Driver's State Monitoring: A Case Study on Big Data Analytics

Shaibal Barua(✉), Shahina Begum, and Mobyen Uddin Ahmed

School of Innovation, Design and Engineering,
Mälardalen University, 72123 Västerås, Sweden
{shaibal.barua, shahina.begum, mobyen.ahmed}@mdh.se

Abstract. Driver's distraction, inattention, sleepiness, stress, etc. are identified as causal factors of vehicle crashes and accidents. Today, we know that physiological signals are convenient and reliable measures of driver's impairments. Heterogeneous sensors are generating vast amount of signals, which need to be handled and analyzed in a big data scenario. Here, we propose a big data analytics approach for driver state monitoring using heterogeneous data that are coming from multiple sources, i.e., physiological signals along with vehicular data and contextual information. These data are processed and analyzed to aware impaired vehicle drivers.

1 Introduction

Automotive industries are devoting to develop autonomous vehicle, however, before achieving that final goal [1], we have to rely on human drivers. Hence, driver state monitoring in terms of distraction, cognitive load, sleepiness, stress, etc. is essential in the transportation research area. These states are identified as causal factors of critical situations that can lead to road accidents and vehicle crashes. These driver impairments need to be detected and predicted in order to reduce critical situations and road accidents.

In the past years, physiological signals along with vehicular data and contextual information have become conventional measures in driver impairment research. Physiological sensors signals, i.e., Electrooculogram (EOG), Electroencephalogram (EEG), Electromyography (EMG), Electrogastrogram (EGG), etc. become the part of big data biological process that are both structural and non-structural, and complex to analyze. Furthermore, vehicular data, e.g., steering wheel movement, lateral position, break, etc. and contextual data such as driving experience, time of driving, road condition, etc. are also vital measures of driver's impairments. Besides, in real-time scenario, it requires a trade-off among resources, for handling the stream of sensor data [2]. Moreover, the advancement of technologies such as Internet, cloud computing, sensors, and wireless networks consequential of generating huge amounts of data stream. Big data analytics is the process that extract useful geometric and statistical pattern, retrieve knowledge, and use for decision making by analyzing and understanding the features of the massive dataset [3, 4]. The challenges in the big data analytics are to manage, process and transform the extracted structured data [5]. Therefore, we propose a multilayer approach for physiological big data analytics that can be used for vehicle

M.U. Ahmed et al. (Eds.): HealthyIoT 2016, LNICST 187, pp. 145–147, 2016.
DOI: 10.1007/978-3-319-51234-1_24

driver monitoring; where usage of machine learning and reasoning with artificial intelligence can facilitate the challenging tasks of big data analytics [6].

2 Approaches and Methods

A schematic diagram of big data analytics for drivers' state monitoring, is shown in Fig. 1. The first layer is the outlier detection and data cleaning layer. Sensor signals can often be contaminated with noises and need to be cleaned before analyzing. For example, EEG signals have gained increasing interest in mobile environments [7] such as vehicle driving, however often contaminated by ocular and muscle artifacts. We have developed a fully automated EEG artifacts handling algorithm called ARTE (Automated aRTifacts handling in mobile EEG) [8]. ECG can be cleaned using existing methods [9]. The next layer processes and creates structural data from the cleaned signals. Vehicular and physiological signals are time synchronized and resampled; categorical and quantitative information are retrieved and formatted from contextual information. Later, features are extracted from all kinds of data and signals, and stored for the data analytics.

