A Review of Social Media Simulation using Multi-agent Models

Editors: List of editors' names

Abstract

Multiagent social network simulations are an avenue that can bridge the communication gap between the public and private platforms in order to develop solutions to a complex array of issues relating to online safety. While there are significant challenges relating to the scale of multiagent simulations, efficient learning from observational and interventional data to accurately model micro and macro-level emergent effects, there are equally promising opportunities not least with the advent of large language models that provide an expressive approximation of user behavior. In this position paper, we review prior art relating to social network simulation, highlighting challenges and opportunities for future work exploring multiagent security using agent-based models of social networks.

Keywords: Social Network, Agent-based Model, Simulation, Recommender System

1. Introduction

Simulators have become integral in various industries Ottogalli et al. (2019) offering virtual environments for training, testing, and experimentation Bousquet et al. (1999) . Their utility spans aviation Sun et al. (2022),Xiong and Wang (2022), Bernard et al. (2022), military Kessels et al. (2021),Sattler et al. (2020) healthcare Kononowicz et al. (2019),Kumar et al. (2020),Croatti et al. (2020) and entertainmentWoolworth (2019). Notably, simulators provide a secure and controlled space for activities, aiding in cost reduction, safety enhancement, and efficiency improvement Green et al. (2016). They contribute to evaluating and optimizing designs before implementation, saving both time and moneyMakransky et al. (2019). Social networks are complex information-sharing systems that have become digital information highways Borgatti et al. (2018),Van Zyl (2009). There have been various attempts to simulate the spread of various types of information on social media using multiagent simulators of user behavior Zhou et al. (2020),Massaguer et al. (2006),Zhao et al. (2015),Prike et al. (2023),Sakas et al. (2019),Dalayli (2020). Multiagent simulators make it possible for external researchers to develop experiments that can be run in the virtual testbed it provides Cardoso and Ferrando (2021).

1.1. Utility of Simulating Social Networks

This paper dives into how simulators are used as tools for understanding information dissemination in social networks, improving our understanding of a variety of complex issues on social networks.

1.2. Modeling Algorithmic Effects

Social media algorithms used to designed to maximize engagement, which tended to amplify human biases towards learning from prestigious, ingroup, emotional, and moral (PRIME) Brady et al. (2023). This can promote misinformation and polarization. Algorithms prioritize engagement over accuracy or truth, leading to the rapid spread of extreme, controversial, or false content. Users often find themselves in "filter bubbles" and "echo chambers," reinforcing their existing views and distorting their perception of group opinions. Users often find themselves in "filter bubbles" and "echo chambers," reinforcing their existing views and distorting their perception of group opinions. The concept of "wisdom of crowds" is compromised by online echo chambers and the presence of fake accounts, bots, and orchestrated networks that manipulate engagement signals Saurwein and Spencer-Smith (2021).

1.3. Modeling Policy Interventions

It is challenging to predict the multifaceted effects of policies on social media. Multiagent simulators offer a virtual testbed for prototyping policy outcomes with the caveat that they don't necessarily reflect all the possible outcomes in perfect accord with the real-world. However, it is useful to be able to model the relative effects of different types of policies prior to their deployent since that permits us to evaluate various policies in the same environment. Simulators can provide a testbed to prototype interventional effects and model how they might influence the system in desirable and undesirable ways.

User Experience and Exposure to Information: The addictive nature Pellegrino et al. (2022) of social media poses challenges to individuals' well-being. Persuasive design techniques contribute to stress, anxiety, and decreased productivity. Filter bubbles and information overload hinder diverse perspectives and contribute to the spread of fake news. Designing healthy user experiences is crucial, emphasizing user control and breaks in user flow.

1.4. Security Testing

Algorithmic Vulnerability Detection: Algorithms used by social media platforms can replicate and amplify human biases, resulting in outcomes that are less favorable to certain groups. Testing strategies include auditing algorithms, diverse design teams, and user feedback. Challenges involve trade-offs between fairness and accuracy, privacy regulations, and the need for algorithmic literacy.

Social Exploitation by Coordinated Networks (CIB): Social bots include, automated accounts imitating human behavior, manipulate public opinion by spreading fake news and divisive content Zhang et al. (2023). Detection and counteracting social bots is challenging due to their sophistication. Orabi et al. (2020) shows that social bots evolve rapidly to evade detection, presenting an "everlasting cat and mouse game" between bot creators and detection methods. Information operations by coordinated networks highlight the global impact of disinformation on social networks.

