Springer Proceedings in Complexity

Zining Yang
Santiago Núñez-Corrales *Editors*

Proceedings of the 2022 Conference of The Computational Social Science Society of the Americas

Springer

Springer Proceedings in Complexity

Springer Proceedings in Complexity publishes proceedings from scholarly meetings on all topics relating to the interdisciplinary studies of complex systems science. Springer welcomes book ideas from authors. The series is indexed in Scopus.

Proposals must include the following:

- name, place and date of the scientific meeting
- a link to the committees (local organization, international advisors etc.)
- scientific description of the meeting
- list of invited/plenary speakers
- an estimate of the planned proceedings book parameters (number of pages/articles, requested number of bulk copies, submission deadline)

Submit your proposals to: Hisako.Niko@springer.com

Zining Yang · Santiago Núñez-Corrales
Editors

Proceedings of the 2022 Conference of The Computational Social Science Society of the Americas

Springer

Editors
Zining Yang
Strategy, Innovation and Advanced Analytics
Southern California Edison
Rosemead, CA, USA

Santiago Núñez-Corrales
National Center for Supercomputing Applications
University of Illinois Urbana-Champaign
Urbana, IL, USA

ISSN 2213-8684 ISSN 2213-8692 (electronic)
Springer Proceedings in Complexity
ISBN 978-3-031-37552-1 ISBN 978-3-031-37553-8 (eBook)
https://doi.org/10.1007/978-3-031-37553-8

Foreword

In October 2022, the Computational Social Science Society of the Americas convened in Santa Fe, NM, for its first in-person conference since the start of the pandemic. As in years past, the aim of the conference was twofold. First, it was to provide a venue to present and discuss cutting-edge work in computational social science. We view computational social science broadly, i.e., it is the *theory-driven, nontrivial use of computational methods to analyze, model, simulation, and explore social systems.* In this sense, we explicitly expand beyond the simple analysis of large data gathered from a social system and explore how these signals correspond to social science theory and often explore the viability of hypothesized generating mechanisms underlying these social dynamics via agent-based simulation. Secondly, it was to create an open dialog among computational social scientists to discuss the current state and future direction of the field. The first, as you will find in the following pages, was as dynamic, broad, and exciting as ever. In the discussions that took place, the second came to light. While not represented in the papers contained in these proceedings, it seems the consensus among the society is that the current state of the field is strong with great future potential. Having said that, many felt the field could have a broader impact—a topic we plan to explicitly address at the 2023 Conference.

It was an exciting event and great to see so many regulars and new members arrive at the Drury Hotel ready to engage with one another face-to-face after such a long break. In addition to our society members, we were joined by two fantastic keynote speakers. During our evening banquet Dr. Patrick Grim discussed the wisdom of crowds and the role of diversity and expertise in problem solving, diving into the details of the concepts popularized by Dr. Page over the past few years. Over lunch Dr. Stephanie Forrest discussed the biology of computation, highlighting many features of biologic systems that shed light on the problem-solving capabilities of human systems. If you missed these talks, check the society's YouTube.com channel for the recordings.

As usual, the breadth of the conference was impressive, spanning from land use to dirty bomb evacuation and from reconceptualizing flocking algorithms to the belief dynamics of a classroom. The discussions at receptions and meals were equally

energetic and wide-ranging. The papers contained in these proceedings demonstrate the passion and extent of expertise of our society's members, as well as that of the field more generally. Among the papers, you will find novel analytic methods to describe and characterize social systems from a numeric perspective and from a linguistic perspective. You will also find many examples of sophisticated agent-based simulations used to explore potential generating mechanisms exploited by social systems to solve problems or mitigate external perturbations.

It is our hope that you will find these papers as energizing to read as we found them to discuss. If so, you are invited to join the society and become more involved with this vibrant community of academic, governmental, and private sector researchers focused on the nontrivial use of computation to better understand social systems and address the hard problems facing our planet and our species.

Matthew Koehler
The Computational Social Science
Society of the Americas
Santa Fe, NM, USA

Contents

A Method to Differentiate 'Fringe' and 'Mainstream' Beliefs, and Its Application to Narratives on Russia, Ukraine, and Putin's 2022 War 1
Peter A. Chew, Matthew H. Fort, and Jonathan A. G. Chew

Evaluating a Crowd Logistics Network Using Agent-Based Modeling 21
Preetam Kulkarni and Caroline Krejci

Introducing Land Constraints to Macroeconomic Agent-Based Models 35
Jacob Kelter, Uri Wilensky, and Joseph Potvin

Interactions, Model Mechanisms and Behavioral Attractors in Complex Social Systems 49
H Van Dyke Parunak and Santiago Núñez-Corrales

Modeling Farmers' Adoption Potential to New Bioenergy Crops: An Agent-Based Approach 63
Kazi Ullah and Andrew Crooks

Investigating Emergency Responders' Roles in a Dirty Bomb Event with an Agent-Based Model 77
Ellie Q. Chen and William G. Kennedy

Replacing Diamond-Dybvig 93
John S. Schuler

Higher-Order Interactions in ABM: A Case Study Using Topologically-Perturbed Voter Models 99
Santiago Núñez-Corrales, Rajesh Venkatachalapathy, Jeffrey Graham, and Srikanth Mudigonda

