# *Prospectus*

# Prospectus

## Semantic Analysis Approaches to Predict User Interactions on Twitter

(Technical Report)

## Examine Impact of Artificial Intelligence on Chinese Telemedicine Startups and US-China Telemedicine Cooperations

(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

Daniel Wang

Fall, 2019

Department of Computer Science

Instructor Name: Sharon Tsai-hsuan Ku, Department of Engineering and Society

Tech advisor: Rich Nguyen, Department of Computer Science

## INTRODUCTION

User behavior Prediction from social media posts could potentially enable more robust active cognition models, and help construct more accurate simulations of the spread of online information to further the understanding of adversarial manipulation of such information. This technical report presents a network-based framework to predict a given Twitter user's reactions to a given set of information. I used natural language processing to tokenize each tweet from our collected data and constructed a semantic network from a given Twitter user's timeline. Then, we used text tokens as nodes, and integrated spreading activation to assign weights to nodes. After that, we calculated the unweighted centrality and weighted centrality with TF-IDF for tweets the user may see, and we used such values with other Twitter-specific features to train the classifier. The classifier implements a Convolutional Neural Network model that takes a list of tweets and assign probabilities of different types of user behaviors. In our evaluation, the implementation of the semantic network has generally increased prediction accuracy from baseline models. I also provide several potential applications of our framework.

In STS Thesis, I would instead talk about telemedicine, a subsection of smart health which involves the practice of caring for patients remotely when the provider and patient are not physically present with each other. Healthcare contributes to any average person's good health and well-being, a fundamental human right. However, through the research and talks to personnels in University of Virginia's health system in Charlottesville, which is a prime example of many small cities in the United States surrounded by largely rural areas, I have discovered that many barriers currently exist for residents in Charlottesville and its surrounding areas, such as language barriers for new Hispanic immigrants, and barriers for rural patients to schedule appointments or seek emergency services due to being far away from the only major healthcare facility in the area. To mitigate such problems, I first looked for international developments about telemedicine, especially Chinese telemedicine startups companies that are apply Artificial Intelligence to advance the development of telemedicine by applying the SCOT diagram and the Value-Sensitive Design Diagram. In addition, I would propose an app-based solution to connect doctors and patients who have difficulties travelling to the hospital by video chat, providing translation software for such video chat to mitigate the language barrier for some patients, and allow doctors to prescribe medicine online and ship such medicine directly to the patient.

## TECHNICAL TOPIC

To combat distortions from foreign entities on Social media, DARPA has created the SocialSim challenge in 2018 to develop innovative technologies for high-fidelity computational simulation of online social behavior, hoping it could enable a deeper and more quantitative understanding of adversaries' use of the global information environment than is currently possible using existing approaches. This study proposes a framework to further enhance the simulation of user behavior by incorporating network analysis, and findings from this study could prove pivotal in informing and improving the simulation of social networks to help combat the manipulation of social networks by foreign entities.

This study seeks to resolve the following premise: Given a Twitter user A, and a set of tweets $t_1$, $t_2$, $t_3$...$t_i$ the user A has interacted with, predict how the user may interact (reply,

retweet, quote, like) to each tweet in the set. I resolve the premise by collecting the tweets the user may see and constructing and analyzing semantic networks of the given user's timeline data. With the rising research interest with Twitter data, there are several works that slightly overlap with our work. With regard to prediction on Twitter, Petrovic, Osborne, and Lavrenko studied the problem of predicting whether a user will retweet a particular tweet, TwitterMancer proposed predicting what type of interaction may take place between two different users with link prediction, and RealGraph utilizes a framework to compute relationship strength for ties based on directed interactions between users. Our work differs from these prior work by being the first to construct a semantic network and utilize NLP techniques with Twitter data to predict how a given user may react to a set of tweets. In addition, this work does employ a similar approach to Kang, Swarup to utilize Semantic Network and Spreading Activation techniques for Twitter Analysis. However, this work seeks to make predictions about user behaviors rather than analyze user sentiment on vaccines.

Our framework is implemented with Python. It has three major components: data collection, to collect Twitter timeline information from the user; feature engineering, that contains the generation of a semantic network, the implementation of Spreading Activation with word embedding vector, and the computation of different centrality and weighted centrality using TF-IDF values; classification, contains the construction of a baseline random forest model and the implementation of a Convolutional Neural Network model for better classification accuracy results.