Polarization: Selective exposure to attitude-confirming information exacerbates confirmation bias and polarizes opinions. Tolerance among users plays a role in mitigating polarization. Information propagation is influenced by the structure of social networks and user activity patterns Haque et al. (2023).

Information Propagation: The structure of social networks impacts how information spreads. User activity patterns follow a power law distribution, with active users playing a significant role in information propagation. Understanding information spread is crucial for better recommendations and identification of manipulation.

2. Existing Applications

Simulators provide a virtual testbed for prototyping various policies as interventions upon the world model underlying a social network. They allow us to encode our understanding of the rules underpinning social interactions, incorporate design affordances from the real world, and create an arbitrarily complex model of information propagation on social networks. Simulators can leverage digital trace data to calibrate their parameters,attempting to provide an accurate representative model of reality. It also allows experimentation in scenarios where there might be ethical challenges deploying randomized control trials. For example, TACIT by Neumann and Wolczynski (2023), enhance fact-checking models and assess their impact on reducing inequalities among online communities. It sparked ethical discussions on prioritizing equity in AI-driven fact-checking.

Algorithmic Auditing: Regulators need detailed evidence on platforms' policies, processes, and outcomes related to misinformation. Algorithm auditing requires a multidisciplinary skill-set and granular data on misinformation spread. Modeling effects of algorithms on social networks can provide insights on balancing societal impacts and technical performance as AI moves from research into real-world applications[1].

Testing User Experience (UX) on Social Media: Simulations Ahlgren et al. (2020) in a social network can assess and predict the type of content users might be exposed to.The open source Misinformation Game simulator Butler et al. (2023) provides a flexible and customizable platform for conducting controlled experiments on factors influencing online misinformation propagation and beliefs. By emulating recommendation systems and user interactions, simulations can reveal issues like filter bubbles, echo chambers, and exposure to harmful content.

Testing Security: Red teaming and blue teaming simulations Seker and Ozbenli (2018)assess and enhance security and privacy measures on social media platforms. These simulations enable proactive identification and mitigation of risks, strengthening the overall security infrastructure.

2.1. The State-of-the-art in Social Network Simulation

Social network simulations have become an integral part of understanding and predicting user behavior on platforms like Facebook, Google, and Twitter. Agent-based simulations with autonomous agents imitate real user actions. Simulators allow safe experimentation with potential changes, ensuring they don't impact real users.

As per Ahlgren et al. (2020), Facebook utilizes a platform called WW (Web-Enabled Simulation) to simulate user interactions and social behaviors within a parallel version of its

1. https://hbr.org/2018/11/why-we-need-to-audit-algorithms

platform. WW employs autonomous software agents or "bots" programmed to imitate real user actions, such as posting, messaging, and making connections. These bots are trained using anonymized logs of real user activity data from Facebook to make their behaviors realistic and statistically match real user statistics. Similarly, Google built RecSim to model the societal effects of recommender systems Ie et al. (2019); Mladenov et al. (2021). RecSim allows configuring agents to represent various types of users, content providers, and other participants in the recommender ecosystem. Twitter also employs a reinforcement learning over agent models in order to optimize user engagement through push notifications O'Brien et al. (2022). As a practical example of the real-world value drawn from simulators, their system manages to successfully maximize long-term user satisfaction by studying user responses to push notifications.

There have been other prior attempts at building expressive forward simulators. HashKat Ryczko et al. (2017) is a dynamic network simulation tool designed to model the growth of information propagation through an online social network. NetSim Stadtfeld et al. (2013) is an R package that allows for the simulation of the co-evolution of social networks and individual attributes. There is a history of common challenges associated with simulating social networks.

2.2. Limitations

The limitations of simulations span a range of challenges. The dependence of simulation outcomes on specific parameter values and internal model structures necessitates sensitivity analysis for a nuanced understanding of variability. Additionally, the inherent complexity of model specifications often limits the transparency of simulation results, hindering a comprehensive grasp of agent trajectories and behaviors. With bespoke simulators for the same task and a lack of transparency into the design of complex simulation procedures, reproducibility becomes challenging due to the absence of standardized procedures and model sharing, impeding knowledge accumulation across studies. The incorporation of social networks introduces complexity, requiring careful calibration with empirical data and risking biases towards replicating existing conditions. Moreover, the abstraction of human behavior in models overlooks crucial psychological and social nuances. Computational power and time constraints further limit the size and complexity of modeled networks. Finally, while achieving macro-level congruence with real-world data is a common goal, it does not guarantee accuracy at the micro-level, posing challenges in validating models against real-world data as per Manzo and Matthews (2014). Additionally, there is ongoing research into LLM-human hybrid models, exemplified by Facebook's Cicero AI (FAIR). The ideas behind integrating LLMs into agent-based modeling extends beyond simulation, towards high-fidelity real-world experimentation introducing new challenges and opportunities at the intersection of artificial intelligence and social science.