Modeling Macaque Fighting Dynamics with the Evolutionary Model Discovery Framework to Understand Its Application and Utility 117
Alex Isherwood, Melanie Jutras, Matthew Koehler, David Slater, William Thompson, and Maria Yelenick

Flocking with Only Two Parameters 129
Dashiell Bhattacharyya and William G. Kennedy

Understanding Genocide Through Emotion Detection in Historic Documents 145
Elizabeth M. von Briesen, Michael Garvin, and Samira Shaikh

An Agent-Based Model to Explore Belief and Behavioural Change in a Classroom 155
Keegan Fernandes, Daniel Davison, and David Wang

The Role of Communication and Network Technologies in the Dynamics of Social Movements 175
Krystyna Marcinek, Rushil Zutshi, Omair Khan, Justin Grana, Marek Posard, Todd Helmus, and Aaron Frank

Foraging Games: Ideal and Not 205
Robin Clark and Steven O. Kimbrough

Entropy-Based Heuristic Approach For The Quantum-Like Generalization of Social Contagion 221
Ece Çiğdem Mutlu and Ozlem Ozmen Garibay

Editors and Contributors

About the Editors

Dr. Zining Yang is a Senior Manager at Southern California Edison. She also works as an Adjunct Professor at California State Polytechnic University, Pomona, and Associate Director at the TransResearch Consortium. She sits on the Board of the Computational Social Science Society of the Americas (CSSSA) and serves as the Scientific Advisory Board Member for Human Factors and Simulations. She received her Ph.D. in Computational and Applied Mathematics and Political Economy from Claremont Graduate University. Her research interests include data analytics, machine learning, modeling and simulation, complex adaptive systems, agent-based models, and network analysis. She has published numerous times in the fields of computer science, economics, public policy, and political science. She has been identified as Outstanding Researcher by the government, worked on a National Science Foundation-sponsored project, and won multiple awards from various organizations, including the Ministry of Education of the People's Republic of China, International Social Computing, Behavioral Modeling and Prediction, and the International Institute of Informatics and Systemics.

Dr. Santiago Núñez-Corrales is a Research Scientist at the National Center for Supercomputing Applications, a Faculty Affiliate at the Center for Global Studies, and a Lecturer at the School of Information Sciences, University of Illinois Urbana-Champaign. He obtained his Doctorate in Informatics University of Illinois Urbana-Champaign with a minor in Global Studies in 2020. He has received multiple awards, including the 2017 SIGHPC ACM/Intel Computational and Data Science Fellowship. His research focuses on developing mathematical and computational methods to unveil statistical physics aspects governing complex multiscale stochastic systems, or strongly coupled, heterogeneous open systems whose description requires events at multiple scales, and whose action is dictated by irreversible thermodynamics. Prior to this, he served as Director General of Research and Technology Development at the Ministry of Science, Technology and Telecommunications in Costa Rica,

where he directed the development of the National Plan for Science, Technology and Innovation 2015–2021 across multiple economic sectors and other science and technology foresight exercises.

Contributors

Dashiell Bhattacharyya George Mason University, Fairfax, VA, USA

Ellie Q. Chen Thomas Jefferson High School for Science and Technology, Alexandria, VA, USA

Jonathan A. G. Chew Galisteo Consulting Group, Inc., Albuquerque, NM, USA

Peter A. Chew Galisteo Consulting Group, Inc., Albuquerque, NM, USA

Robin Clark Department of Linguistics, University of Pennsylvania, Philadelphia, PA, USA

Andrew Crooks University at Buffalo, Buffalo, NY, USA

Daniel Davison University of Waterloo, Waterloo, ON, Canada

Keegan Fernandes University of Waterloo, Waterloo, ON, Canada

Matthew H. Fort Galisteo Consulting Group, Inc., Albuquerque, NM, USA

Aaron Frank RAND Corporation, Arlington, VA, USA

Ozlem Ozmen Garibay Department of Industrial Engineering and Management Systems, University of Central Florida, Florida, FL, USA

Michael Garvin UNC Charlotte, Charlotte, NC, USA

Jeffrey Graham Department of Mathematical Sciences, Susquehanna University, Selinsgrove, PA, USA

Justin Grana RAND Corporation, Arlington, VA, USA

Todd Helmus RAND Corporation, Arlington, VA, USA;
RAND Corporation, Santa Monica, CA, USA

Alex Isherwood The MITRE Corporation, McLean, VA, USA

Melanie Jutras The MITRE Corporation, McLean, VA, USA

Jacob Kelter Northwestern University, Evanston, IL, USA

William G. Kennedy Department of Computational and Data Sciences, George Mason University, Fairfax, VA, USA

Omair Khan RAND Corporation, Santa Monica, CA, USA

Steven O. Kimbrough Department of Operations, Information and Decisions, University of Pennsylvania, Philadelphia, PA, USA

Matthew Koehler The MITRE Corporation, McLean, VA, USA

Caroline Krejci University of Texas at Arlington, Arlington, TX, USA

Preetam Kulkarni University of Texas at Arlington, Arlington, TX, USA

Krystyna Marcinek RAND Corporation, Arlington, VA, USA;
RAND Corporation, Santa Monica, CA, USA

Srikanth Mudigonda School for Professional Studies, Saint Louis University, St. Louis, MO, USA

