Prospectus

Deep Learning Based Occupants' Activity Prediction in a Smart Building Assistant System

(Technical Report)

Investigation on How to Teach Individuals to Identify, Avoid and Stop Fake News

(STS Research Paper)

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Haoran Zhu

Fall, 2019

Department of Computer Science

Signed: _____

Approved: Sean Ferguson _____ Date 12/11/2019

Tsai-Hsuan Ku, Department of Engineering and Society

Approved: _____ Date 11/19/2019

Haiying Shen, Department of Computer Science

Introduction

My technical research is about the application of deep learning on indoor environment control system in office buildings. Based on the research, several students at Dr. Haiying Shen's Pervasive Communication Lab and I designed a smart building assistance system that consists of various analysis on sensor data and a deep neural network (DNN) based prediction model. From experimental study on collected data, we found that our proposed assistance system can increase occupants' comfort levels and contribute to energy saving.

The authenticity of information has become a longstanding concern. On social networks, distorted, inaccurate and false information spreads at such a fast speed that it has tremendous potential for significant real-world impacts within a short period of time (Figueira, & Oliveira, 2017). My STS project mainly focuses on information literacy education. Our team hopes to design a systematic approach, which is understandable and doable for ordinary community members, on how to identify fake news and conduct fact-checking with existing available resources.

Technical Report

Nowadays, smart buildings equipped with fully or partially automated system to control the physical environments are becoming increasingly prevalent. Specifically, the physical working environments in office buildings, such as illumination, air quality, temperature and humidity have long been recognized as key features in ambient working environments that exert significant influence on occupants' comfort level and health conditions (Allen, MacNaughton, Laurent, & Eitland, 2015; Deuble, Dear, 2012). Moreover, there have been studies showing that buildings usually account for 30% to 40% of the total energy consumption and carbon dioxide emissions worldwide. Thus, proper indoor-environmental settings can greatly reduce energy usage and thus create more environmental-friendly buildings (Chen, Chou, Duri, Lei, & Reason, 2009).

With the purpose of realizing automated controlling system in smart office buildings, several students in Dr. Haiying Shen's Pervasive Communication Lab and I proposed a smart building assistance system that consists of sensor data analysis and deep neural network (DNN) based prediction model. Specifically, the proposed system predicts the physical features for current activity so that occupants feel comfortable and their levels of desire are satisfied based on their previous activities and the corresponding physical features. Meanwhile,
the proposed system also saves energy as a by-product of offering more comfort to occupants.

We first collect one-year-long smart building datasets from four different data sources (i.e., sensors, calendar, weather, and survey) All datasets are combined and extensive feature engineering is then performed on the collected dataset in order to prepare the raw data for prediction model. Specifically, this process includes the application of (1) discretization and one-hot encoding; (2) multiple-feature combination; (3) extraction of statistical values such as minimum and maximum to expand the feature space; (4) addition of time-series features; (5)

feature selection. We then propose a support vector regression (SVR) based prediction model and a hybrid DNN model, which consists of several long short-term memory (LSTM) blocks and a feed-forward DNN block so that different physical features (e.g., lighting, shading, air quality, and temperature) can be predicted for different building activities (e.g., meeting, lunch, and research activities). Eventually, we conduct extensive experimental studies to evaluate the performance of the proposed prediction models with other existing machine learning models in terms of mean absolute percentage error (MAPE) and accuracy.

We also conduct experiments using our collected data to figure out the activity-wise comfort levels and energy saving rates. From the experiments, we conclude that the proposed smart building assistance system is able to increase occupants' comfort level in office buildings by adaptively adjusting indoor environment features such as temperature and humidity. What's more, the proposed system also reduces the difference between indoor and outdoor temperatures and contributes to energy saving by smartly reducing the workloads of control systems such as air conditioning system when the office is not occupied.

In the future, we will focus on the following tasks: (1) Investigation on the interpretability of our automatic control system to make each adaptation of factors understandable to humans by adding reasoning with changes. (2) Integration of human knowledge or customized settings in separated building assistance systems. For example, users in different rooms may have different preferences. We will further explore different machine learning algorithms to take personal preferences into consideration. (3) Improvement on the robustness of our proposed system. DNN-based applications are sensitive to minor changes in the inputs and thus might be vulnerable to adversarial attacks. Improving the robustness of the system and avoiding malicious manipulations of indoor factors might be essential concerns in safety-critical scenarios such as a room that contains explosive which might explore when exposed to improper temperature conditions.