3. Conclusion

Multiagent simulators are expressive models of online interaction and have demonstrably yielded value in varied applications. While there are limitations from scale and complexity, there is significant value that is likely to be unlocked by advances in computational modeling and machine learning for this area.

References

John Ahlgren, Maria Eugenia Berezin, Kinga Bojarczuk, Elena Dulskyte, Inna Dvortsova, Johann George, Natalija Gucevska, Mark Harman, Ralf Laemmel, Erik Meijer, et al. Wes: Agent-based user interaction simulation on real infrastructure. In *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops*, pages 276–284, 2020.

Stefano Armenia, Marco Angelini, Fabio Nonino, Giulia Palombi, and Mario Francesco Schlitzer. A dynamic simulation approach to support the evaluation of cyber risks and security investments in smes. *Decision Support Systems*, 147:113580, 2021.

Yanki Aslan, Salman Salman, Jan Puskely, Antoine Roederer, and Alexander Yarovoy. 5g multi-user system simulations in line-of-sight with space-tapered cellular base station phased arrays. In *2019 13th European Conference on Antennas and Propagation (EuCAP)*, pages 1–5. IEEE, 2019.

Steffen Bangsow. *Tecnomatix plant simulation*. Springer, 2020.

Fabien Bernard, Xavier Bonnardel, Raphael Paquin, Martial Petit, Killian Marandel, Nicolas Bordin, and Françoise Bonnardel. Digital simulation tools in aviation maintainability training. *Computer Applications in Engineering Education*, 30(2):384–395, 2022.

Stephen P Borgatti, Martin G Everett, and Jeffrey C Johnson. *Analyzing social networks*. Sage, 2018.

François Bousquet, Olivier Barreteau, Christophe Le Page, Christian Mullon, and Jacques Weber. An environmental modelling approach: the use of multi-agent simulations. *Advances in environmental and ecological modelling*, 113(122), 1999.

William J Brady, Joshua Conrad Jackson, Björn Lindström, and MJ Crockett. Algorithm-mediated social learning in online social networks. *Preprint at OSF preprints. https://doi.org/10.31219/osf. io/yw5ah*, 2023.

Stuart A Bremer. *The GLOBUS model: Computer simulation of worldwide political and economic developments*. Routledge, 2019.

Lucy H Butler, Padraig Lamont, Dean Law Yim Wan, Toby Prike, Mehwish Nasim, Bradley Walker, Nicolas Fay, and Ullrich KH Ecker. The (mis) information game: a social media simulator. *Behavior Research Methods*, pages 1–22, 2023.

Rafael C Cardoso and Angelo Ferrando. A review of agent-based programming for multi-agent systems. *Computers*, 10(2):16, 2021.

Stephen Casper, Jason Lin, Joe Kwon, Gatlen Culp, and Dylan Hadfield-Menell. Explore, establish, exploit: Red teaming language models from scratch. *arXiv preprint arXiv:2306.09442*, 2023.

Pietro Cipresso. Modeling behavior dynamics using computational psychometrics within virtual worlds. *Frontiers in psychology*, 6:1725, 2015.

Paul Covington, Jay Adams, and Emre Sargin. Deep neural networks for youtube recommendations. In *Proceedings of the 10th ACM Conference on Recommender Systems*, New York, NY, USA, 2016.

Angelo Croatti, Matteo Gabellini, Sara Montagna, and Alessandro Ricci. On the integration of agents and digital twins in healthcare. *Journal of Medical Systems*, 44:1–8, 2020.

Jamie Ian Cross, Christine Boag-Hodgson, Tim Ryley, Timothy Mavin, and Leigh Ellen Potter. Using extended reality in flight simulators: a literature review. *IEEE Transactions on Visualization and Computer Graphics*, 2022.

Feyza Ünlü Dalayli. Representation of robots in the social media with the simulation universe: social media influencers and influencer robot miquela sousa. *International Journal of Social Science*, 3(2):87–102, 2020.

Enrico Di Minin, Henrikki Tenkanen, and Tuuli Toivonen. Prospects and challenges for social media data in conservation science. *Frontiers in Environmental Science*, 3, 2015. ISSN 2296-665X. doi: 10.3389/fenvs.2015.00063. URL https://www.frontiersin.org/articles/10.3389/fenvs.2015.00063.

