Driving Behavior Model Considering Driver's Over-Trust in Driving Automation System

Hailong Liu
Graduate School of Informatics,
Nagoya University
Furo-cho, Chikusa-ku,Nagoya,464-8601,Japan
liuh@murase.m.is.nagoya-u.ac.jp

Toshihiro Hiraoka
Institute of Industrial Science,
The University of Tokyo
4-6-1, Komaba, Meguro-ku,
Tokyo 153-8505, Japan
thiraoka@iis.u-tokyo.ac.jp

AutomotiveUI '19 Adjunct , September 21–25, 2019, Utrecht, Netherlands

ACM ISBN 978-1-4503-6920-6/19/09.
https://doi.org/10.1145/3349263.3351525

Abstract
Levels one to three of driving automation systems (DAS) become more and more sophisticated, not only the driver's driving skills will be reduced, but also the problem of over-trust will become serious. To prevent the over-trust, this paper discusses the followings: 1) the definition of over-trust in the DAS, 2) the occurrence conditions and the mechanism of over-trust in the DAS, and 3) a driving behavior model considering driver's over-trust in driving automation system.

Author Keywords
Driving behavior model; over-trust; risk homeostasis theory; mental model; driving automation systems

CCS Concepts
•Security and privacy → Trust frameworks; •Human-centered computing → User models; Collaborative interaction; HCI theory, concepts and models;

Introduction
Driving automation system (DAS) can be regarded as a dynamically control system of a vehicle, and it has been developed by imitating the driving process of a human driver considering interactions among the driver-vehicle-environment. A driver generally recognizes the vehicle state (e.g.: velocity, acceleration) and the driving environment (e.g.: the location, the shape of road, objects) and hazards,

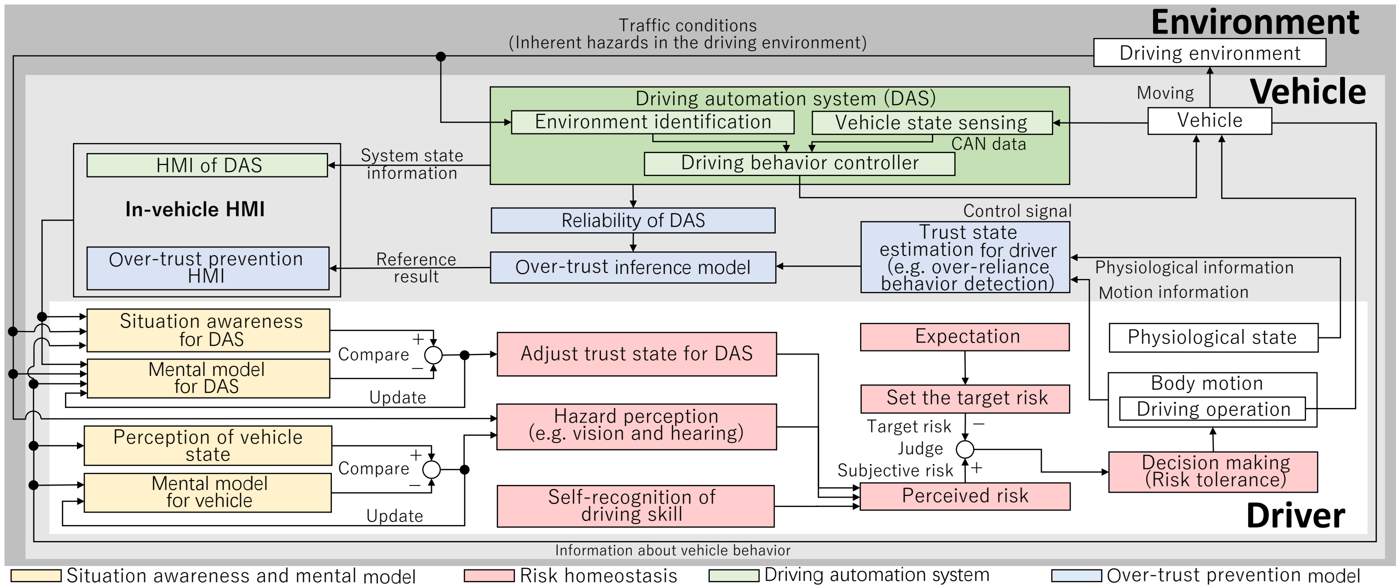

Figure 1: Driving behavior model considering driver's trust for vehicle system and risk compensation behavior when the driver is driving vehicle with driving automation system which equips an over-trust prevention HMI.

via body sensory information (e.g.: visual information, auditory information). After that, the driver takes the recognized vehicle state and the driving environment into account to decide a driving plan, and therefore he/she drives the vehicle. As for the DAS, the driving process is the same as the driving process of the human driver. The DAS equips various sensors such as a camera, G-sensors, gyros, a LiDAR, and a microphone. These devices are used to identify the vehicle state and the traffic conditions instead of the driver's three semicircular canals, eyes, and ears. The DAS analyzes the information observed by sensors and controls the vehicle. The above process is common to all DAS basically.

