Computers in Industry 99 (2018) 226–240

Contents lists available at ScienceDirect

Computers in Industry

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

A novel Big Data analytics and intelligent technique to predict driver's intent

Lech Birek[a], Adam Grzywaczewski[a], Rahat Iqbal[b], Faiyaz Doctor[c], Victor Chang[d,*]

[a] Jaguar Land Rover, UK
[b] Coventry University, UK
[c] University of Essex, UK
[d] Xi'an Jiaotong-Liverpool University, China

ARTICLE INFO

Keywords:
Driver's intent prediction
Big Data
Big Data analytics
Computational intelligence
E-calendar
Geo referencing

ABSTRACT

Modern age offers a great potential for automatically predicting the driver's intent through the increasing miniaturization of computing technologies, rapid advancements in communication technologies and continuous connectivity of heterogeneous smart objects. Inside the cabin and engine of modern cars, dedicated computer systems need to possess the ability to exploit the wealth of information generated by heterogeneous data sources with different contextual and conceptual representations. Processing and utilizing this diverse and voluminous data, involves many challenges concerning the design of the computational technique used to perform this task. In this paper, we investigate the various data sources available in the car and the surrounding environment, which can be utilized as inputs in order to predict driver's intent and behavior. As part of investigating these potential data sources, we conducted experiments on e-calendars for a large number of employees, and have reviewed a number of available geo referencing systems. Through the results of a statistical analysis and by computing location recognition accuracy results, we explored in detail the potential utilization of calendar location data to detect the driver's intentions. In order to exploit the numerous diverse data inputs available in modern vehicles, we investigate the suitability of different Computational Intelligence (CI) techniques, and propose a novel fuzzy computational modelling methodology. Finally, we outline the impact of applying advanced CI and Big Data analytics techniques in modern vehicles on the driver and society in general, and discuss ethical and legal issues arising from the deployment of intelligent self-learning cars.

1. Introduction

The automotive industry is an ever-evolving competitive sector, which was propelled forward through human ingenuity and scientific advances. A comparison between today's vehicles with the ones that drove off the assembly line just 100 years ago would demonstrate that: today's cars last longer, are more fuel efficient, are able to provide their owners with a comfortable driving experience, while offering a variety of services through the integration of networking and ITC solutions. Vehicle manufacturing is one of the many industrial success stories of the 20th century and set to be one of the most rapidly evolving industries in the 21st century. The freedom offered by a car is now of paramount importance to modern citizens and the modern vehicle is emerging to be an extension of peoples' everyday living and working environments. This growing reliance and interdependency means that there is high demand for a wide range of quality services to be seamlessly offered by the vehicles. As it was stated by Jonathan Ive the head of design of Apple Computers "The car has become an article of dress without which we feel uncertain, unclad, and incomplete in the urban compound". The exponential growth of miniaturization in computing technologies and the rapid advancements in communication technologies have led a variety of sensing equipment and interactive information systems to rapidly find their way into modern vehicles. This wealth of information offered through the invisible web of smart electronics, along with the opportunities offered by state of the art features installed in the cabins and powertrain systems of modern vehicles, deliver the promise of providing an enhanced driving experience with the help of specifically designed embedded computing applications. Modern research and every day experience, demonstrates that being in a good mood is one of the best precondition for safe driving, and that happy drivers produce fewer accidents [1]. On the contrary, when the car, or the environment causes the driver to feel negative emotions such as aggressiveness, anger and stress, this influences their concentration and skills, as for example preventing the driver from concentrating on the road ahead, which can consequently cause accidents, and life threatening situations [2].

In recent years, developing novel applications, which can promote driver's experience by effectively anticipating and reacting to their

* Corresponding author.
E-mail address: vchang1_76@yahoo.co.uk (V. Chang).

https://doi.org/10.1016/j.compind.2018.03.025
Received 6 June 2017; Received in revised form 22 February 2018; Accepted 20 March 2018
0166-3615/

intent, have gained an immense boost from the Big Data and Internet of things revolution. Accumulating and utilizing data from the various data sources available in the modern hi-tech environment allows for gaining detailed insights on the drivers needs, preferences and habits, thus offering opportunities for the development of improved driver assistive systems. The importance of Big Data to this and wider application areas is highlighted in a recent statement by IBM's chief executive officer that “Big Data is the new oil” [3]. The analogy of data to the liquid that has fueled our cars and our lives in the previous century has been identified and highlighted by various studies demonstrating the power and impact of Big Data to modern economy and societies. As pointed out by Hashem et al. Big Data has three main characteristics. Firstly, the data itself is numerous and high dimensional. Secondly, it is not possible to categorize the data into regular relational databases, and finally data streams are created, captured, and analyzed rapidly [4]. As Gerhard mentions Big Data is a revolutionary leap forward from traditional analysis, which possesses three main characteristics volume, variety, and velocity [5]. Volume refers to the amount of data, which are created and stored. Variety is related to the various types of and sets of data collected, and velocity can be defined as the speed of data generation, streaming and aggregation [6]. In Kaissler et al.'s study, data value, and complexity are also proposed as Big Data characteristics. Data value is a measure of the usefulness of data in decision-making process, while complexity is a measure of the degree of interdependence and interconnectedness in Big Data structures. The inherent difficulty concerning the handling of these large amounts of data results in major challenges concerning its storage and analysis, as well as cost and time associated to the efficient delivery of results. Moreover, the results of this analysis should be delivered in an interpretable and easily visualized way [7]. Big Data analytics refers to the techniques utilized in order to examine and process the data so that hidden underlying patterns are revealed and interesting relations and other insights concerning the application context under investigation are exposed.