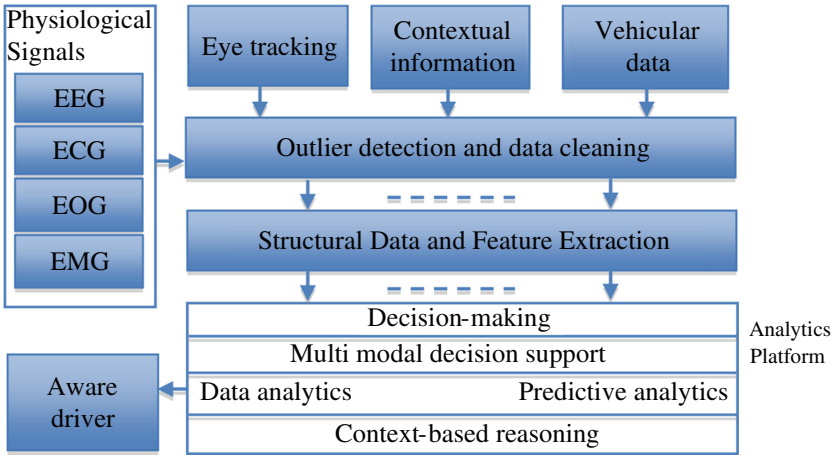

Fig. 1. Schematic diagram of proposed system

The core module of this approach is the analytics platform. The analytics platform is the combination of several sub-modules, i.e., decision-making, multimodal decision support, data analytics, predictive analytics, and context-based reasoning. Decision-making and data analytics can be achieved using statistical analysis, machine-learning algorithms. Data mining and machine learning tools and methods can retrieve hidden patterns in the data and also can be used for knowledge discovery [10, 11]. Deep learning based approach will provide classification and prediction from the data. Case-based decision-making is one possible use for the decision-making where previous decision come useful and will be integrated with the multimodal decision support module.

3 Discussion

Impaired driving due to driver's mental states, i.e., sleepiness, inattention, cognitive load, stress, etc. are one of the main researches of Safety driving in the transportation authorities and automotive industries. Most of the studies of driver's state monitoring are based on a single parameter, e.g., heart rate variability, eye tracking, or EEG signal analysis. Moreover, now a day because of the sensors availability, cloud computing, and IoT, datasets become huge in volume and also consist of a large variety in characteristics, and data are gathered with high velocity. In our study, we have considered multimodal approach, where several physiological signals, along with contextual information and vehicular data can be used to classify driver's state in real time. Using adaptable artificial intelligence and machine learning algorithms, knowledge representation, reasoning, and information retrieval can be handier and precise. Here, several sub-modules are combined within the analytics platform, using various machine learning algorithms that can provide constructive awareness to the impaired drivers.

References

1. Greenblatt, J.B., Shaheen, S.: Automated vehicles, on-demand mobility, and environmental impacts. Curr. Sustain./Renew. Energy Rep. **2**, 74–81 (2015)
2. Baccarelli, E., Cordeschi, N., Mei, A., Panella, M., Shojafar, M., Stefa, J.: Energy-efficient dynamic traffic offloading and reconfiguration of networked data centers for big data stream mobile computing: review, challenges, and a case study. IEEE Netw. **30**, 54–61 (2016)
3. Suthaharan, S.: Big data classification: problems and challenges in network intrusion prediction with machine learning. SIGMETRICS Perform. Eval. Rev. **41**, 70–73 (2014)
4. Leary, D.E.O.: Artificial Intelligence and Big Data. IEEE Intell. Syst. **28**, 96–99 (2013)
5. Cuzzocrea, A., Song, I.-Y., Davis, K.C.: Analytics over large-scale multidimensional data: the big data revolution! In: Proceedings of the ACM 14th International Workshop on Data Warehousing and OLAP, pp. 101–104. ACM, Glasgow (2011)
6. Maltby, D.: Big data analytics. In: 74th Annual Meeting of the Association for Information Science and Technology (ASIST), pp. 1–6 (Year)
7. Lin, C.T., Ko, L.W., Chiou, J.C., Duann, J.R., Huang, R.S., Liang, S.F., Chiu, T.W., Jung, T.P.: Noninvasive neural prostheses using mobile and wireless EEG. Proc. IEEE **96**, 1167–1183 (2008)
8. Barua, S., Begum, S., Ahmed, M.U.: Intelligent automated EEG artifacts handling using wavelet transform, independent component analysis and hierarchal clustering. In: Workshop on Embedded Sensor Systems for Health through Internet of Things (ESS-H IoT) at 2nd EAI International Conference on IoT Technologies for HealthCare (2015)
9. Kaufmann, T., Sütterlin, S., Schulz, S.M., Vögele, C.: ARTiiFACT: a tool for heart rate artifact processing and heart rate variability analysis. Behav. Res. Methods **43**, 1161–1170 (2011)
10. Condie, T., Mineiro, P., Polyzotis, N., Weimer, M.: Machine learning for big data. In: Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data, pp. 939–942. ACM, New York (2013)
11. Depeige, A., Doyencourt, D.: Actionable Knowledge As A Service (AKAAS): leveraging big data analytics in cloud computing environments. J. Big Data **2**, 1–16 (2015)