Different types of DASs have different requirements for usage conditions such as driving task and operational domain. *SAE international* defined the DASs into five levels [13]. Levels one to three of the DASs require the driver to be fully or partially involved in the driving tasks (e.g.: part of the vehicle control and emergency intervention). It means that the driver must not only "properly monitor" the driving environment and the state of the DAS all the time while driving, but also need to "properly handles" driving task when the situation exceeds the functional limitation of the DAS or when the DAS fails. Due to the above constraints, it is easy to imagine that the following problems will occur more frequently as the system performance improves: **1)** The driver's driving skill [17] and situation awareness [14] will decline as the driving opportunities decrease. **2)** The prolonged monitoring control for the DAS is extremely likely to increase the mental and physical workload of the driver when the DAS does not (or less) requires driving operation of the driver [11]. **3)** As the DAS functionality becomes more diverse and more reliable, the driver may easily over-trust in the DAS especially when the driver does not clearly understand the usage conditions and the functional boundaries of the DAS [12, 15, 3].

What is over-trust in the DAS?

Inagaki defined over-trust as an incorrect situation of a diagnostic decision where the user believes that the object is trustworthy although it actually is not trustworthy [7]. The over-trust in the DAS is that the DAS cannot respond to driving tasks but the driver trusts it can. Consequently, there are two judgment conditions of over-trust: **1) the driver is trusting in the DAS**, **2) the DAS can not respond to driving tasks**. Based on the two conditions of over-trust and the four dimensions of trust [9], there are four types of over-trust: over-trust in foundation, purpose, process, and performance. One reason for over-trust is that the user has misconceptions about the four dimensions of the system. The future goal of this study is to propose the method to prevent over-trust by reducing those misconceptions.

Two problems in the judgment of over-trust

Problem 1: Non-real-time judgment If the situation where the DAS cannot respond happens when the driver trusts in the DAS, he/she can observe the fact only after the situation happens. The DAS can detect the system error only after it happens, even if it has an error detection function. Therefore, from both viewpoints of the driver and the system, it is difficult to judge the over-trust in real-time.

Problem 2: Un-observable property of over-trust The judgment conditions of over-trust include two factors: the driver and the DAS. The driver cannot fully know the performance of the DAS and the situations where the DAS cannot respond. The DAS cannot measure the driver's trust state directly because the trust state is an internal psychological activity. Therefore, neither the driver nor the monitoring system cannot observe the over-trust directly.

There are two countermeasures to solve the above-mentioned problems; 1) a driver monitoring system to estimate the driver's trust state, and 2) a system to predict whether the DAS can respond to the driving situation or not. Combined with these two systems, over-trust will be predicted before a dangerous event happens.

Mechanisms of over-trust and driving behavior modeling

In order to prevent over-trust, the mechanisms of over-trust should be discussed first. This paper proposes a model, as shown in Fig. 1, which consists of a driving behavior generating process, a process of the DAS, and a process of over-trust prevention HMI. This model is used to interpret the mechanisms of over-trust.

While driving a vehicle, the driver always tries to understand the vehicle states such as velocity and acceleration and the traffic environment. Moreover, the driver perceives the hazards by comparing the recognized vehicle state with the mental model [6] for the vehicle. The driver also perceives the hazards through the driving environment by considering traffic conditions. The mental model for the vehicle is an explanation of the thought process of the driver about how to operate the vehicle and move it in the real world. It is based on the long term driving experience of the driver by estimating the relationship between the operation (system input) and the change of the vehicle state (the output of the system). "*Stepping on the gas pedal leads to vehicle accelerating*" is the most simple example of the mental model obtained from driving experience. The mental model can be updated by comparing the predicted vehicle state via the mental model and the recognized vehicle state. Similarity, a mental model of the DAS also can be generated through the experience of the driver who drives the vehicle with the DAS many times. The driver recognizes the situation of the DAS by the feedback via the HMI, the information of vehicle behavior, and the hazards in the driving environment. Then the driver compares the predicted result by using the mental model with the actual situation of the DAS in order to adjust his/her trust state.

The driver will perceive the risk from the hazards in driving environment and his/her trust in the DAS. The self-recognized driving skill also affects the risk perceived by the driver. Note that the perceived risk is a subjective one. According to the risk homeostasis theory [16], the driver decides his/her driving behavior (decision making) by comparing perceived risk with the acceptable risk level. The acceptable risk level is defined by the expected utility which is affected by the long-term experience. If the perceived risk is lower than the acceptable risk level, then the driver's driving behavior will become riskier. The driver's behavior will become more careful in the opposite case. This way of adjusting behavior based on the perceived risk is called risk compensation behavior.