Ece Çiğdem Mutlu Department of Industrial Engineering and Management Systems, University of Central Florida, Florida, FL, USA

Santiago Núñez-Corrales National Center for Supercomputing Applications, University of Illinois, Champaign, IL, USA ;
NCSA, University of Illinois Urbana-Champaign, Urbana, IL, USA

H Van Dyke Parunak Parallax Advanced Research, Beavercreek, OH, USA

Marek Posard RAND Corporation, Arlington, VA, USA

Joseph Potvin Xalgorithms Foundation, Ottaw, ON, Canada

John S. Schuler Kukun Home Investment Intelligence, Menlo Park, CA, USA

Samira Shaikh UNC Charlotte, Charlotte, NC, USA

David Slater The MITRE Corporation, McLean, VA, USA

William Thompson The MITRE Corporation, McLean, VA, USA;
The Vermont Complex Systems Center at the University of Vermont, Burlington, VT, USA

Kazi Ullah University at Buffalo, Buffalo, NY, USA;
Oak Ridge National Laboratory, Oak Ridge, TN, USA

Rajesh Venkatachalapathy Systems Science Graduate Program, Portland State University, Portland, OR, USA

Elizabeth M. von Briesen Elon University, Elon, NC, USA

David Wang University of Waterloo, Waterloo, ON, Canada

Uri Wilensky Northwestern University, Evanston, IL, USA

Maria Yelenick The MITRE Corporation, McLean, VA, USA

Rushil Zutshi RAND Corporation, Santa Monica, CA, USA

A Method to Differentiate 'Fringe' and 'Mainstream' Beliefs, and Its Application to Narratives on Russia, Ukraine, and Putin's 2022 War

Peter A. Chew, Matthew H. Fort, and Jonathan A. G. Chew

Abstract We introduce and describe an unsupervised text-analytic method which we hypothesize could help an analyst seeking to make sense of the information landscape to be able to pinpoint, from a high level, the sources of fringe narratives and beliefs. The method is premised on the idea set forth in [1] that cults can often be recognized 'in the wild' through their distinctive vocabulary. Building on concepts from unsupervised learning, we generalize this idea to posit that even if we did not know a priori which particular 'distinctive vocabulary' identified an as-yet unknown 'fringe belief', the texts which are representative of 'fringe' beliefs should, in general, stand out from the 'mainstream' of text when anomaly detection techniques are applied. We test our hypothesis first by constructing an artificial dataset in which we hand-select 27 snippets of text representative of a set of 'fringe' beliefs about Russia and approximately the length of Twitter posts, add approximately 20,000 Twitter posts gathered using 'neutral' words relevant to the topic of the 27 texts, and then applying our technique while withholding from it which source each text came from. We find that the technique can indeed direct the attention of the analyst to the 27 texts, the 'needle in the haystack'. We then test the hypothesis again by applying the technique to 2,838 articles from Russian-language media sources in 2022, including three based in Moscow, and one which relocated to Latvia in 2014 to escape Kremlin control and censorship. We hypothesize that in the Russian news media landscape, the independent media organization should appear as 'fringe'. Again, our hypothesis is confirmed. There were also some surprises in the results, which we discuss—along with how related techniques can also pinpoint what about the 'fringe' in each case differentiates it from the mainstream.

P. A. Chew (✉) · M. H. Fort · J. A. G. Chew
Galisteo Consulting Group, Inc., 4004 Carlisle Blvd NE Suite H, Albuquerque, NM 87107-4566, USA
e-mail: pachew@galisteoconsulting.com

J. A. G. Chew
e-mail: jonathan.chew@galisteoconsulting.com

Z. Yang and S. Núñez-Corrales (eds.), *Proceedings of the 2022 Conference of The Computational Social Science Society of the Americas*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-031-37553-8_1

1 Background

'First they ignore you, then they laugh at you, then they fight you, then you win.' This quote, attributed (but maybe misattributed) to Mahatma Gandhi, neatly demonstrates the facets of a 'fringe' belief or theory. Copernicus's heliocentric model of the Solar System was, at one point, a fringe belief in the sense that it was not held to in the mainstream. Indeed, Galileo was subjected to an Inquisition by the Catholic Church which maintained that heliocentrism was heretical since it allegedly contradicted the Bible. Today, Copernicanism has become uncontroversial; indeed, mainstream (including Catholic and other mainstream religious) thought has no trouble simultaneously holding in mind the concept of a Solar System in which the Sun is at the center, but also a galaxy in which the Sun is not at the center.

Though it turned out that Copernicus was onto something with his 'fringe beliefs', it would be just as much of a mistake to think that those on the fringe are inevitably those who are on the right side of history, as it would be to think that the fringe is, by definition, 'lunatic'. There is also such a thing as deception which can take hold on the fringe just as easily as it can in the mainstream. The insurrectionists who stormed the U.S. Capitol on January 6, 2021 were motivated by a set of beliefs, including false beliefs, which—though shared by the President of the United States—were not, we believe it is safe to say, representative of a majority of Americans.

