Human Trust Study on Autonomous Car

(Technical Paper)

Disinformation in Targeted Health Advertisements and Posts in Social Media

(STS Paper)

A Thesis Prospectus Submitted to the

Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree

Bachelor of Science, School of Engineering

Zi Zeng

Fall, 2019

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Signature _____ Date _____

Zi Zeng

Approved _____ Date _____

Lu Feng, Department of Computer Science & Engineering Systems and Environment

Approved _____ Date _____

Sean Ferguson, Department of Engineering and Society

Part I: Technical Report Prospectus

Introduction

Autonomous cars have many advantages, such as giving the driver more free time and significantly decreasing traffic congestion and the risk of transportation-related accidents, provided, of course, that the autonomous system functions correctly. Most autonomous cars still offer the option for drivers to switch to manual driving mode whenever they would prefer. However, it is expected that a driver will only take over when an incident is detected and an alarm is raised and then, only if the driver does not trust the ability of the autonomous car to avoid the incident in autonomous mode. To get the best use out of an autonomous system, it needs to be designed in such a way that it will be trusted by its drivers. Our research group plans to conduct a study on human trust in an autonomous driving simulator, using human subjects. I hope to gain a better understanding of the topic by participating in the study and by learning to use the PreScan software to design driving scenarios for use in a driving simulator.

Technical Topic

Researchers define human trust in an autonomous system slightly differently. To avoid ambiguity when conducting the human trust experiments, our research group defined trust as the participants' delegation of their responsibility when driving to the autonomous system. This indicates that they are willing to accept the uncertainties and risks that may be encountered during autonomous driving activities. To study human trust in an autonomous driving system, a group of participants will be invited to operate a driving simulator. Each participant will drive in four different scenarios. In each scenario, several different types of incident will be introduced, such as stationery vehicles and pedestrians.

The autonomous car relies on various forms of advanced sensor and on the algorithms used to calculate the meaning of the sensor response. The outputs from the sensors assist in adjusting the behavior of the throttle, brake pedals, and steering wheel of the driving simulator. All incidents are designed to be sensor detectable, and the driver is given two options when they hear the alarm from the autonomous system. If they don't trust the autonomous system, they are free to change to manual mode, and they are thereafter able to take control of the steering, brake, and throttle. Alternatively, they can choose not to switch and to let the autonomous system handle the incident. Throughout the whole experiment, the driver will indicate their level of trust in the autonomous system by pushing buttons on the simulator. The changes in their trust rate will be recorded. Such human trust ratings are frequently used in evaluating human trust in an autonomous driving system. Strauch et al. (2019) report research where ratings are recorded before, during, and after both manual and autonomous driving modes (Strauch, 2019). Other studies show that participants tend to believe that the reactions of autonomous cars when facing incidents are unpredictable. This impression lowers the trust rating that drivers give to the autonomous system and this could represent a huge challenge for manufacturers. If they want to commercialize autonomous cars, it is essential for the public to trust and accept them (Sadvandi & Halkias, 2019).

All scenarios to be used in the simulation will be created on a driving simulation platform called PreScan. PreScan has the ability to execute designed scenarios using its GUI-based pre-processor (TASS International, 2019). With its powerful functionality, PreScan is widely accepted in both the industry and academic research as a valid tool for the design and study of

autonomous driving. It allows researchers to test with ease, driving scenarios within a virtual system and also the efficiency of their driving algorithms. Our research group has been using PreScan to design a series of autonomous driving scenarios (Feng, 2019). The ongoing project will involve designing scenarios to be used by another research group that has the ability to collect and analyze, inter alia, the time taken for a driver to switch from autonomous driving mode to manual driving mode when encountering an incident. Additional outputs, such as lane type and distance between each type of lane, may be recorded and exported by modifying the MATLAB code to allow for further calculation of things such as determining whether or not the autonomous car drives out of its lane.

Automation is the future of transportation, and as researchers, it is to be expected that every study will have its limitations in addressing the problems and developing solutions. By assisting in these simulation-based experiments and by learning from the group how to design reusable scenarios with PreScan, I hope to achieve a deeper understanding of the autonomous system and also learn the approaches that engineers and scientists are currently working on to improve the issue of human trust in the autonomous system so that future driving experiences may be improved.