After that, the driver controls the body to complete the driving maneuver. On the side of the DAS, it generates driving signals based on environmental information and vehicle status. The vehicle status can be measured by in-vehicle sensors via CAN bus, and environmental information can be measured by cameras and LiDARs. Finally, the vehicle is eventually controlled by combining the the driving maneuver and the control signals from the DAS.

After the vehicle moving, the driver will recognize the updated situation of the vehicle, the DAS and the driving environment, through the movement of the vehicle. These processes will be looped during driving.

A machine learning model may estimate the driver's trust state for the DAS based on the driver's posture, the driving maneuver, and his/her physiological information. Moreover, a evaluation system may estimate the ability of the DAS on real-time. For example, a defect detecting system could be seen as an evaluation system. It may be used to find the abnormal control signal generated by the DAS and evaluate the ability of the DAS [10]. Meanwhile, the reliability of the DAS should be evaluated too [2]. If the predicted ability of the DAS becomes too low to respond to the situation, and the machine learning model estimates that the driver still trusts in the DAS, then an over-trust prevention HMI will be activated to present alerts and some useful information about the ability of the DAS. Moreover, the DAS will control the vehicle to perform a minimal risk maneuver such as slowing down to stop, if a driver does not respond to the alert in a period of time.

Conclusion

This paper discussed the definition and properties of over-trust in the DAS, and proposed the mechanisms of over-trust and the driving behavior model. As a future work, authors will try to make a concrete over-trust estimation model, e.g., estimating the driver's trust states in DAS [1, 4, 8]. Moreover, authors will design an over-trust prevention HMI which provides effective information, e.g., 1) perception of the elements in the environment, 2) comprehension of the current situation, and 3) projection of future state [5].

REFERENCES

1. Genya Abe, Kenji Sato, and Makoto Itoh. 2018. Driver trust in automated driving systems: The case of overtaking and passing. *IEEE Transactions on Human-Machine Systems* 48, 1 (2018), 85–94.

2. Naoki Akai, Luis Yoichi Morales, and Hiroshi Murase. 2018. Simultaneous pose and reliability estimation using convolutional neural network and Rao–Blackwellized particle filter. *Advanced Robotics* (2018), 1–15.

3. Oliver Carsten and Marieke H. Martens. 2018. How can humans understand their automated cars? HMI principles, problems and solutions. *Cognition, Technology & Work* (12 May 2018).

4. Munjal Desai, Poornima Kaniarasu, Mikhail Medvedev, Aaron Steinfeld, and Holly Yanco. 2013. Impact of robot failures and feedback on real-time trust. In *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*. IEEE Press, 251–258.

5. Mica R Endsley. 1995. Toward a theory of situation awareness in dynamic systems. *Human factors* 37, 1 (1995), 32–64.

6. Stefan N Grösser. 2012. Mental model of dynamic systems. In *Encyclopedia of the sciences of learning*. Springer, 2195–2200.

7. Toshiyuki Inagaki and Makoto Itoh. 2013. Human's overtrust in and overreliance on Advanced Driver Assistance Systems: a theoretical framework. *International journal of vehicular technology* 2013 (2013).

8. Makoto Itoh. 2012. Toward overtrust-free advanced driver assistance systems. *Cognition, technology & work* 14, 1 (2012), 51–60.

9. John Lee and Neville Moray. 1992. Trust, control strategies and allocation of function in human-machine systems. *Ergonomics* 35, 10 (1992), 1243–1270.

10. Hailong Liu, Tadahiro Taniguchi, Kazuhito Takenaka, and Takashi Bando. 2018. Defect-Repairable Latent Feature Extraction of Driving Behavior via a Deep Sparse Autoencoder. *Sensors* 18, 2 (2018), 608.

11. Akira Ohtani, Yoshinori Egami, Kenji Sato, Kazushi Mitsui, and Masaaki Abe. 2016. Human Machine Interface with Driver in Automated Driving. *Transactions of Society of Automotive Engineers of Japan* 47, 4 (2016), 961–966.

12. Raja Parasuraman and Victor Riley. 1997. Humans and automation: Use, misuse, disuse, abuse. *Human factors* 39, 2 (1997), 230–253.

13. SAE. 2016. J3016:Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. (2016), 1–30.

14. Neville A Stanton and Mark S Young. 2005. Driver behaviour with adaptive cruise control. *Ergonomics* 48, 10 (2005), 1294–1313.

15. Hiroshi Uno, Akira Ohotani, Tsutomu Asoh, and Yukinobu Nakamura. 2010. Comparison of Task Demand Assessment Techniques for In-vehicle Driver Interface to Present Iformation. *Transactions of Society of Automotive Engineers of Japan* 41, 2 (2010), 539–544.

16. Gerald JS Wilde. 1982. The theory of risk homeostasis: implications for safety and health. *Risk analysis* 2, 4 (1982), 209–225.

17. Mark S Young and Neville A Stanton. 2007. What's skill got to do with it? Vehicle automation and driver mental workload. *Ergonomics* 50, 8 (2007), 1324–1339.