We must also note that what is fringe relative to one group may be mainstream relative to another. Galileo's contretemps with the Catholic Church of course had a religious dimension, and there were those among the January 6 insurrectionists who also, like the seventeenth-century Catholic mainstream, assumed a religious mantle. Suppose, for the sake of argument, that among some particular Bible-believing religious group in America, that 80% supported the fringe ideas motivating January 6, but 20% did not; if that were true, then what is on the 'fringe' nationally would be 'mainstream' within that group, but what would be fringe for the group might then align with the national mainstream. To determine 'fringe', one must have some kind of 'mainstream' to act as a point of reference.

The point of all this is that whether an idea is 'fringe' now says nothing about whether it is 'fringe' for all time, and it does not say anything, necessarily, about whether that idea is true or false, good or bad, safe or dangerous to society. Ensuring we set the terms of discussion narrowly like this helps set the stage for a productive and highly general application of unsupervised learning to distinguishing the fringe from the mainstream. But a deliberately narrow statement of our terms of reference does more than that. It enables us to be precise about what problem we are setting out to solve, in terms that will make sense to an intelligence analyst seeking to make sense of the information environment. We are, in effect, focusing on a problem of helping the analyst make sense of the forest and not get distracted by the trees. An analyst will want to know broadly what in a dataset (for example of social media posts) is representative of a mainstream set of beliefs, and what is representative of the fringe, irrespective of which side of that divide 'hostile' and 'friendly' fall. The contribution of this paper, then, is to demonstrate a method to let the data itself

(and we focus on natural language data) tell us the answer to this, without knowing anything a priori about the data.

2 Prior Literature

Our work here was against the background of our general immersion in unsupervised data analytics as applied to natural language. It was specifically motivated, however, by reading 'Cultish: the Language of Fanaticism' [1], which is not a work from the field of data analytics at all, but is more of a general-interest discussion by a linguist, Amanda Montell, about how cults[1] are often recognizable by the distinctive vocabulary they use—often quite deliberately.

Through examples of cults ranging from multilevel marketing (MLM) to fitness to Scientology to mainstream American evangelical Christianity, Montell documents how cults are consistently associated with specific and distinctive vocabulary—for example, '#BossBabe' (MLM), 'change your body' (fitness), 'enturbulated' (Scientology), 'on my heart' (evangelicalism). Montell points out that the use of such language is deliberate in that it helps cult members delineate between the in-group and the out-group (ibid, pp. 42–43) and sometimes is even enshrined, as in Scientology (ibid, p. 139), as a prerequisite to being considered a bona fide member of the cult. In common with us, Montell does not assume that cults need take on sinister dimensions, but a recurring thread throughout Montell's discussion, nevertheless, is the danger that these distinctive vocabularies can be used to manipulate cult members, often for pecuniary gain.

We focus in this paper just on Montell's notion that each cult is recognizable 'in the wild' by its distinctive vocabulary (and that distinctiveness would be assessed relative to the background noise of how words are used by the mainstream of society). We set aside as out of our own scope the potential manipulative use of language in cults (though assuming we could come up with a way of identifying cults through their use of language, manipulativeness could be something an intelligence analyst would then be in a better position to assess). How could unsupervised machine learning help with this? And it is one thing to identify a fringe belief we already know about (like Scientology), but there is a harder related problem: supposing a new cult were emerging that we were unaware of. Could machine learning identify the 'unknown unknown' in the information landscape, through unknown but still distinctive *new* combinations of words? For these, we can turn to existing work in the field of unsupervised machine learning.

In the IEEE VAST Challenge 2009 [4], the following problem was posed to participants:

[1] We deliberately leave quite open the question of how to define 'cult', which for us is less central an issue than it may be for Montell. Montell's book simply provides a springboard for our work, which centers in the end around 'fringe' versus 'mainstream'. This allows generalization of our research to finding *any* sociocultural entity or group that stands out from the crowd.

> There are problems at the U.S. Embassy in Flovania... The network security team recently found irregularities while reviewing Embassy network traffic logs.
>
> Finding no security issues that could explain the anomalies, they notified the Embassy Counterintelligence Officer. Upon further investigation, the CI officer identified certain espionage. You have been requested to help him with the investigation...
>
> A month's worth of network traffic logs is available. Each employee has been assigned a desktop computer with a static IP address for use in their daily duties. The network traffic log data consists of the computer IP address, the employee number of the assigned user, outgoing and incoming activity from the computer including destination site, payload (request and response data) and port number.

A neat machine learning solution to this problem [2] assumed that the overwhelming majority of unproblematic network traffic—for which the entire dataset, including anomalies, can be thought of as an approximation—represented background noise, and that against this backdrop, the instances of data exfiltration would then be identifiable as anomalies in the overall data landscape. Using unsupervised analysis, Robinson [ibid] identified 19 network events out of the total of 115,414 that appeared anomalous, and this was the result he submitted to the Data Challenge. Of these nineteen, it turned out that eighteen were the answer that IEEE had predetermined to be the correct answer (representing the counterintelligence threat). There was an interesting twist to the VAST data challenge story, though, that brings to mind the observation attributed to Gandhi, and suggests that even the world of data science today has its 'mainstream' and 'fringe'. Robinson shared with us in personal communication that one reviewer had stated that he did not understand Robinson's math, while another, incredulous that the 18 anomalous events could have been detected so easily, insinuated that Robinson's solution must have involved cheating. We, however, are confident that it was not so.