Bibliography

Feng, L. (2019). Project: NSF CRII: CPS: Cognitive Trust in Human-Autonomous Vehicle Interactions. Retrieved 2019, from https://www.cs.virginia.edu/~lufeng/project-cps-trust.html

Sadvandi, S., Halkias, D. (2019). Challenges of Human Factors Engineering in the Coming Transition to Autonomous Vehicle Technologies: A Multiple Case Study. ISM Journal of International Business, 3(1), 3–8. Retrieved from https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=139003520&site=eds-live

Strauch, C., Mühl, K., Patro, K., Grabmaier, C., Reithinger, S., Baumann, M., & Huckauf, A. (2019). Real autonomous driving from a passenger's perspective: Two experimental investigations using gaze behaviour and trust ratings in field and simulator. Transportation Research Part F: Psychology and Behaviour, 66, 15–28. https://doi.org/10.1016/j.trf.2019.08.013

TASS International (2019). PreScan Overview. Retrieved from https://tass.plm.automation.siemens.com/prescan-overview

Part II: STS Thesis Prospectus

Introduction and Background

The goal of this project is to reduce the influence of fake news on the public and find ways to promote trustworthy information on the internet. Increasingly, many people seek health and medical advice online. Targeted algorithms take advantage of this to bring users information that sounds relevant but may contain misinformation. Among the communication channels used, social media is the worst-hit area. As a not-for-profit network of scientists that aims to promote scientifically supported health and medical news to the public on the internet, Health Feedback conducted research into the top hundred most frequently shared health articles. Their results were shocking: a large majority of the health news shared on social media contains misleading or even false health information, and as a social media platform, Facebook contributes 96% of those articles (Teoh, 2019). The top three topics among the articles investigated involved disease or the treatment of disease, food and nutrition, and vaccinations. Of these three, the food and nutrition articles had the lowest credibility ratings because many were published on sites that had no scientific credentials (Teoh, 2019).

A combination of the bandwagon effect and preconceived ideas makes people believe in the targeted health news, although many of the writers are unable to support their claims with detailed information or credible origins. These fake advertisements and health posts are becoming increasingly hazardous to the public's health. People buy unnecessary and potentially harmful supplements and worse, practice dubious treatments that make them delay getting proper medical help (Booth, 2019). These factors constitute a huge challenge to health authorities and medical professionals to deliver verified health information to the public that will prevent misinterpretations and erroneous beliefs from spreading. My research, therefore, aims to study how targeted health advertisements and posts on social media that contain misinformation or emotionally manipulate their information affect the public.

Literature Review

Rowe and Alexander point to a possible cause of this problem: this is the loss of professional journalism coupled with the diversification of information platforms. The number of knowledgeable health and nutrition reporters has decreased in recent years due to the emergence of new information technologies and the changes in the social environment brought about by the internet (Rowe & Alexander, 2019). Furthermore, the lack of expertise by many writers in their subject matter has made things considerably worse. New journalists frequently lack knowledge in health-related matters and write articles that are misleading. However, Rowe and Alexander argue that the journalists are not the only people to blame: the marketing environment presses journalists to post information that has economic value to them (Rowe & Alexander, 2019). Last, but not least, there has been a shift in the quality of scientific studies, with many using data trends rather than lab tests from which to draw their conclusions. Data analysis algorithms providing "statistically significant" results are highly efficient as compared to a traditional research approach where hypothesis and experimental design need to be established and then tested. This can result in conclusions

that are less scientific, and the authors refer to this type of research practice as the "misuse of data analysis" and say that it should be discouraged (Rowe & Alexander, 2019).

Sommariva et al. (2018) conducted a case study to investigate the phenomenon of fake health articles in the social media. In this study, the social media sites were referred to as "powerful health communication platforms." The study analyzed stories shared on social networking sites relating to the Zika virus, and then categorized them into either verified health information or rumors. They discovered that rumors were shared three times more frequently than verified news (Sommariva et al., 2018). The purpose of this study was not to find ways to prevent fake health news from spreading on the social media, but rather to help health professionals take advantage of the social platforms and learn to make use of the mechanisms used to spread the fake news. That knowledge could help them break down the current communication barrier that prevents the public from accessing trustworthy health information (Sommariva et al., 2018).

Wang et al. (2019) describe how to track content through communication platforms. They introduced a systematic way of investigating the misinformation propagated by health-related papers circulating on the social media. This systematic approach categorizes the health articles by using multiple analysis methods, such as data extraction, co-citation analysis, and social network analysis. They concluded from their findings that unverified health news may lead to badly thought out social movements, which can have bad consequences. Therefore, remedial action needs to be taken by social and technology scientists as well as by health organizations and professionals (Wang et al., 2019).

Framework

Learning from Wang's approach, a theoretical framework for analyzing targeted health information on social media has been developed. There are two main components in this framework.

First, health advertisements and articles collected from social media for later analysis should be categorized into three main fields: medical treatment, disease prevention, and public health. The theory behind this division is that people react differently to health information in these different categories. They tend to believe health articles that contain advanced healthcare information, such as new medical interventions and disease prevention. People may be skeptical about public health articles initially, but repeated propagation of the information may influence them as time goes by. The reason for this is that public health articles account for a large proportion of online health articles, and they spread faster than articles in the first two categories. Using this theoretical framework will help in studying people's reactions toward health articles in the different categories and in understanding the psychology behind their acceptance of the views expressed.

