AFRL-AFOSR-UK-TR-2024-0016

Networked Social Influence and Acceptance in a New Age of Crises

Guo, Weisi
CRANFIELD UNIVERSITY
COLLEGE ROAD
BEDFORD, , MK43 0AL
GBR

01/09/2024
Final Technical Report

DISTRIBUTION A: Distribution approved for public release.

Air Force Research Laboratory
Air Force Office of Scientific Research
European Office of Aerospace Research and Development
Unit 4515 Box 14, APO AE 09421

REPORT DOCUMENTATION PAGE

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE 20240109	**2. REPORT TYPE** Final	**3. DATES COVERED** **START DATE** 20200901	**END DATE** 20230831

4. TITLE AND SUBTITLE
Networked Social Influence and Acceptance in a New Age of Crises

5a. CONTRACT NUMBER	**5b. GRANT NUMBER** FA8655-20-1-7031	**5c. PROGRAM ELEMENT NUMBER**
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**

6. AUTHOR(S)
Weisi Guo

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) CRANFIELD UNIVERSITY COLLEGE ROAD BEDFORD MK43 0AL GBR	**8. PERFORMING ORGANIZATION REPORT NUMBER**

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) EOARD UNIT 4515 APO AE 09421-4515	**10. SPONSOR/MONITOR'S ACRONYM(S)** AFRL/AFOSR IOE	**11. SPONSOR/MONITOR'S REPORT NUMBER(S)** AFRL-AFOSR-UK-TR-2024-0016

12. DISTRIBUTION/AVAILABILITY STATEMENT
A Distribution Unlimited: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
In this project, we set out the ambition to create a modeling framework, grounded in real and diverse online social network data, on how to combine complex social network analysis with nonlinear human behaviour dynamics. This has the advantage of being able to uncover hidden influencers. The objectives are:
1. Who are the real influencers: finding influencers on social networks as a function of network topology (graph algebra) and nonlinear influence behaviour dynamics (differential equations, data embeddings, etc.).
2. Explainability: identify what parameters cause influencers to be influential: e.g. network place (eigenvalue) and personal behaviour dynamics. Analytical explainability also provides a pathway towards understanding how to sparsely monitor vast networks, e.g., which key influencers should be monitored and why.
This has widespread applications in counterterrorism, social media analysis, and understanding how we can effectively spread policies and create acceptance to technologies in an age of emerging crises.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	SAR	14

19a. NAME OF RESPONSIBLE PERSON NANDINI IYER	**19b. PHONE NUMBER** *(Include area code)* 314-235-6161

Standard Form 298 (Rev.5/2020)
Prescribed by ANSI Std. Z39.18

Section 2: Technical Report

Project Title: Networked Social Influence and Acceptance in a New Age of Crises
FA9550-17-1-0056 P00003
Start Date: September 2020
PI: Prof. Weisi Guo
Student: Mrs. Bailu Jin

Accomplishments

Research Objectives: Please list the main research objectives of this project

In this project, we set out the ambition to create a modeling framework, grounded in real and diverse online social network data, on how to combine complex social network analysis with nonlinear human behaviour dynamics. This has the advantage of being able to uncover hidden influencers. The objectives are:

1. Who are the real influencers: finding influencers on social networks as a function of network topology (graph algebra) and nonlinear influence behaviour dynamics (differential equations, data embeddings, etc.).
2. Explainability: identify what parameters cause influencers to be influential: e.g. network place (eigenvalue) and personal behaviour dynamics. Analytical explainability also provides a pathway towards understanding how to sparsely monitor vast networks, e.g., which key influencers should be monitored and why.

This has widespread applications in counterterrorism, social media analysis, and understanding how we can effectively spread policies and create acceptance to technologies in an age of emerging crises.

Please provide details of accomplishments during this reporting period.

During this report period, we are focusing on the following three tasks:

Task 1: Data Driven Modeling Social Media Influence using Differential Equations (Oct 2021 – Mar 2022)
Objective: We aim to model the personal opinion evolution process under the effect of the online social influence as a function.
Development: In online social networks, we use the text content posted by one individual to represent the individual's opinion. In our case study, we gather the COVID-19 specific tweets content as the initial input, then represent tweets using uni-dimensional continuum by word-embedding and dimensional reduction.

Significant Result: Our research on the COVID-19 topic and for the account analysed shows that social media users primarily shift their opinion based on influencers they follow and self-evolution of opinion over a long-time scale is limited.