In modern times advances in software and hardware technologies provide a wealth of diverse data sources, which generate a wealth of Big Data information relating to the drivers intent and preferences, allowing for patterns of behavior of the driver to be revealed. It is common knowledge that most people use the same routes when commuting to work or other everyday locations (i.e. home, petrol station, supermarket etc.) at the same times of the day for the same days of the week. Similarly, although a person may shop on different days or at different times, they will often visit the same grocery store(s) [8]. Therefore, it is possible to predict a person's future preferences and actions if they are correlated with past behavior [9]. User profiling, and data provided by GPS, or geographical data accumulated by mobile devices can be utilized in order to reveal the driver's preference concerning specific roads and destinations thus allowing for the necessary adjustments to the vehicle's configurations. Data streams containing physiological signal information such as (EEG, HR, GSR) captured with the use of sensory equipment installed in the car, or from unobtrusive wearable sensors can provide insights of the driver's physical and mental state, which can aid the car in anticipating the driver's intent and actions. Smart phones and modern mobile devices are another platform that poses embedded sensors (such as GPS, WiFi, accelerometers etc.) and can collaborate with on-board computer systems to offer valuable information in order to reveal the user's needs and behavioral patterns [10]. Video input through cameras inside the car monitoring the driver's eyes and upper body posture can provide video streams which can reveal different aspects of the driver such as their sleepiness or fatigue levels [11], while cameras outside the car can provide information concerning the position of the car on the lane thus contributing in revealing the driver's intents. For example, this information may be used to identify the driver's intent to stop the vehicle and take a break to rest, or their intentions to overtake a slow moving vehicle ahead. Audio equipment can be utilized for capturing audio signals relating to the driver's behavioral patterns or psychosomatic state. Previous research have detected yawning to monitor the driver's drowsiness levels [12] and other audio signals can be utilized so that the car can anticipate the driver's state of mind in context of their vehicle usage at a given point in the day, i.e. an intention for a relaxed drive home after a tiring day at work. A rich source of information to reveal driver's intent can be provided by the ever-expanding use of social networks. Before entering their cars, the drivers interact with social networks, where they post comments and status updates which if harvested can be utilized by the car to automatically detect their driver's desire to drive to a certain destination or to reveal their overall affective state. Personal data from electronic calendars and mailing lists is another information source, which can be used in order to provide insights of the driver's driving patterns. This can be combined with data concerning the volume of traffic, which can be utilized to inform the driver and allow the car to plan an optimal route to the desired location. This data can be provided from RSS feeds or other related publicly available news websites. The behavior and intent of the driver can also be related to the way they use their car. Controller Area Network (CAN bus) data captured by modern cars, relating to the actions of the driver, such as the amount of pressure applied in the acceleration or brake pedals, the angle rotation of the steering wheel, and others parameters, can also be utilized to discover behavior patterns and construct effective predictive models of the driver's response to different situations and their expectations from the vehicle.

Most of the data sources identified above are related to the ability of modern car to connect with other smart networks and devices and has led the automotive industry to a very important point of its history. Connectivity is not new to the car and it has been introduced as a premium feature to vehicles for years. It is only now that the world wide regulation for emergency call and stolen vehicle tracking is mandating connectivity across all new vehicles introducing the need for automotive companies to compete even more in this market. The smart phone revolution has also created a big market for infotainment systems, which rely on this connectivity infrastructure, to provide the same experience within vehicles. Vehicular Ad Hoc Networks (VANET) for example is a powerful technology which has emerged from the need to support the utilization of the various wireless products and devices in the car and it has been used by applications in intelligent transportation systems (ITS), surveillance systems, and safety alerts on the road [13,14]. The introduction of vehicle connectivity to the modern automotive industry creates vehicle specific opportunities. Current luxurious cars have significantly complex computer systems with an array of internal computer networks, embedded devices as well as hundreds of non-trivial sensors gathering gigabytes of data every minute. State of the art vehicles are being equipped with significant amounts of networking infrastructure to provide connectivity to other vehicles, road infrastructure, home networks, power grid data networks and end user mobile devices (Bluetooth, NFC, Wi-Fi and its derivatives). The self-learning car developed by Jaguar Land Rover is an example of a state of the art vehicle exploiting some of these capabilities [15]. This emerging technological environment creates a massive opportunity for vehicles to take full advantage of the concept of the Internet of Things by collecting, processing and aggregating this information and transforming it into knowledge required by their cars, other drivers, vehicles, and the wider society. For example having a distributed sensor network of cars being able to assess the density of traffic on the road can be used to optimize the traffic within a region, maximizing the road and public transport throughput.

For the advantages of Big Data and Internet of Things to be exploited there is an increasing need for implementing intelligent systems and computational techniques which can reduce the complexity and cognitive burden on accessing and processing the large volumes of data generated in both embedded hardware and software based data analytics [16,17]. Big challenges stem from the utilization of Big Data in real world problems such as the prediction of the driver's intent, since