Second, to differentiate between good and misleading health information, an accurate definition of "good" is needed. To simplify the process of analyzing online health articles, good health information is defined as that having credible sources, scientific support, and a minimum of subjective opinion. Credible sources refer to written materials and websites that are supported by the government, by research, and by academic

institutions and that have been published within the last ten years. Scientific support refers to scientific evidence provided by conducting experiments rather than from mere hypotheses. Subjective opinions should be minimal and account for less than ten percent of the information presented.

Using this framework can help in identifying and understanding misleading health information on the social media. It can also significantly reduce unnecessary time spent on research and prevent people from working with irrelevant data.

Methodology

To study this topic, the analysis will be divided into three phases, and each will use a different set of methodologies. This approach is inspired by Griffiths' study on the impact of online social networks on health systems (Griffiths, 2015).

In the first phase, health articles and advertisements will be collected from online social platforms, such as Facebook, Twitter, and Instagram, and they will be categorized into one of three groups: medical treatment, disease prevention, and public health. All articles will be documented in Microsoft Excel with basic information such as the publishing platform, the publishing date, and the category they belong to. Only articles and advertisements published in 2019 will be collected. For the advertisements, only those containing scientifically-related claims will be included.

In the second phase, further analysis will be conducted on the chosen articles and advertisements. Each one will be reviewed to extract more detailed information, such as the sources it refers to and the scientific evidence it uses. Data analysis and visualization software, such as Tableau, will be used.

In the third phase, interviews and questionnaires will be used to collect additional data. To get a better understanding of how health information in social media affects local student groups, college students will be interviewed. Questionnaires will be given to get a large study sample within a short period of time. The interviews and questionnaires will focus on understanding how students react to the different types of health information on social media and the impact of that information on the students' emotions.

Throughout the whole research process, to get a deeper understanding of health information on social media, the propagation of a single health news item will be tracked, namely vaping. Vaping is both a domestic and international health concern due to the confusing information given on social media. People have been debating the safety of vaping on the internet, and even the scientists have different views. Some claim that vaping is the "biggest health opportunity", while others believe that vaping causes potential health risks that may be more dangerous than people think, especially for young adults (CBS Interactive, 2019).

Timeline and Future Accomplish

The first phase of the research process will be completed by the end of January 2020, in the Spring semester. I will then spend the month of February working on phase

two. Toward the end of February, I will start to prepare for the interviews and questionnaire. I plan to include vaping as one of the topics that will be discussed during the interviews and included in the questionnaire. The interviews and questionnaire distribution will be conducted in March. The last week of March will be used to organize the data collected and to prepare for a presentation. With this research, I hope to make the audience aware of the impact of targeted health advertisements and posts on social media and help them understand the potential harm of the misinformation and emotionally manipulative information, which is often overlooked.

Bibliography

Booth, S. (2019, January 9). Fake Health News Dangerous. Retrieved from https://www.healthline.com/health-news/how-fake-health-news-may-be-influencing-you- to-make-dangerous-decisions

CBS Interactive. (2019, September 20). E-cigarette panic is ruining "biggest public health opportunity" in 120 years, scientist says. CBS News. https://www.cbsnews.com/news/vaping-deaths-us-blowing-single-biggest-public-health-opportunity-in-120-years-as-debate-e-cigarettes-scientist/.

Chrissyfarr. (2019, June 2). This doctor is recruiting an army of medical experts to drown out fake health news on Instagram and Twitter. Retrieved from https://www.cnbc.com/2019/05/31/doctor-recruiting-doctors-to-fight-fake-health-info-on- social-media.html.

Rowe, S., & Alexander, N. (2019). Fake Nutrition/Health News, Part 3: How (and Why) Did We Get Here? Nutrition Today, 54(4), 170–173. https://doi.org/10.1097/NT.0000000000000355

Sommariva, S., Vamos, C., Mantzarlis, A., Đào, L. U.-L., & Martinez Tyson, D. (2018). Spreading the (Fake) News: Exploring Health Messages on Social Media and the Implications for Health Professionals Using a Case Study. American Journal of Health Education, 49(4), 246–255. Retrieved from https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1183667&site=eds-live

Griffiths, F., Dobermann, T., Cave, J. A., Thorogood, M., Johnson, S., Salamatian, K., … Goudge, J. (2015). The Impact of Online Social Networks on Health and Health Systems: A Scoping Review and Case Studies. Policy & internet, 7(4), 473–496. doi:10.1002/poi3.97

Teoh, F. (Ed.). (2019, June 28). The Most Popular Health Articles of 2018, a Scientific Credibility Review. Retrieved from https://healthfeedback.org/the-most-popular-health-articles-of-2018-a-scientific-credibility-review/

Wang, Y., McKee, M., Torbica, A., & Stuckler, D. (2019). Systematic Literature Review on the Spread of Health-related Misinformation on Social Media. Social Science & Medicine, 240. https://doi.org/10.1016/j.socscimed.2019.112552