Mohamed El-Sefy, Mohamed Ezzeldin, Wael El-Dakhakhni, Lydell Wiebe, and Shinya Nagasaki. System dynamics simulation of the thermal dynamic processes in nuclear power plants. *Nuclear Engineering and Technology*, 51(6):1540–1553, 2019.

Meta Fundamental AI Research Diplomacy Team (FAIR)†, Anton Bakhtin, Noam Brown, Emily Dinan, Gabriele Farina, Colin Flaherty, Daniel Fried, Andrew Goff, Jonathan Gray, Hengyuan Hu, et al. Human-level play in the game of diplomacy by combining language models with strategic reasoning. *Science*, 378(6624):1067–1074, 2022.

Chen Gao, Yu Zheng, Wenjie Wang, Fuli Feng, Xiangnan He, and Yong Li. Causal inference in recommender systems: A survey and future directions. *arXiv preprint arXiv:2208.12397*, 2022.

Carlos A Gomez-Uribe and Neil Hunt. The netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems (TMIS)*, 6(4):1–19, 2015.

Michael Green, Rayhan Tariq, and Parmis Green. Improving patient safety through simulation training in anesthesiology: Where are we? *Anesthesiology Research and Practice*, 2016:4237523, 2016. doi: 10.1155/2016/4237523.

Amanul Haque, Nirav Ajmeri, and Munindar P Singh. Understanding dynamics of polarization via multiagent social simulation. *AI & society*, pages 1–17, 2023.

Eugene Ie, Chih-wei Hsu, Martin Mladenov, Vihan Jain, Sanmit Narvekar, Jing Wang, Rui Wu, and Craig Boutilier. Recsim: A configurable simulation platform for recommender systems. *arXiv preprint arXiv:1909.04847*, 2019.

Kawaljeet Kaur Kapoor, Kuttimani Tamilmani, Nripendra P Rana, Pushp Patil, Yogesh K Dwivedi, and Sridhar Nerur. Advances in social media research: Past, present and future. *Information Systems Frontiers*, 20:531–558, 2018.

Ilona Kessels, Bart Koopman, Nico Verdonschot, Marco Marra, and Kaj Gijsbertse. The added value of musculoskeletal simulation for the study of physical performance in military tasks. *Sensors*, 21(16):5588, 2021.

Andrzej A Kononowicz, Luke A Woodham, Samuel Edelbring, Natalia Stathakarou, David Davies, Nakul Saxena, Lorainne Tudor Car, Jan Carlstedt-Duke, Josip Car, and Nabil Zary. Virtual patient simulations in health professions education: systematic review and meta-analysis by the digital health education collaboration. *Journal of medical Internet research*, 21(7):e14676, 2019.

Adarsh Kumar, Rajalakshmi Krishnamurthi, Anand Nayyar, Kriti Sharma, Vinay Grover, and Eklas Hossain. A novel smart healthcare design, simulation, and implementation using healthcare 4.0 processes. *IEEE access*, 8:118433–118471, 2020.

Tan Duy Le, Adnan Anwar, Seng W Loke, Razvan Beuran, and Yasuo Tan. Gridattacksim: A cyber attack simulation framework for smart grids. *Electronics*, 9(8):1218, 2020.

Dawen Liang, Laurent Charlin, and David M Blei. Causal inference for recommendation. In *Causation: Foundation to Application, Workshop at UAI. AUAI*, 2016.

Guangdong Liu. Space plasma instrument concept and analysis using simulation and machine learning techniques. 2022.

Yang Liu and Songhua Xu. Detecting rumors through modeling information propagation networks in a social media environment. *IEEE Transactions on computational social systems*, 3(2):46–62, 2016.

Zhuoran Liu, Leqi Zou, Xuan Zou, Caihua Wang, Biao Zhang, Da Tang, Bolin Zhu, Yijie Zhu, Peng Wu, Ke Wang, et al. Monolith: real-time recommendation system with collisionless embedding table. *arXiv preprint arXiv:2209.07663*, 2022.

Guido Makransky, Richard E Mayer, Nicola Veitch, Michelle Hood, Karl Bang Christensen, and Helen Gadegaard. Equivalence of using a desktop virtual reality science simulation at home and in class. *Plos one*, 14(4):e0214944, 2019.

Gianluca Manzo and Toby Matthews. Potentialities and limitations of agent-based simulations. *Revue française de sociologie*, 55(4):653–688, 2014.