Task 2: Behaviour Informed Influence Ranking Measures (Apr 2022 – Dec 2022)
Objective: We aim to address this multi-disciplinary research area by introducing and connecting the diverse methodologies for identifying influential nodes. The key novelty is to review connections and cross-compare different multi-disciplinary approaches that have origins in graph theory and sampling, control theory, natural language processing, and social psychology.
Development: We categorise the opinion identification methods into four main categories, Topology-based Centrality, Topic-sensitive Centrality, Control- and Sampling-based Centrality. These categories define opinion leaders in distinct ways and ingest different data features.
Significant Result: Through a case study, we perform a comparative analysis of multiple methodologies. The result shows that a horizontal comparison among different ranking strategies is challenging, due to the disparate criteria utilised by the methods. There may be some overlap between the identified opinion leaders through various methods, yet their correlation and causality require further studies.

Task 3: Developing an influential Tweet Bot (Jan 2023 – Dec 2023)
Objective: Build an influential tweet bot using reinforcement learning with ChatGPT. This task builds upon the findings and methodologies of to create a practical application that leverages the principles of opinion evolution and user influence in online social networks.
Development: Build a social community simulator: 1) build a simulator using ChatGPT; 2) evaluate the platform based on the previous findings and methodologies. Build an influential Bot: 1) offline train the RL agent on the community simulator; 2) evaluate the RL performance by measuring the improvement of ranking it generates in the simulated environment.
Significant Result: The probability distribution shows that artificial dataset may be too uniform and lacks variation. The regression result shows that the simulated environment also fit on the opinion evolution model (Task 1) with lower R^2 value.

How were the results disseminated to communities of interest?
Nothing to report.

What do you plan to do during the next reporting period to accomplish the goals and objectives?
Nothing to report.

Impacts

Development of the principal discipline(s) of the project
Current algorithms are not suitable for modeling complex networked influence, as they do not contain meaningful explicit dynamics for social behaviour. Our project proposed incorporating behavioural models that can yield more precise insight into networked online social influence. Analytical explainability also provides a pathway towards understanding how to sparsely monitor vast networks, e.g., which key influencers should be monitored and why.

Other disciplines
Nothing to report.

Describe the impact in this reporting period on the development of human resources
Nothing to report.

Describe the impact on teaching and educational experiences
Nothing to report.

Describe the impact in this reporting period on physical, institutional, and information resources that form infrastructure
Nothing to report.

Impact on society beyond science and technology
This project has widespread applications in counterterrorism, social media analysis, and understanding how we can effectively spread policies and create acceptance to technologies in an age of emerging crises.

Changes

Changes in approach
Nothing to report.

Problems or delays
Nothing to report.

Expenditure Impacts
Nothing to report.

Significant changes in the use or care of human subjects, vertebrate animals and/or biohazards
Nothing to report.

Changes to the primary place of performance from that originally proposed
Nothing to report.

Technical Updates

Task 1: Data Driven Modeling Social Media Influence using Differential Equations (Oct 2021 – Mar 2022)

Aim: Individuals modify their opinions towards a topic based on their social interactions. Empirical research in psychology has shown that individuals influence each other by seeking similarity or conforming under social pressure. Opinion evolution models conceptualize the change of opinion as a uni-dimensional continuum, and the effect of influence is built by the group size, the network structures, or the relations among opinions within the group. In this research, we aim to model the personal opinion evolution process under the effect of the online social influence as a function.

Background Theory: Starting from the formal model, the change of opinion is always conceptualized as a uni-dimensional continuum and determined by the size, the network structures, or the relations among opinions within the group. Based on previous psychology models, we model the opinion evolution process as a function of online social influence using an ordinary differential equation (ODE) that reflects the social network influencer interactions. We outline our system model in Figure 1.