In detail, Robinson's approach [2] is to treat the network events as if they were 'documents' and the characteristics of those events (IP addresses, etc.) as if they were words in the documents—in effect, to treat the network traffic data like natural language data. A Latent Dirichlet Allocation is then computed on the 'term-by-document' matrix, and pairwise distances between each pair of network events computed. The landscape can then be analyzed via heatmaps and/or dendrograms, and it was via this method that Robinson found 'the needle in the haystack': the 19 network events were easily identifiable as the anomalous cluster in the overall landscape.

3 Application of Unsupervised Analysis to Identifying Fringe Beliefs

Building upon Montell's insights [1], we can make one or two tentative assumptions that allow us then to frame, and solve, the problem of differentiating between fringe and mainstream beliefs in very similar terms to those of Robinson [2], whose goal is to differentiate between anomalous and background network traffic. Here, our

approach differs from most previous research on comparative text analysis, where set dictionaries [13] or extrinsic information [14, 15] are often required to identify outliers. Our tentative assumptions are:

(1) that the language used in text such as newspaper articles, Twitter posts, etc., at some level reflects the beliefs of the writer. This is a very loose assumption and it should be emphasized that we expect the machine learning methods we employ to take in stride any dual meanings, codewords that may be encountered;
(2) that beliefs significantly outside the mainstream would then be reflected as anomalous within the data, without our even having to know in advance which specific words go with which beliefs (e.g. 'enturbulated'—Scientology)—just as Robinson did not have to know in advance, for example, which specific IP addresses were associated with data exfiltration. To the extent assumption (1) is invalid, we should then find that out if the results diverge from those we might expect.

Note that we need not assume in any of this that the fringe represents a threat, as it did in the data exfiltration instance. All we need assume is that unsupervised analysis is capable of identifying the fringe *because of its differentness as reflected in the data.*

Our approach proceeds similarly to [2], although (out of personal preference) we choose to use Singular Value Decomposition (SVD) [12] instead of Latent Dirichlet Allocation (LDA), and several other small changes flow from that decision. From a collection of texts, we perform the following steps:

1. Calculate how many occurrences of each word are in each text and populate the cells of a term-by-text matrix with those values. Each cell (i, j) in this matrix will then represent the $f(i, j)$, the frequency of word i in text j.
2. Weight the values in the matrix using Pointwise Mutual Information weighting, in which:

$$pmi(i, j) = \log \frac{p(i, j)}{p(i)p(j)} = \log \frac{p(i|j)}{p(i)}$$

3. L2-normalize the values in the matrix column-wise (document-wise), such that the sum of the squared values in each column is 1.0. The effect of this step is to balance the information contribution of each text to the overall signal.
4. Compute the SVD of the matrix output from step 3. The SVD factorization is as follows:

$$X = U \cdot S \cdot V^H = \sum_{i=1}^{r} s_i \cdot u_i \cdot v_i^H$$

5. Calculate pairwise similarities between each pair of texts, based on the output of SVD. (Whereas Hellinger distance is used in [2], which is appropriate for LDA given that the output of LDA represents probability distributions, we use cosine here.)

6. Plot heatmaps and/or dendrograms to identify clusters of outliers among the texts.
7. Use additional analytic techniques to characterize differences between the mainstream and the fringe clusters—in terms of the vocabulary they use.

All the above can be implemented relatively straightforwardly in a Jupyter Python notebook [3].

We then test the hypothesis that this approach can indeed differentiate usefully between 'fringe' and 'mainstream' in two ways, each with a different dataset. First, we artificially introduce a set of a few 'fringe belief' texts into a 'mainstream' of general Twitter traffic, and assess the usefulness of the approach in differentiating the two. We then hypothesize that independent Russian media should be represented as a 'fringe' in the overall Russian-language media landscape dominated by Kremlin narratives and censorship. To test this hypothesis, we assess the extent to which the output of the approach above indeed suggests that independent Russian media are fringe. We also use the unsupervised framework flesh out what 'fringe' means in this case, drawing upon Montell's ideas [1], to identify specific differences in vocabulary between the Russian fringe and mainstream.

4 Application (1): 'Gog and Magog'

4.1 What is 'Gog and Magog'?

> **Obi-Wan Kenobi**: With all due respect, Master, is he not the Chosen One? Is he not to destroy the Sith and bring balance to the Force?
>
> ***Mace Windu***: *So the prophecy says.*
>
> **Yoda**: A prophecy that misread could have been.

It was 1993, and the author (Chew), who had grown up, lived and studied in the United Kingdom, and never spent any significant time in the United States, spent a summer volunteering at an evangelical summer camp in upstate New York. While there, I spent my spare time leafing through the books at the camp bookstore. I had recently completed an undergraduate degree in Russian language and literature, with a year living in the USSR, and one book caught my attention: 'The Beginning of the End' [5]. In this book, the author LaHaye made a claim that riveted me: that chapters 38 and 39 in the Biblical book of Ezekiel prophesied a literal, still-future attack by the Soviet Union, which LaHaye claimed was referred to by Ezekiel as 'Gog' and/or 'Magog', on a country to its south, Israel. To me, this was a fringe belief—something I had never encountered before—but a fascinating one. I later learned that Ronald Reagan subscribed to essentially the same belief. Reagan reportedly raised some eyebrows at a 1971 dinner with California legislators where he said:

> Ezekiel says that fire and brimstone will be rained upon the enemies of God's people. That must mean that they'll be destroyed by nuclear weapons. They exist now, and they never did in the past. Ezekiel tells us that Gog, the nation that will lead all of the other powers of

> darkness against Israel, will come out of the north. Biblical scholars have been saying for generations that Gog must be Russia. What other powerful nation is to the north of Israel? None. But it didn't seem to make sense before the Russian revolution, when Russia was a Christian country. Now it does, now that Russia has become communistic and atheistic, now that Russia has set itself against God. Now it fits the description of Gog perfectly. [6, 7]

As it turns out, of course, both LaHaye and Reagan must have been off the mark at least in one respect: in 1991, the 'communistic and atheistic' Soviet Union ceased to exist, so at the very least, Ezekiel 38–39 was 'a prophecy that misread could have been'. Still, I also later learned that this interpretation of Ezekiel 38–39 did not start with LaHaye, Reagan, or even LaHaye's mentor in the American evangelical world, Hal Lindsey, but has actually been reappearing in one form or another for centuries, in different religious traditions, and in Europe and the Middle East as well as America; and while the idea of 'Russia = Gog and/or Magog' may always have been on the fringe, it was also unlikely to die along with the Soviet Union, given its long pedigree. Indeed, a Haaretz article [8] referencing this idea dates from as recently as 2014, mentioning Gog and Magog in connection with the Russian annexation of Crimea.

The very names, 'Gog and Magog', suggest that the 'Russia = Gog and/or Magog' idea is a good candidate for the type of unsupervised analysis we propose, because 'Gog' and 'Magog' are highly distinctive words in English, quite specific in combination to their application to this particular fringe belief. It would seem that these, as well as other associated words used by LaHaye, Reagan and other ('brimstone', 'nuclear weapons', 'atheistic') should stand out, marking the texts in which they occur as anomalous compared to a background of texts which might relate more generally to Russia and war. This is our hypothesis: does an implementation of unsupervised anomaly detection support it?

4.2 Curation of Dataset

To test the hypothesis, we selected a sample of 27 snippets of text, each between 101 and 267 characters long (to approximate to tweet-length text), each consisting of at least one complete sentence, from (1) Reagan's 1971 speech quoted above, (2) LaHaye's discussion of Russia and Magog [5], (3) Hal Lindsey's discussion of the same subject [9] (quoted at length in [5]), and (4) excerpts from what the Vilna Gaon was reported to have said [8]. We then ran a simple analysis to count the number of occurrences of each distinct word represented in the 27 text snippets and determine which words occurred most frequently. Among the most frequent were the following:

Russia—20 occurrences.

nations—5 occurrences.

allies—3 occurrences.

war—3 occurrences.

Russian—3 occurrences.

countries—2 occurrences.

These words can be thought of as a 'bridge' between the general topic area of this particular fringe belief, and mainstream ideas on the same topic.

To obtain the 'mainstream' portion of the dataset, we then used combinations of the above words to return a total of 33,470 tweets using the Twitter Developer API, which we then randomly downsampled using selective percentages of each set of search results, to end up with approximately 20,000 tweets, made up as follows:

Twitter API query	# Results returned	Downsampling (%)	Selected tweets
Russian allies	452	0	–
Russia countries	18,000	50	9,000
Russia war	10,469	75	7,852
Russia nations	2,940	100	2,940
Russia countries	1,609	50	805
Total	33,470		20,597

Some tweets were returned by multiple different queries above and duplicates were eliminated, reducing the 20,597 to 20,114. The hand-curated 27 snippets of 'fringe belief' text were then mixed into the Twitter dataset to create a new dataset of 20,141 texts.

4.3 Analytical Results

First, to give a sense of the analytical results, in Fig. 1 we present a heatmap showing inter-text similarities. In this heatmap, for presentation purposes, we further downsample the 20,141 × 20,141 adjacency matrix to a 527 × 527 matrix, to include just the first 500 Twitter texts plus the 27 'Gog and Magog' texts.

The color scheme in this heatmap represents pairs of identical texts (in a bag-of-words sense) as yellow points, and dissimilar texts as purple, with a continuum in between. There is an art to interpreting the heatmap and other outputs, but this does not imply interpretation of the outputs is a completely subjective matter–one might compare reading these heatmaps to how two skilled radiologists might read X-rays or CT scans and arrive at similar diagnostic conclusions.