The technical advisor, Professor Rich Nguyen has generously provided access to Rivanna, UVA's research computing high-performance cluster for the computation of multiple features, as well as the training for classification models. So far, we collected a total of 27 million tweets created before August 15, 2019. In all, about 68.23 % of the sampled users' posts from their timelines were interactions, while about 31.77% of the sampled users' posts were original. For the tweets that are interactions, we labeled each tweet based on different types of interactions. Like was the most common interaction overall, in which about 48.37% of sampled users' interactions were likes, followed by retweet(30.23%), reply(19.00%), and quote(2.40%), respectively. After performing the experiments on the collected data. On average, models with subgraph centralities achieved a 0.5% increase in prediction accuracy over the baseline model, with a maximum improvement of 1.635%. By the end of the year, this technical project hopes to further increase the accuracy of such predictions, followed by a technical paper published on the methods employed with this project.

## STS THESIS: EXAMINE IMPACT OF ARTIFICIAL INTELLIGENCE ON CHINESE TELEMEDICINE STARTUPS AND US-CHINA TELEMEDICINE COOPERATIONS

### INTRODUCTION

In the field of telemedicine, a subsection of healthcare that involves the cooperation between traditional medical services and the fast-involving Internet, China is one of the world leaders in both policy implementation and technological development. In terms of the policy, China's national government has been encouraging the application of medical big data since 2015, as well as rolling out policies dealing with artificial intelligence development in the healthcare industry(Kong et al., 2019). In terms of technological development, in August 2019, robots guided by doctors in a remote location completed major portions of three simultaneous

orthopedic surgeries through 5G wireless technology (Juan, 2019), after a successful brain surgery done by a Chinese doctor who inserted a stimulation device in the brain of a Parkinson's patient nearly 1,900 miles away in March (Frost, 2019). Both surgeries are world-first. From medical training to providing emergency assistance, such operations may hopefully help people in areas that have difficulty to access healthcare and cut the cost of surgeries in the long term. Meanwhile, the United States has engaged in a trade war with China. In 2019, the Trump administration has placed a list of Chinese companies such as Huawei and Hikvision (Swanson, 2019) on the entity list to stop China's unfair trade practices and state-subsidized development in advanced technology, such as 5G, Artificial Intelligence, and telemedicine. Besides, major tech companies in the US have also allocated resources to develop various telemedicine platforms to compete with Chinese counterparts. For example, Amazon recently launched Amazon Care, a virtual primary care clinic with an option for nurses to visit employees in the home (Farr, 2019), and Google has also launched Google Health, described as using Artificial Intelligence to assist in diagnosing cancer, predicting patient outcomes, preventing blindness, and much more (Google, 2019).

With China's dramatic advancement in Artificial Intelligence and telemedicine and United State's swift response to counter China's rise, I would seek to answer the following research questions in this project:

1. How does the national government's policy in healthcare and Artificial Intelligence influence the development of telemedicine startup companies in China?
2. How do US telemedicine startup companies compete or cooperate with their Chinese counterparts?

## LITERATURE REVIEW

First, I researched the definition of telemedicine to fully understand the subject. According to a World Health Organization's recent report, telemedicine is defined as "The delivery of health care services, where distance is a critical factor, by all health care professionals using information and communication technologies for the exchange of valid information for diagnosis, treatment and prevention of disease and injuries, research and evaluation, and for the continuing education of health care providers, all in the interests of advancing the health of individuals and their communities" (Ryu, 2012). This long and broad definition suggests that telemedicine, just the technology it relies on, is an open and constantly evolving subject, as new advances in both medical and informational technology open new possibilities to help address new and existing challenges for the health needs of societies and communities. Naturally, with the recent technological breakthroughs in Artificial Intelligence, it's in both government and private enterprises' best interest to employ this advanced model in telemedicine to address both existing and future challenges in healthcare, and countries who failed to fully take advantage of this technology would be left years behind and unable to cooperate or compete with other countries in the future.

Then, I researched about China's policy about encouraging the development of more advanced healthcare solutions and Artificial Intelligence. On October 25, 2016, China formally passed the blueprint of "Healthy China 2030," working towards the national goal of reaching a health standard on par with developed countries by 2030. In this initiative, a shift in concentration towards coordinated development of health-based economy from a previous pursuit of rapid

economic growth was underlined, and the advancement of telemedicine, a subsection of smart health which involves the practice of caring for patients remotely when the provider and patient are not physically present with each other, is at the center of this new development (Tan et al., 2017). Also, China's State Council in July 2017 published "A Next-Generation Artificial Intelligence Development Plan", aiming to become the world leader in Artificial Intelligence in a domestic industry worth US$150 billion by 2030 (Estevadeordal et al., 2018). With these stated official policies, China may use such national initiative to encourage more students to study technology and medicine, therefore building a high-quality talent pool for domestic companies to develop state-of-the-art healthcare solutions based on telemedicine. Besides, China's unique state capitalism model means the Chinese government may invest in domestic companies that specialize in the integration of healthcare and Artificial Intelligence, to both fulfill the policy agenda and help these companies better compete internationally.