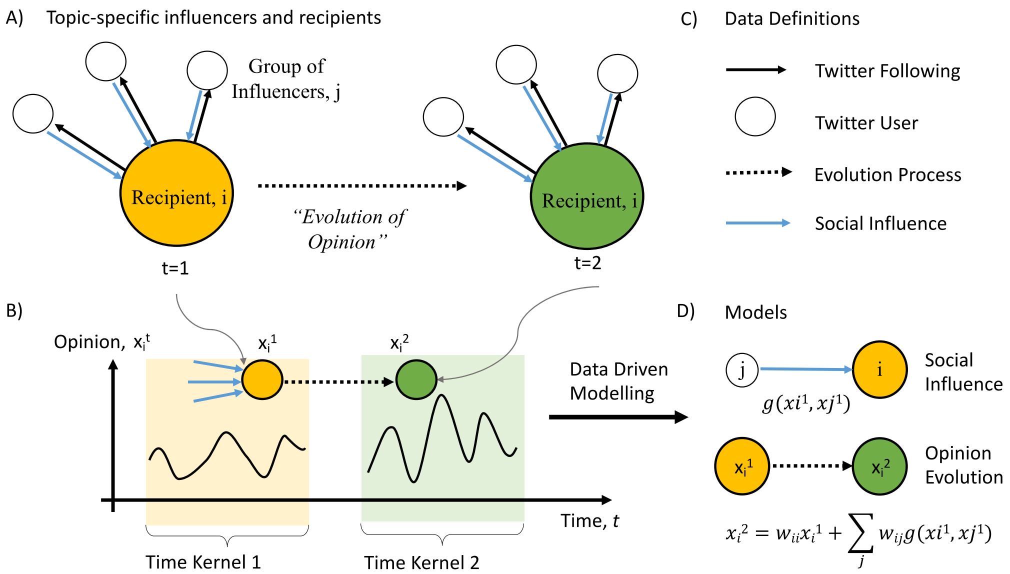


Figure 1: A) shows the recipient i and the influencers j (determined by Twitter following), and the influencers provide the forces of social influence on the recipient's opinion over time. B) shows the evolution process of recipient i under the forces from influencers with a time kernel. C) shows data definitions. D) shows definitions of the

social influence and opinion evolution models.

Data: Here we choose COVID-19 as our specific topic. COVID-19 pandemic has been an ongoing global pandemic since December 2019. Discussions on disease symptoms, prevention, vaccine, and local policies are widely spread online. We used the Twitter API to gather each user's "Following" relationship and tweet contents. Figure 2 shows the number of all Tweets and topic-specific Tweets generated from active users from March 2020 to Feb 2022, including 175624 tweets and 85946 topic-related tweets in total.

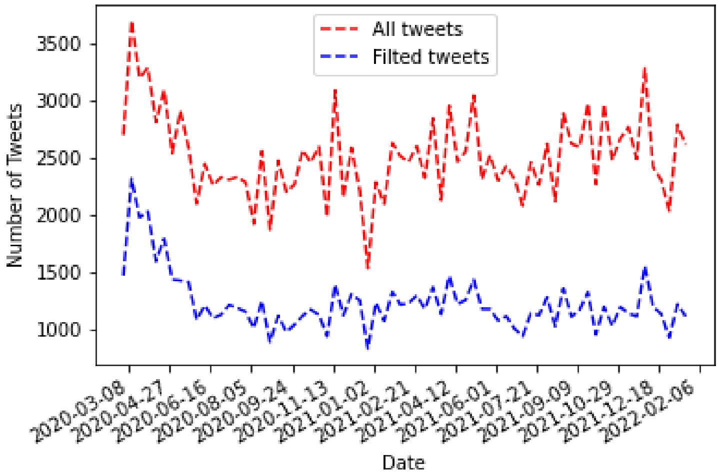


Figure 2: Number of all Tweets and topic-specific Tweets from March 2020 to Feb 2022.

Pipeline: The previously introduced social influence model mainly conceptualizes the opinion using pro-event and post-event psychology survey questions. In online social networks, we use the text content posted by one individual to represent the individual's opinion. In our case study, we gather the COVID-19 specific tweets content as the initial input, then represent tweets using uni-dimensional continuum by word-embedding and dimensional reduction. The process is shown in Figure 3.