Here, the yellow square near the center of the heatmap represents a cluster of identical texts, which in the context of Twitter, are likely to be retweeted content. This is of less interest than the blue cluster at the bottom right, which represents a group of thematically similar, but not identical, texts. It turns out that this cluster is the 27 'Gog and Magog' texts. We therefore have good prima facie evidence that the technique can successfully distinguish between 'fringe' and 'mainstream'. However, the evidence is not conclusive, because in a sense this heatmap made identification

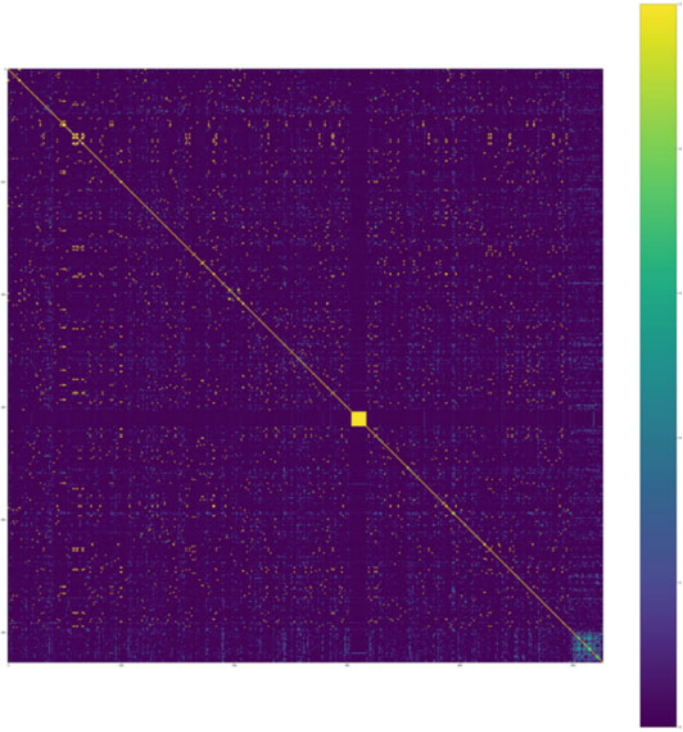


Fig. 1 Heatmap of inter-text similarities, Twitter plus 'Gog and Magog' texts

of the anomaly easier by ordering the tweets first and the 'Gog and Magog' texts second. A more conclusive piece of evidence is provided by, in effect, randomizing the order of the texts and then applying hierarchical clustering, which it is possible to do straightforwardly in a Jupyter notebook using the Python 'plotly' library to plot a heatmap and dendrogram together,[2] to determine if the cluster of fringe belief texts can still be identified. This output is presented in Fig. 2.

Within this heatmap, again, the bright yellow clusters are of less interest because these are clusters of identical or near-identical texts, which we expect to find in Twitter in the form of retweeted content. In texts representing fringe beliefs, one would expect to find more 'noise'. In this regard, the cluster highlighted on both the horizontal and vertical axes with a red brace (note that the right-to-left ordering on the horizontal axis is identical to the top-to-down ordering on the vertical axis) catches our attention, both because the hierarchical clustering represented in the dendrogram captures its anomalous nature, and because the associated part of the heatmap indicates the texts in this group are similar to one another, but not identical or near-identical. Using the Jupyter notebook, we can zoom in on the dendrogram (Fig. 3) to determine which texts are in this group and determine that the indices of

[2] See https://plotly.com/python/dendrogram/. (Retrieved on 6/20/2022.).

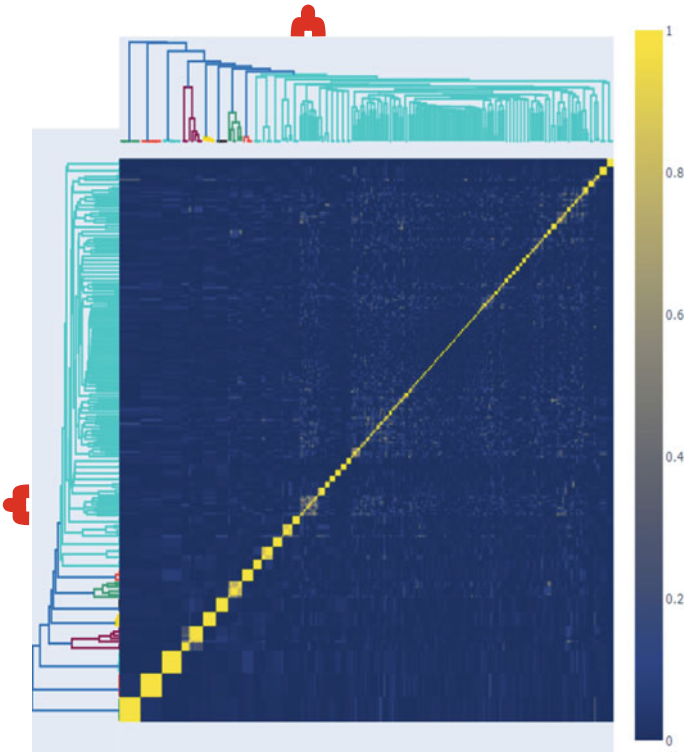


Fig. 2 Heatmap + dendrogram of 'Gog and Magog' + Twitter dataset

these documents are 135, 320, 500, 502, 504–506, 509–510, 512–516, 519, and 524. Indices 0–499 correspond to the general Twitter data, and 500–526 correspond to the 'Gog and Magog' texts. In other words, within the anomalous cluster we were able to identify visually, 14 of the 27 'Gog and Magog' texts are present, plus two others. Interestingly, of those two others, text 320 reads: 'Those Countries which Russia considers Sovereign will be re annexed by Russia thrusting the world into WW3'—not completely out of the 'ballpark', perhaps, for the 'Gog and Magog' fringe belief.