STS Thesis

Before we dive into the discussion about fake news, we need to be clear about what fake news exactly is. Ireland (2018) offered a rather accurate definition of fake news by listing what fake news is not. Fake news is not just bias. It's neither an adequately labelled opinion piece, nor a piece with logical fallacies. The term fake news should only apply to pieces that intentionally ignore, distort, or fabricate facts. Fake news is similar to yellow journalism in the 1890s, which does not contain truth and facts are missed or twisted (Mott, 2013).

Literature Review

Nowadays, while we are constantly absorbing information from a wide range of platforms every day, the chance of getting disinformed by fake news on those platforms has significantly increased as a result. Exciting results have been achieved by various technology in automatically identifying and flagging disinformation on online media platforms. However, as Neagu (2019, p.1) stated, "the only way to stop fake news is for you to take responsibility," it's important for modern citizens themselves to become critical consumers of news and information.

There has been research that investigates how to teach individuals to identify and avoid fake news. Some research has done a great job at depicting the interactions between different entities in the education on information literacy and these research has inspired me to develop an actor network that describes the spread of fake news as shown in Fig. 1. Regulations or instructions on the interactions among actors in the network will help individuals to identify, avoid and stop fake news.

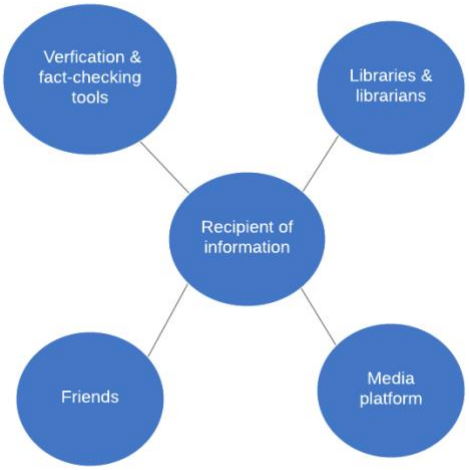

Fig. 1: Actor Network in Receiving and Stopping Fake News

Burkhardt (2017) and Davis (2016) proposed several useful ways for identifying disinformation, including "Be Careful about Who You Friend on Social Media", "Pay Attention to Domain and URL", "Read before Sharing", "Seek Information beyond Filter Bubble" and "Use Verification and Fact Check Tools". She analyzes the functionality of each method by identifying some actors that play significant roles in spreading or stopping fake news. The methods they proposed provide useful instructions for information recipients to regulate their interaction with other entities. For example, the "Be Careful about Who You Friend on Social Media" principle focuses on the interactions between recipient and recipients' friends. When we accept friend requests from strangers on social media, we make a psychological transition from regarding friend requestors as strangers to regarding them as friends. Such psychological shift will result in a certain amount of trust between the recipient and requestors, which makes the recipient more vulnerable to the information shared by those requestors, which might be bot accounts programmed to gain trust from social media users and spread disinformation. When people are more careful about accepting friend requests they received from strangers, the possibility of being influenced by fake news shared by friends decreases.

Moreover, Burkhardt (2017) targets the recipients of information literacy education as new generations, who are mostly students. The specificity in Burkhardt's approach reveals the diversity in both how people tend to receive information and how people interpret fake news. Matteson (2019) also suggests that information literacy educators should employ different theoretical understandings that describe and support information literacy when targeting at different groups of people.

STS Framework

Social construction of technology (SCOT) theory can be applied to analyze how we can deliver information literacy education to different groups effectively. Take fact-checking tool as an example. Undoubtedly, different people have different interpretations toward fact checking tools. While some people view these tools as truth provider, there are people that simply treat them as

another source of information, whose authenticity is also not guaranteed. The various interpretations will shape the design of fact checking tools. If people tend to hold suspicious attitude toward fact-checking tools, the probability of their information to be less authentic will increase as these fact-checking websites might have to loose their standards for validating information and include eye-catching news in order to attract users. Hence, for people who hold suspicious attitudes toward fact-checking attitude, information literacy programs specifically need to cultivate their confidence about authenticity of fact-checking tools.

This result is also justified by Actor Network Theory (ANT). According to ANT, the success of fact-checking tools and information literacy education programs in helping people identify fake news is not because it's the ultimate truth. In other words, we cannot expect the recipients of information literacy education programs to believe from the bottom of their hearts that fact-checking tools or information literacy is the panacea for identifying fake news and thus do exactly what they are advised to do. Instead, if we want people to accept information literacy, we have to understand the combination and interaction of elements that make it successful. This idea can also be illustrated by below strategies.