In addition, I researched about China's telemedicine startups and explored such companies' use of Artificial Intelligence. One of the telemedicine startups I examined is called Ping An Doctor. This platform has developed "One-Minute Clinics", which combines a small waiting room about the size of a phone booth with an interactive AI-based system that solicits patient information, builds a medical history, and then communicates the relevant information to a remote cloud-based physician. The patient would then be able to purchase medicine from the adjoining "Smart Medicine Cabinet" or a nearby pharmacy (Price et al., 2019). From this example, it's apparent that the Chinese government's policy has succeeded in encouraging some Chinese telemedicine startups to adopt the integration of healthcare and Artificial Intelligence. It's also noteworthy, however, that the lack of privacy protection laws in China may also help Chinese medicine startups to achieve world-first innovations. In western democracies, it's very unlikely that the government would permit the use of people's medical records by just a startup company. However, it seems the Chinese government places the advancement in technology over an individual's privacy rights, which in this case did contribute to the otherwise unlikely success of this startup.

Finally, I also researched about telemedicine startups in the US and such startup's corporations or competitions with Chinese telemedicine startups. One of the telemedicine startups I examined is called DocFlight. For a couple of thousand U.S. dollars, DocFlight affords patients in China access to its network of several hundred physicians in the U.S. from top-flight hospitals. That includes the Dana-Farber Cancer Institute and NYU Langone Medical Center, among others. During the virtual visit, the patient may meet an appropriate specialist online via video chat, with Medical records translated by computer. After any follow-up questions about the patient, findings and recommendations are summarized by the U.S.-based physician. The consultation usually takes around 45 minutes (NBC, 2019). From this example, it's apparent that this US-based startup attempts to attract wealthy Chinese patients, which many patients distrust about China's healthcare system (Meyers, 2018). If this startup was to succeed and attract an enormous amount of Chinese patients, it would provide meaningful competition with China's local hospitals and cause them to lose valuable patients, thus having fewer resources to hire the best medical professionals or purchase the latest medical equipment, which would cause the reduction of quality in China's healthcare system. However, this startup also has the potential to foster the cooperation of Chinese and American doctors to provide the most accurate health assessment for the patient.

## FRAMEWORK

This project employs both Social Construction of Technology (SCOT) diagram and Value-Sensitive Design (VSD) as the framework to better understand how China's government policy, Chinese telemedicine startups' implementation of Artificial Intelligence, and US telemedicine startups' cooperation or competition with Chinese counterparts affect the corporate structure of telemedicine startups and the respective impact on relevant stakeholders. In addition, further interviews with Chinese students would be conducted to understand their perspectives about government policy and the development of Chinese telemedicine startups.

## METHODS

To acquire the information necessary to use the framework and answer the research question, I plan to conduct individual case studies with Chinese and US telemedicine startups and apply the SCOT diagram and the VSD to better understand how stakeholders and design of the startup may reflect the influence from Chinese government, Artificial Intelligence, and competition or cooperation between the US and China. In addition, I plan to conduct interviews with Chinese students to better understand their views on the effectiveness of the government policy with healthcare and Artificial Intelligence, and their views on the US-China competition and cooperation.

1. Case Study.
    a. SCOT Diagram. Example: DocFlight

    In the case of DocFlight, the stakeholders are the patients, healthcare professionals in China and the US, and startup company executives. During a virtual visit by the patient, as the patient most likely distrust the healthcare system in China, it expects to gain insightful diagnosis about the disease. At the same time, the doctors in the US would reap extra compensation outside of their work hours and may gain a deeper understanding of some of the common diseases in China. In addition, startup company executives would profit from portions of the consultation fee paid by the patient, to expand and reach more Chinese patients and connect with more doctors in the US. However, the more virtual visit, the more patients that doctors in China would lose, which could have helped many struggling doctors or hospitals in China to stay afloat.

    b. VSD. Example: Ping An Good Doctor

    In the case of Ping An Good Doctor, the value is the accessibility to healthcare, the norms are easy access to prescription, patient and physician convenience, access to medication, and affordable, and the design requirements are allowing users to efficiently find their prescription and pick up the medication quickly, enable healthcare professionals to help more patients than the traditional hospital setting, and give better prescription recommendations to patients from rural areas than inferior rural hospitals and clinics, and reduce the extra fee incurred for patients during prescription. The first design requirement was satisfied by automating the traditionally time-consuming process of retrieving and analyze relevant patient information to AI to reduces patients' total wait time spent in the clinic, the second design requirement was satisfied by allowing one physician to

handle more patients by remote connections to the mini-clinics, which reduce physicians idle time in the traditional hospital setting, the third design requirement was satisfied by help rural patients that may not be able to travel to world-class hospitals in China's major cities to prescribe directly from the physicians in these hospitals online, which would often mean prescriptions that have better understanding of the exact ailment of the patient, and help the patient for a speedier recovery. At last, just like booking taxis with Uber using an app, such mini-clinic may reduce other necessities in a traditional hospital setting for prescription, which would mean cheaper prescriptions for patients. As all the design requirements were satisfied, the values of the design to provide better access to healthcare for patients was achieved.