Daniel Massaguer, Vidhya Balasubramanian, Sharad Mehrotra, and Nalini Venkatasubramanian. Multi-agent simulation of disaster response. In *ATDM workshop in AAMAS*, volume 2006, 2006.

Martin Mladenov, Chih-Wei Hsu, Vihan Jain, Eugene Ie, Christopher Colby, Nicolas Mayoraz, Hubert Pham, Dustin Tran, Ivan Vendrov, and Craig Boutilier. Recsim ng: Toward principled uncertainty modeling for recommender ecosystems. *arXiv preprint arXiv:2103.08057*, 2021.

Terrence Neumann and Nicholas Wolczynski. Does ai-assisted fact-checking disproportionately benefit majority groups online? In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, pages 480–490, 2023.

Randall K Nichols, Candice M Carter, Jerry V Drew II, Max Farcot, Captain John-Paul Hood, Mark J Jackson, Peter D Johnson, Siny Joseph, Saeed Kahn, Wayne D Lonstein, et al. Space systems modeling and simulation [diebold]. *Cyber-Human Systems, Space Technologies, and Threats*, 2023.

Conor O'Brien, Huasen Wu, Shaodan Zhai, Dalin Guo, Wenzhe Shi, and Jonathan J Hunt. Should i send this notification? optimizing push notifications decision making by modeling the future. *arXiv preprint arXiv:2202.08812*, 2022.

Mariam Orabi, Djedjiga Mouheb, Zaher Al Aghbari, and Ibrahim Kamel. Detection of bots in social media: a systematic review. *Information Processing & Management*, 57(4): 102250, 2020.

Kiara Ottogalli, Daniel Rosquete, Aiert Amundarain, Iker Aguinaga, and Diego Borro. Flexible framework to model industry 4.0 processes for virtual simulators. *Applied Sciences*, 9(23):4983, 2019.

Joon Sung Park, Joseph C O'Brien, Carrie J Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. Generative agents: Interactive simulacra of human behavior. *arXiv preprint arXiv:2304.03442*, 2023.

Alfonso Pellegrino, Alessandro Stasi, and Veera Bhatiasevi. Research trends in social media addiction and problematic social media use: A bibliometric analysis. *Frontiers in Psychiatry*, 13:1017506, 2022.

Ethan Perez, Saffron Huang, Francis Song, Trevor Cai, Roman Ring, John Aslanides, Amelia Glaese, Nat McAleese, and Geoffrey Irving. Red teaming language models with language models, 2022. *URL https://arxiv. org/abs/2202.03286*.

Nadia Politi, Athanasios Sfetsos, Diamando Vlachogiannis, Panagiotis T Nastos, and Stylianos Karozis. A sensitivity study of high-resolution climate simulations for greece. *Climate*, 8(3):44, 2020.

Toby Prike, Lucy Butler, and Ullrich Ecker. Source-credibility information and social norms improve truth discernment and reduce engagement with misinformation online. 2023.

Kristen Lani Rasmussen, Andreas F Prein, Roy M Rasmussen, Kyoko Ikeda, and Changhai Liu. Changes in the convective population and thermodynamic environments in convection-permitting regional climate simulations over the united states. *Climate Dynamics*, 55:383–408, 2020.

Michael A Rosen, Elizabeth A Hunt, Peter J Pronovost, Molly A Federowicz, and Sallie J Weaver. In situ simulation in continuing education for the health care professions: a systematic review. *Journal of Continuing Education in the Health Professions*, 32(4): 243–254, 2012.

Kevin Ryczko, Adam Domurad, Nicholas Buhagiar, and Isaac Tamblyn. Hashkat: large-scale simulations of online social networks. *Social Network Analysis and Mining*, 7:1–13, 2017.

Kalle Saastamoinen and Kasper Maunula. Usefulness of flight simulator as a part of military pilots training–case study: Grob g 115e. *Procedia Computer Science*, 192:1670–1676, 2021.

Damianos P Sakas, Dimitrios K Nasiopoulos, and Panagiotis Reklitis. Modeling and simulation of the strategic use of social media networks in search engines for the business success of high technology companies. In *Strategic Innovative Marketing: 6th IC-SIM, Pafos, Cyprus 2017*, pages 227–236. Springer, 2019.

Lauren A Sattler, Chad Schuety, Mark Nau, Daniel V Foster, John Hunninghake, Tyson Sjulin, and Joshua Boster. Simulation-based medical education improves procedural confidence in core invasive procedures for military internal medicine residents. *Cureus*, 12 (12), 2020.