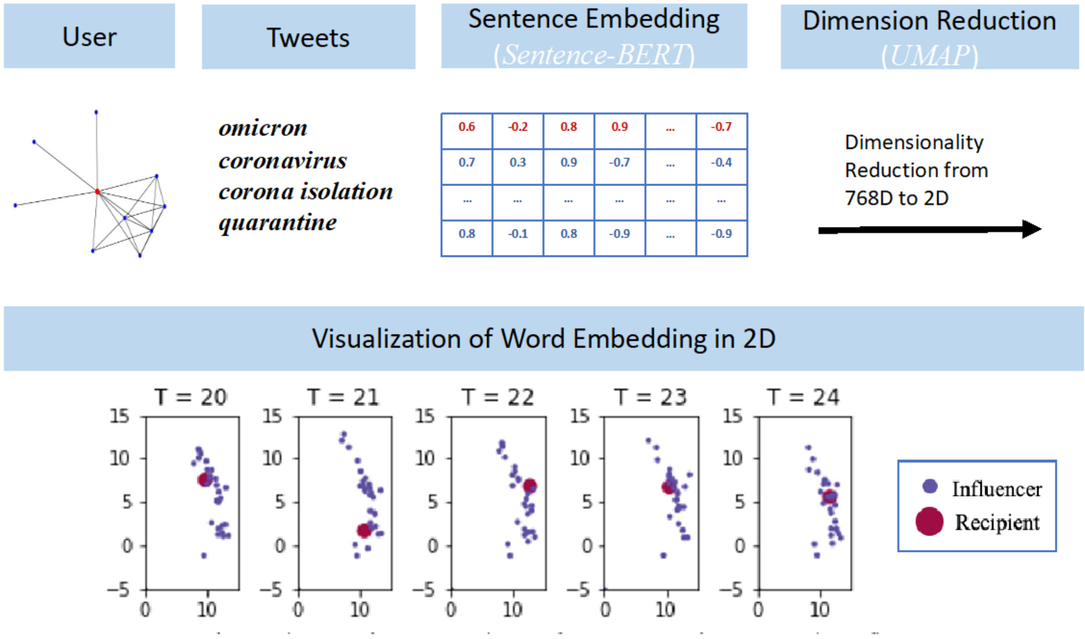


Figure 3: Process of word-embedding, dimensionality, and visualization of users' opinions on COVID-19 topic.

Result: We perform our analysis on 87 active users with corresponding influencers on the COVID-19 topic from 2020 to 2022. As shown in Table 1, the regression results demonstrate that 99% of the variation in the quantified opinions can be explained by the way we model the connected opinions from their influencers. Our research on the COVID-19 topic and for the account analysed shows that social media users primarily shift their opinion based on influencers they follow and self-evolution of opinion over a long-time scale is limited.

No. of Influence Models	Observations per Model
87	69
$\hat{R}$ Mean	Adj. $\hat{R}$ Var
0.98232	0.00769
Pro F-statistic Mean	Pro F-statistic Var
0.00012	1.26e-06

Table 1: Regression Result.

Task 2: Behaviour Informed Influence Ranking Measures (Apr 2022 – Dec 2022)

Background: Online social networks (OSNs) provide a platform for individuals to share information, exchange ideas and build social connections beyond in-person interactions. For a specific topic or community, opinion leaders are individuals who have a significant influence on others' opinions. Detecting and modeling opinion leaders is crucial as they play a vital role in shaping public opinion and driving online conversations. Existing research have extensively explored various methods for detecting opinion leaders, but there is a lack of consensus between definitions and methods. It is important to note that the node importance term "node centrality" in graph theory does not necessarily align with the concept of "opinion leader" in social psychology.

Aim: We aim to address this multi-disciplinary research area by introducing and connecting the diverse methodologies for identifying influential nodes. The key novelty is to review connections and cross-compare different multi-disciplinary approaches that have origins in graph theory and sampling, control theory, natural language processing, and social psychology. We discuss how they tell a different technical tale of network influence and also propose how some of the approaches can be combined via networked dynamical systems modeling, underpinned by social psychology models. An OSN case study is performed on Twitter data to compare the performance of different methodologies discussed to elucidate the research progression and inspire further research in this cross-disciplinary area.

Methodologies:
We categorise the opinion identification methods into four main categories, Topology-based Centrality, Topic-sensitive Centrality, Control- and Sampling-based Centrality. These categories define opinion leaders in distinct ways and ingest different data features. Topology-based centrality mainly concentrates on the network structure. In this context, opinion leaders are defined as individuals who occupy the most significant position within the social group. When user semantic content is taken into consideration, the Topic-Sensitive Centrality facilitates the identification of opinion leaders within specific topics. This approach helps identify influential users capable of disseminating topic-related information and influencing opinions within specific contexts. Additionally, real-time content can be utilised as a representation of the dynamic opinion states of users, which can be used to build a mathematical model to describe the evolution of opinion states. Leveraging the dynamic influence model, control methodologies aim to identify individuals who can steer the direction of overall opinion. Finally, graph sampling methodologies focus on identifying a specific subset of opinion leaders who, despite their limited numbers, can be instrumental in reconstructing the comprehensive opinion network.