2. Interviews

Interviews with Tsinghua and Zhejiang University students would be conducted to gain insights about their opinions on Chinese government's policy on encouraging the development of telemedicine and Artificial Intelligence, their personal experience about interacting with apps utilizing telemedicine and Artificial Intelligence, and their attitudes and opinions about the cooperation or competition from telemedicine startups in the US. Such interviews may provide a diverse array of opinions that help understand the first-hand experience of Chinese people's interaction with telemedicine.

## DISCUSSION AND NEXT STEPS

Ultimately, the evaluation of the effects of China's healthcare and Artificial Intelligence policy would hopefully help policymakers around the world to better establish guidelines or recommendations for either government or private enterprises to achieve progress in telemedicine. In addition, the evaluation of US telemedicine startups' corporations or competitions with their Chinese counterparts would hopefully help the general public to better understand the pros and cons of meaningful engagement with technology from China. In the future, I plan to introduce an app-based platform similar to those found in Chinese telemedicine startups to the Charlottesville area to help mitigate barriers and enable more residents in Charlottesville to gain access to quality healthcare.

## BIBLIOGRAPHY

Kong, X., Ai, B., Kong, Y., Su, L., Ning, Y., Howard, N., … Fang, Y. (2019). Artificial intelligence: a key to relieve China's insufficient and unequally-distributed medical resources. American journal of translational research, 11(5), 2632–2640.

Juan, D. (2019, August 29). In a first, three surgeries conducted using robots, 5G. Retrieved December 5, 2019, from China Daily website: https://www.chinadaily.com.cn/a/201908/29/WS5d670e17a310cf3e355686fa.html

Frost, C. (2019, August 16). 5G is being used to perform remote surgery from thousands of miles away, and it could transform the healthcare industry. Retrieved December 5, 2019, from Business Insider website: https://www.businessinsider.com/5g-surgery-could-transform-healthcare-industry-2019-8

Swanson, A. (2019, October 17). US to blacklist 28 Chinese entities over abuses in Xinjiang. Retrieved December 5, 2019, from The New York Times website:

https://www.nytimes.com/2019/10/07/us/politics/us-to-blacklist-28-chinese-entities-over-abuses-in-xinjiang.html

Farr, C. (2019, September 24). Amazon launches Amazon Care, a virtual medical clinic for employees. Retrieved December 5, 2019, from CNBC website https://www.cnbc.com/2019/09/24/amazon-launches-employee-health-clinic-amazon-care.html

Google, Inc. Retrieved December 5, 2019, from Google Health website https://health.google/

Ryu S. (2012). Telemedicine: Opportunities and Developments in Member States: Report on the Second Global Survey on eHealth 2009 (Global Observatory for eHealth Series, Volume 2). Healthcare Informatics Research, 18(2), 153–155. doi:10.4258/hir.2012.18.2.153

Xiaodong Tan, Qian Wu, Haiyan Shao, Global commitments and China's endeavors to promote health and achieve sustainable development goals, (Published online 2018 Apr 12), J Health Popul Nutr. 2018; 37: 8. Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5898031/#

Estevadeordal, Antoni; Beliz, Gustavo; Estevez, Elsa; Integration and Trade Journal: Volume 22: No. 44: July, 2018: Planet Algorithm: Artificial Intelligence for a Predictive and Inclusive form of Integration in Latin America. Available at https://publications.iadb.org/en/integration-and-trade-journal-volume-22-no-44-july-2018-planet-algorithm-artificial-intelligence

Price, William Nicholson, Artificial Intelligence in the Medical System: Four Roles for Potential Transformation (February 25, 2019). Yale J.L. & Tech. (Forthcoming); Yale J. Health Pol'y L. & Ethics (Forthcoming); U of Michigan Public Law Research Paper No. 631. Available at SSRN: https://ssrn.com/abstract=3341692

Fuchs, C. (2017, August 7). This Startup Connects U.S. Doctors with Patients in China. Retrieved December 5, 2019, from NBC website https://www.nbcnews.com/news/asian-america/startup-connects-u-s-doctors-patients-china-n788961

Meyers, J. (2018, January 11). Distrusting China's medical system, patients turn to U.S. doctors online. Retrieved November 9, 2019, from Los Angeles Times website: https://www.latimes.com/world/asia/la-fg-china-cancer-us-20180111-story.html