Florian Saurwein and Charlotte Spencer-Smith. Automated trouble: The role of algorithmic selection in harms on social media platforms. *Media and Communication*, 9(4):222–233, 2021.

Ensar Seker and Hasan Huseyin Ozbenli. The concept of cyber defence exercises (cdx): Planning, execution, evaluation. In *2018 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)*, pages 1–9. IEEE, 2018.

Omar Shaikh, Valentino Chai, Michele J Gelfand, Diyi Yang, and Michael S Bernstein. Rehearsal: Simulating conflict to teach conflict resolution. *arXiv preprint arXiv:2309.12309*, 2023.

Hing Yu So, Phoon Ping Chen, George Kwok Chu Wong, and Tony Tung Ning Chan. Simulation in medical education. *Journal of the Royal College of Physicians of Edinburgh*, 49(1):52–57, 2019.

Christoph Stadtfeld, Maintainer Christoph Stadtfeld, Depends Rcpp, LinkingTo Rcpp, and GNU SystemRequirements. Package 'netsim'. 2013.

Elizaveta Stavinova, Alexander Grigorievskiy, Anna Volodkevich, Petr Chunaev, Klavdiya Bochenina, and Dmitry Bugaychenko. Synthetic data-based simulators for recommender systems: A survey. *arXiv preprint arXiv:2206.11338*, 2022.

Xing Sun, Yuan Yuan, Ting Tan, Tingting Jing, Fei Qin, and Hua Meng. Large eddy simulations of heat transfer and thermal oxidative coking of aviation kerosene in vertical u-tube at a supercritical pressure. *International Journal of Heat and Mass Transfer*, 195: 123205, 2022.

Hélène Trinon et al. Immersive technologies for virtual reality-case study: Flight simulator for pilot training. 2019.

Vineet Vajpayee, Victor Becerra, Nils Bausch, Jiamei Deng, SR Shimjith, and A John Arul. Dynamic modelling, simulation, and control design of a pressurized water-type nuclear power plant. *Nuclear Engineering and Design*, 370:110901, 2020.

Anria Sophia Van Zyl. The impact of social networking 2.0 on organisations. *The Electronic Library*, 27(6):906–918, 2009.

Xiaofeng Wang, Zheng Zhu, Guan Huang, Xinze Chen, and Jiwen Lu. Drivedreamer: Towards real-world-driven world models for autonomous driving. *arXiv preprint arXiv:2309.09777*, 2023.

David S Woolworth. Acoustics simulations to inform the designs of large worship and entertainment spaces to the client and contractor. *The Journal of the Acoustical Society of America*, 145(3):1854–1854, 2019.

Yican Wu. Development and application of virtual nuclear power plant in digital society environment. *International journal of energy research*, 43(4):1521–1533, 2019.

Minglan Xiong and Huawei Wang. Digital twin applications in aviation industry: A review. *The International Journal of Advanced Manufacturing Technology*, 121(9-10):5677–5692, 2022.

Yaming Zhang, Wenjie Song, Yaya H Koura, and Yanyuan Su. Social bots and information propagation in social networks: Simulating cooperative and competitive interaction dynamics. *Systems*, 11(4):210, 2023.

Liang Zhao, Jiangzhuo Chen, Feng Chen, Wei Wang, Chang-Tien Lu, and Naren Ramakrishnan. Simnest: Social media nested epidemic simulation via online semi-supervised deep learning. In *2015 IEEE international conference on data mining*, pages 639–648. IEEE, 2015.

Lixin Zhou, Jie Lin, Yanfeng Li, and Zhenyu Zhang. Innovation diffusion of mobile applications in social networks: A multi-agent system. *Sustainability*, 12(7):2884, 2020.

Appendix A. Appendix

A.1. Simulators Across Industries

Simulators find widespread utility across industries. In aviation, flight simulators, ranging from basic desktop models to complex full-motion systems, are crucial for pilot training Cross et al. (2022),Saastamoinen and Maunula (2021),Trinon et al. (2019). Healthcare relies on simulators for teaching medical workers and replicating complex medical procedures, spanning from fundamental task trainers to advanced patient simulators mimicking human physiology Rosen et al. (2012) ,So et al. (2019). Space exploration benefits from simulations for spacecraft design and mission scenario testing, allowing for development and optimization before actual missionsNichols et al. (2023),Aslan et al. (2019),Liu (2022).