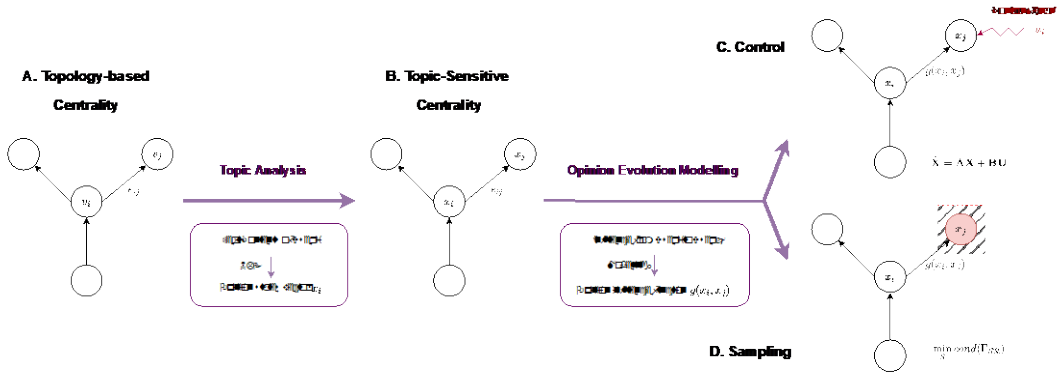


Figure 4: Opinion Leader Definition under various user ranking measure. A) Topology-based: Individuals who occupy the most significant position within the social group. B) Topic-Sensitive: Individuals who disseminate topic-related information or influence the opinions of others within a specific context. C) Control: Individuals who can steer the direction of overall opinion. D) Sampling: Individuals who can be instrumental in reconstructing the comprehensive opinion network.

Conclusion: Through a case study, we perform a comparative analysis of multiple methodologies. The result shows that a horizontal comparison among different ranking strategies is challenging, due to the disparate criteria utilised by the methods. There may be some overlap between the identified opinion leaders through various methods, yet their correlation and causality require further studies.

Definition	Concept Type	Detection Method
The information was first known by the opinion leaders, then spread by them to others.	Spread	Topology-based
Individuals who disseminate topic-related information effectively.	Spread	Topic-sensitive
Utilise an informal approach to influence others' attitudes or behavior to achieve a desired result.	Control	Control-based
Opinion leaders are well-informed, trusted in groups, and their opinions are representative.	Representative	Sampling-based

Table 2: Comparative analysis of opinion leader definitions and their detection methodologies.

Task 3: Developing an influential Tweet Bot (Jan 2023 – Dec 2023)

Aim: Build an influential tweet bot using reinforcement learning with ChatGPT. This task builds upon the findings and methodologies of to create a practical application that leverages the principles of opinion evolution and user influence in online social networks.

Background: LLM such as GPT was used as a source of simulated respondents in several cases. GPT has been tested for multiple tasks and proved to be effective without gradient updates or fine-tuning. Natural-language agent-based model of argumentation has been used to simulate the argumentative opinion dynamics. The explicit stance used in traditional simulation of opinion dynamics can be represented by implicit argument from natural-language ABMAs. In our case, we combine the agent-based modelling and LLM to extend the validity of agents and try to explain the nonlinear mechanisms behind the societal emergence.

Process:

Task 3.1: Build a social community simulator
(1) Build a simulator using ChatGPT.
(2) Evaluate the platform based on the previous findings and methodologies.
Task 3.2: Build an influential Bot
(1) Offline train the RL agent on the community simulator.
(2) Evaluate the RL performance by measuring the improvement of ranking it generates in the simulated environment.

Task 3.1: Build a social community simulator

As shown in Figure 5, the process of building simulator involves 1). Personas: Given the initial, limited personas with descriptions, generate a group of personas. 2). Contents: Use the description of each persona to generate tweet content on a specific topic. 3). Interactions: Follow the predefined links, generate new tweet content for each persona.