Future Research Directions

Existing information literacy projects mainly focus on social media. However, these projects usually work less effectively with people such as senior citizens who might not be very familiar with modern technology. For example, it's certainly harder for people with limited knowledge about software engineering to learn how to analyze domain and URL while trying to identify fake news. Therefore, information literacy education projects need to categorize methods of identifying fake news based on the degree of required technology knowledge and take this difference into consideration while delivering information literacy knowledge. Furthermore, while social media is prevalent in today's society, we cannot neglect the group of people who receive information through traditional media such as newspaper. Hence, it's crucial to come up with projects that target at traditional media to involve more people into the combat against fake news.

Beside proposals mentioned above, more research should be focused on how to incorporate technology into information literacy education in a larger extent. For example, online courses will help to spread information literacy education to a much larger user base than libraries alone. Also, more innovative ways should be applied in the combat against fake news. For example, Ireland (2018) came up with a project that delivers information literacy education through memes. Creative ways like this will catch the public's interests and attract more people to be educated. Specifically, I will conduct interviews and spread out surveys to Charlottesville residents and UVa students to collect data about what people's favorite ways of receiving education are. With those collected data, further improvements on information literacy education projects will be acquired and implemented.

Bibliography

Allen, J. G., MacNaughton, P., Laurent, J. G. C., Flanigan, S. S., & Eitland, E. S. (2015). Green Buildings and Health. *Current Environmental Health Reports*, *2*, 250–258. doi: 10.1007/s40572-015-0063-y

Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 U.S. Presidential election online discussion. *First Monday*, *21*(11). doi: 10.5210/fm.v21i11.7090

Burkhardt, J. M. (2017). Combating Fake News in the Digital Age. *Library Technology Reports* , *53*, 5–33.

Chen, H., Chou, P., Duri, S., Lei, H., & Reason, J. (2009). The Design and Implementation of a Smart Building Control System. *2009 IEEE International Conference on e-Business Engineering*, 255–262. doi: 10.1109/icebe.2009.42

Davis, W. (2016, December 5). Fake Or Real? How To Self-Check The News And Get The Facts. Retrieved November 5, 2019, from https://www.npr.org/sections/alltechconsidered /2016/12/05/503581220/fake-or-real-how-to-self-check-the-news-and-get-the-facts.

Deuble, M. P., & Dear, R. J. D. (2012). Green occupants for green buildings: The missing link? *Building and Environment*, *56*, 21–27. doi: 10.1016/j.buildenv.2012.02.029

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The Rise of Social Bots. *Communications of the ACM*, *59*, 96–104.

Figueira, Á., & Oliveira, L. (2017). The current state of fake news: challenges and opportunities. *Procedia Computer Science*, *121*, 817–825. doi: 10.1016/j.procs.2017.11.106

Davis, W. (2016, December 5). Fake Or Real? How To Self-Check The News And Get The Facts. Retrieved November 5, 2019, from https://www.npr.org/sections/alltechconsidered /2016/12/05/503581220/fake-or-real-how-to-self-check-the-news-and-get-the-facts.

Ireland, S. (2018). Fake news alerts: Teaching news literacy skills in a meme world. *The Reference Librarian*, *59*(3), 122–128. doi: 10.1080/02763877.2018.1463890

Karadzhov, G., Gencheva, P., Nakov, P., & Koychev, I. (2017). We Built a Fake News & Click-bait Filter: What Happened Next Will Blow Your Mind! *RANLP 2017 - Recent Advances in Natural Language Processing Meet Deep Learning*. doi: 10.26615/978-954-452-049-6_045

Matteson, M. L., & Gersch, B. (2019). Unique or ubiquitous: information literacy instruction outside academia. *Reference Services Review*, *47*(1), 73–84. doi: 10.1108/rsr-12-2018-0075

Mott, K. (2013, February 13). Yellow Journalism - Present and Past. Retrieved November 4, 2019, from https://www.americanhistoryusa.com/yellow-journalism-present-and-past/.

Neagu, S. (2019, March 4). The only way to stop fake news is for you to take responsibility. Retrieved November 5, 2019, from https://thenextweb.com/contributors/2019/01/12/ the-only-way-to-stop-fake-news-is-for-you-to-take-responsibility/.

Neely-Sardon, A., & Tignor, M. (2018). Focus on the facts: A news and information literacy instructional program. *The Reference Librarian*, *59*(3), 108–121. doi: 10.1080/02763877.2018.1468849

Varol, O., Ferrara, E., Davis, C. A., Menczer, F., & Flammini, A. (2017). Online Human-Bot Interactions: Detection, Estimation, and Characterization. *Eleventh International AAAI Conference on Web and Social Media*.