knowledge of the system safety. Lack of knowledge or improper knowledge about the system may lead to either ignoring possible risk (which may lead to false negatives) or their exaggeration (which may lead to false positives). This introduces uncertainty in the risk analysis which is not taken into consideration while making decisions (Goerlandt and Reniers, 2016). Additionally, the knowledge of the hazards and possible failures helps guide the design process of the systems by providing the ability to make informed design decisions in the design phase of the product (Björnsson, 2017; Villa et al., 2016).

Reliability refers to the "*extent to which a framework, experiment, test, or measuring instrument yields the same results over repeated trials*" (Carmines and Zeller, 1979). In a review of Quantitative Risk Analysis (QRA) method applications, Goerlandt et al. (2016) found that significant differences existed in the results of QRA conducted by different teams/groups of experts. While mandating a specific QRA method could reduce variation (Van Xanten et al., 2013), they argued that this would not ascertain the accuracy of the results, but make results converge and more comparable.

For HARA to be scientific, it needs to be reliable (Hansson and Aven, 2014). In this paper, the authors adopt the "reliability" definition and types of reliability as defined by Aven and Heide (2009) *(pg. 1863)*:

- *"The degree to which the risk analysis methods produce the same results at reruns of these methods (R1).*
- *The degree to which the risk analysis produces identical results when conducted by different analysis teams, but using the same methods and data (R2)*
- *The degree to which the risk analysis produces identical results when conducted by different analysis team with the same analysis scope and objectives, but no restrictions on methods and data (R3)*"

1.2. Reliability through objectivity

According to Cambridge English Dictionary ("Cambridge English Dictionary," 2017), "objectivity" is defined as "*the state or quality of being objective and fair*", where "objective" is defined as "*based on real facts and not influenced by personal beliefs or feelings*". In order to prevent the influence of personal beliefs and mental models of experts leading to varied and unreliable HARA ratings, the authors propose the introduction of a rule-set to introduce objectivity in the process. Objectivity could potentially be a tool to help provide consistency and convergence of HARA ratings, thus providing increased reliability.

1.3. Automotive functional safety

In the automotive domain, the ISO 26262-2011 standard (automotive functional safety international standard) lacks a quantified and a robust process for automotive certification (Yu et al., 2016). The standard refers to ASIL as a metric for hazard analysis which is influenced by Severity (S), Exposure (E) and Controllability (C) rating. However, the methodology for determining these parameters and their quantification is not mentioned. Instead a set of sample tables has been provided (Ellims and Monkhouse, 2012). SAE J2980 provides some guidance to certain degree of objectivity to automotive HARA. But it too falls short in defining various aspects influencing severity, exposure and controllability rating (SAE International, 2015). SAE J2980 provides one table to parametrise severity using speed and collision type as parameters. It doesn't provide any guidance for controllability and exposure ratings. Even for severity, the parameters used are not exhaustive enough.

Thus, there is a need for creating a method for extracting patterns and creating templates for safety case development which would influence the HARA (Kelly, 2004). While ISO 26262 (2011)

Please cite this article in press as: Khastgir, S., et al. Towards increased reliability by objectification of Hazard Analysis and Risk Assessment (HARA) of automated automotive systems. Safety Sci. (2017), http://dx.doi.org/10.1016/j.ssci.2017.03.024