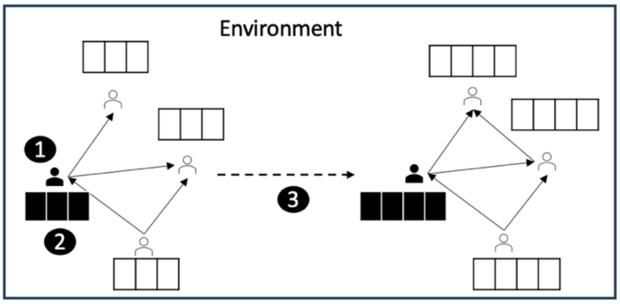


Figure 5: Simulator environment.

Task 3.2: Build an influential Bot

The simulated environment allows us to investigate the process of how to generate an opinion leader in the community. Our next research question is can we use reinforcement learning to find a way to create an opinion leader for this group. As illustrated in figure 7, we state the setting of our reinforcement learning training process. State Space: Opinion and link states of each user; Observation State Space: Opinion states of initial following accounts of agent I; Action Space: [Narrow, Creative]; Update Link Method: Users may follow others who have the same opinion states with themselves in 2 time-steps; Reward: Ranking movement; RL Model: q-learning.

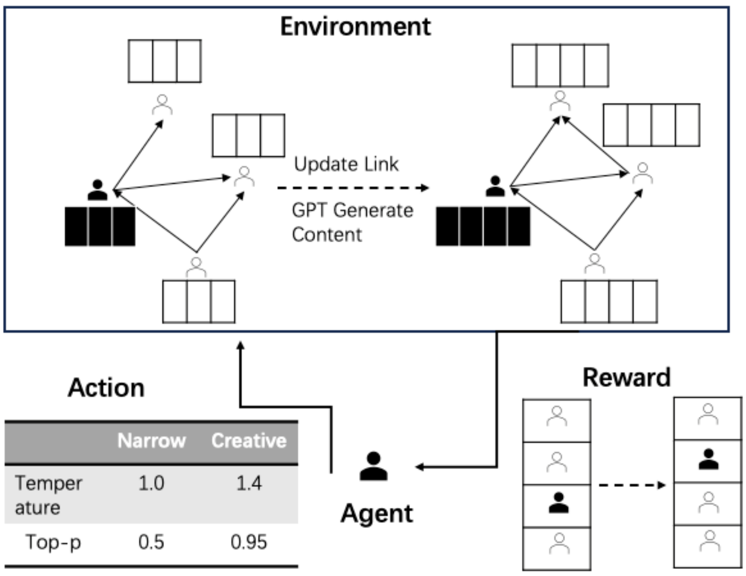
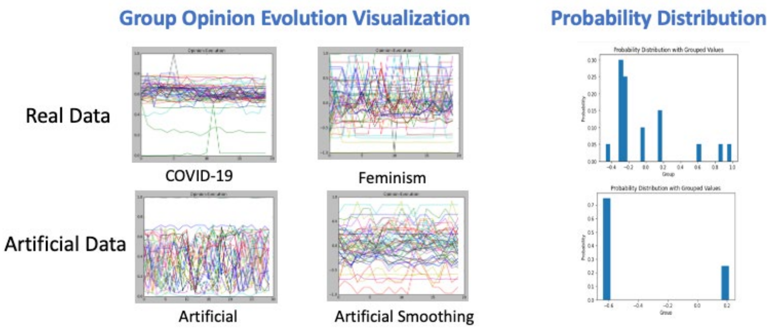


Figure 7: Offline train the RL agent on the community simulator.

Validation Result: As shown in Figure 6, the probability distribution shows that artificial dataset may be too uniform and lacks variation. The regression result shows that the simulated environment also fit on the opinion evolution model (Task 1) with lower R^2 value.



Regression on Opinion Evolution Model

	Covid-19	Feminism	Artificial	Artificial Average
R^2	45%	25%	21%	19%

Figure 6: Visualization, probability distribution and R^2 value of group opinion evolution in real data set and artificial dataset.

Summary

Task 0: Model review and creation of linear networked ODE influence models (2020-21)
Task 1: The influence model validation on the Twitter dataset (2021-22)
Paper accepted by IEEE/ACM Adv Social Network and Data Mining
Task 2: Behaviour Informed Influence Ranking Measures (2022-22)
Paper under review for Journal of Network and Computer Applications
Task 3: Developing an influential Tweet Bot (2023-23)

Note: The completion of Task 3 has been extended by six months due to an extension of the student's PhD studies. Currently, the student is in the process of drafting the accompanying paper. Given the tight deadline, it is not possible to provide the paper at this moment. The paper will be submitted upon completion.