Beyond aviation and healthcare, simulations play pivotal roles in cybersecurityArmenia et al. (2021), Le et al. (2020), economicsBremer (2019), climate science Politi et al. (2020),Rasmussen et al. (2020) and nuclear power plant design Vajpayee et al. (2020),El-Sefy et al. (2019),Wu (2019) . These virtual environments assist in modeling, testing safety measures, and simulating emergency situations. Illustrative examples include the Wang et al. (2023) - a diffusion based simulator driven by real world autonomous driving data, CERN's computational psychometricsCipresso (2015) for real-world and virtual behavior integration, and Siemens' use of TECNOMATIX (Bangsow, 2020) for optimizing production systems and logistics processes. These applications showcase the diverse and impactful uses of simulators across various domains, demonstrating their potential in understanding and modeling complex scenarios.

Appendix B. Open Challenges

While multi-agent simulators are a promising approach for studying information for studying information propagation on social networks, current techniques have significant limitations. Even the most advanced simulators have significant gaps in their capabilities when compared to the complexity of real-world social platforms.

B.1. Inference at Scale

Inferring accurate simulation parameters from real-world social network data is extremely challenging, especially at the massive scales of modern platforms. Each social media site has a unique architecture and focuses on different types of user interactions and content. For example, Twitter emphasizes short messages, broadcasting and news, while Instagram centers on the visual photo and video sharing. The core algorithms driving each platform are opaque and keep changing.

Computational social science techniques aim to improve large-scale inference by combining machine learning with insights from disciplines such as sociology, psychology, and communications theory. For example, research into cognitive biases that influence how users spread misinformation. Even with advances in big data and artificial intelligence, capturing every aspect of human behaviour remains difficult.

B.2. Data Limitations

While simulations rely on real-world data for inference and evaluation, comprehensive social media data is increasingly difficult for researchers to access. Platforms like Facebook and Twitter have become more restrictive in sharing data, due to concerns around privacy, ethics and potential misconduct. For instance, in an interview by Undark.[2], Meta representatives said that common researcher practices like web scraping or third-party APIs can now lead to accounts being blocked or banned if done without permissions. Even when data is granted, it is often limited in scope or heavily sampled across several channels. This makes collecting large, unbiased datasets to train accurate simulations acutely challenging. The study conducted by Liu and Xu (2016), mentioned how they used web crawling for their study but it was difficult to adapt it to simulations for large-scale data. In summary,

2. See https://undark.org/2022/04/18/why-researchers-want-broader-access-to-social-media-data/

expanding platform restrictions on data access increasingly hinder the simulation research needed to understand and improve social media. Mechanisms to enable responsible data sharing with researchers are needed. As per Di Minin et al. (2015), Social media provides a wealth of real user conversation data that could help in conversational science research. But access is restricted so responsible data sharing under ethics would enable leveraging these conversations to advance conversational agents. With proper safeguards, social media data presents opportunities to develop dialogue systems grounded in natural human exchanges. Finding the balance between privacy protection and research access remains challenging but important for progress in conversational AI. Kapoor et al. (2018) also highlights the value of social media data, showing how user-generated content on these platforms provides a rich source of natural conversations and social interactions that could inform research across diverse fields, including information systems.

B.3. Algorithmic Auditing

B.3.1. Modeling Algorithmic Effects

Algorithms play a crucial role when it comes to social media recommendation systems in delivering appropriate content to users. However, the intricacy of these algorithms can be challenging to model in simulations. We can take a look at social media platforms like Twitter, TikTok, Netflix and YouTube as general examples of a complex recommendation system. TikTok's recommendation system, Monolith, is a real-time recommendation system by Liu et al. (2022) that incorporates data structures such as collision-less embedding tables with distributed architectures for training and serving. The concurrent data flows, timing-sensitive operations, failure handling, and sheer data volumes presented in the paper are non-trivial to model.

YouTube represents one of the largest scale and most sophisticated recommendation systems in existence as shown in Covington et al. (2016). The recommendation system at Twitter[3] is composed of many interconnected services and jobs, which aims to distill roughly 500 million tweets posted daily down to a small number of top tweets that ultimately show up on a user's "For You" section. On the other hand, the recommender system at Netflix, described in Gomez-Uribe and Hunt (2015), is not just one algorithm but rather a variety of algorithms that collectively define the Netflix experience.

The paper by Gao et al. (2022) is a survey of the literature on causal inference-based recommendation, that aims to enhance recommender systems by utilizing causal inference to extract causality from data. To simulate algorithms used in this paper, we would require modeling complex causal relationships, accounting for potential confounders and biases, and balancing the trade-offs between the accuracy and fairness of the recommendations.

Simulating the various algorithms used by this Liang et al. (2016) paper recommendation system requires capturing complex behaviors and dynamics, such as the user discovery process, the user preference function, the causal effects of the recommendations, and the feedback loop between the users and the items.

3. https://blog.twitter.com/engineering/en_us/topics/open-source/2023/twitter-recommendation-algorithm

Overall, simulating the various algorithms used by social media recommender systems requires capturing intricate behaviors and dynamics, which is a complex task that will require considerable engineering effort.

B.3.2. Benchmarking Recommender Systems

Simulators can be used to model how information spreads on social networks and to examine the effects of recommender systems on the virtual sharing ecosystem when it comes to benchmarking recommendation systems. Stavinova et al. (2022) demonstrates one of the utilities of simulators which is to test the limits of existing recommender systems of different types (including Reinforcement Learning ones) and to complex user preference formation.

There are several available online recommendation systems that can be used as examples to examine the effects of recommender systems on the virtual information-sharing ecosystem. One example is this repository.[4], which contains different sophisticated recommendation systems available for simulation environments. The repository provides baselines and reproducible code for standard recommendation techniques, and the datasets available in the repository could serve as a starting point to generate simulated user-item interactions and feedback. Another example is RecSim Ie et al. (2019), a configurable platform for authoring simulation environments to facilitate the study of RL algorithms in recommender systems.

B.4. Security

Simulators can help analyse vulnerabilities in social media algorithms. In recent times, social media platforms heavily rely on recommendation algorithms to curate content for users. These algorithms tend to be vulnerable and can be exploited for malicious purposes such as spreading misinformation or harmful content. Other than that responses and recommendations generated by LLMs, and recommenders can also prove to be unreliable in several cases. In a comprehensive study by Casper et al. (2023), researchers applied Red teaming methods to study the GPT-2 and GPT-3 models and discovered prompts and responses that were toxic when compared to real-world knowledge. Red Teaming, a process of testing the effectiveness of algorithms by simulating attacks on them for multi-agent simulators can help identify vulnerabilities and improve the security of the algorithms. Casper et al. (2023) explains a three-step red teaming approach which they utilised for enabling simulators to refine and extend the study of data propagation on social media. Steps involved identifying the model's behaviour in the desired context, then establishing a measurement of undesired behaviour and finally exploiting or attacking the model's flaws using pre-established red teaming methodology. Another study by Perez et al. focused on evaluating potential harms from language models by using another language model to generate adversarial test cases. The study discusses challenges in algorithms being gamed and exploited and proposes red teaming as a way to discover potential issues.

4. See https://github.com/recommenders-team/recommenders

B.5. Large Language Model based Agents and Generative Social Science

The recent advancements in natural language processing and machine learning viz. LLMs has given new life to the idea of 'agent'-based models and spurred interest in generative social simulations. Large Language Models (LLMs) are increasingly employed in generative social science, particularly in agent-based modeling. [5] For example, the approach by Shaikh et al. (2023) aims to explain macro-level social phenomena by simulating interactions among individual agents following simple rules. The goal is to grow macro structures from micro foundations, emphasizing dynamic processes and the sufficiency of micro-level rules. To enhance realism, models are embedded with empirical data about agents and environments. Agent-based modeling is instrumental in studying diverse social phenomena, from disease spread to urban development, by unraveling complex systems' emergence from individual agent interactions. Integrating LLMs into this framework, as seen in projects, like Stanford's generative agents by Park et al. (2023), aims to introduce human-like behaviors and natural language communication in simulated environments, modeling on information diffusion, relationship formation, and coordination in these societies.

Appendix C. Conclusion

With the advent of regulation in the sphere of social media and digital services such as the Digital Services Act, the Online Safety Bill, and other legislation, global lawmakers are hoping to strike a balance and deliver effective policy mechanisms to improve the online experience for users of social platforms. This is challenging in the absence of access to platform data, knowledge of business considerations, and a lack of clarity into what is called 'impossible tradeoffs' in the field of trust and safety–which often balances availability of resources against provision of additional mechanisms to ensure user safety. Simulators offer a high-fidelity solution that trades off complexity from real-world algorithms with the intuitive understanding they offer to non-domain experts, about the workings of a complex system such as social networks bridging the public-private information gaps to effectively address online safety issues.

5. There is frequently an overloading of the term 'agents' where it may reference an independently operating LLM; in this context an LLM might represent a user's activity independently in a complex system.