One objection may be that 14 out of 27 is only just above 50% recall (not as good as Robinson's in [2], possibly reflecting that this problem is a harder one given the specifics of the data), but considering that our objective was only to differentiate the fringe belief from the mainstream, not necessarily to retrieve *all* texts relating to the fringe belief, we think it is still reasonable for us to claim that Fig. 2 demonstrates that that objective has been achieved. Note that greater precision could potentially be achieved by a bootstrapping process. Having found the 'tip of the iceberg' of the Gog/Magog cluster, finding other similar documents is a well-understood and easily solvable problem in text analytics.

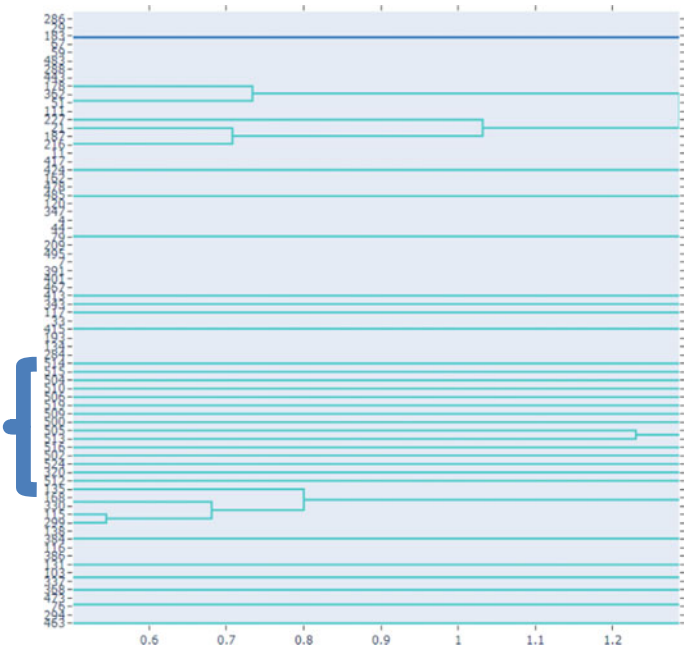


Fig. 3 Detail of anomalous cluster from Fig. 2

5 Application (2): Russian-Language Media Landscape in 2022

5.1 The Russian Media Landscape in 2022

2022 has been a pivotal year in Russia, with the February 24, 2022 invasion of Ukraine by Russia, the resulting Western sanctions placed on Russia, the flight of Western capital and firms from Russia, and increasingly stringent control by the Russian State over the media in the country. Emblematic of this control is the March 4, 2022 law passed by the Russian Duma [10] which subjects anyone who spreads what the Kremlin considers 'fake' information about Russian armed forces and the war in Ukraine to fines or prison terms of up to 15 years. One of the most reported-upon aspects in Western media of the Russian government's control over the portrayal of the war in Ukraine is that Russian media are not allowed to use the term 'war' but must instead refer to Russian involvement in Ukraine as a 'special military operation'.

In reality, State control over information in modern Russia has been consolidating for some time, as evidenced by a 2014 statement by Galina Timchenko. Timchenko was fired in 2014 from her job as chief editor of Lenta.ru, an online Russian news source based in Moscow. However, with some Lenta.ru colleagues, Timchenko relocated to Latvia and set up an independent Russian-language news organization called

Meduza. At the time, Timchenko explained her decision as follows: 'It's just the way it is: right now, in Latvia one can set up an independent Russian-language publication, but not in Russia' ('Просто так вышло, что в Латвии сейчас можно сделать независимое русскоязычное издание, а в России нет.') [11]. To the present day, Meduza continues to publish, in Russian, from its headquarters in Latvia.

Amidst the Kremlin's attempts to exert tight control over the Russian-language—a level of control which, as Timchenko said, does not reach into Latvia, albeit a neighboring country and former Soviet republic—our hypothesis is that the 'mainstream' of Russian-language media would be represented by sources based on Russian territory, and Meduza would, relative to that mainstream, appear to be on the fringe. To remind the reader, by referring to one type of source as 'mainstream' and another as 'fringe', we are not necessarily implying agreement either with the former or the latter. An intelligence analyst might disagree vehemently with the positions taken by Russian mainstream media, but still find it valuable to be able to characterize the Russian-language information environment and understand what that looks like from the perspective of residents of the Russian Federation.

5.2 Curation of Dataset

To test this hypothesis, we apply the same technique as for the 'Gog and Magog' and Twitter data. We collected close to 2,838 news articles from four different Russian news sources: three based in Russia, which we can assume to be more or less subject to the control over information exerted by the Kremlin, and Meduza.

We targeted collection of data towards articles posted from January 1, 2022 onwards, using a web scraper (a free, Chrome-extension version of Web Scraper, https://webscraper.io/), and we aimed to collect articles posted on each Monday in the period.[3] We hoped this would result in a somewhat representative sample of news articles, assuming a roughly weekly news cycle.

These procedures resulted in the following data collected:

Data source	Description	Number of articles
1tv.ru	Russia's 'Channel One' (Первый канал), a Russian state-controlled television channel, based in Moscow	407
gazeta.ru	A Russian news website based in Moscow and owned by the state-owned company Sberbank	1087
meduza.io	An independent Russian-language news website based in Latvia	611
rt.com	Also known as RT (formerly Russia Today), a Russian state-controlled international television network, likely intended to compete with Western cable news such as CNN or BBC World	733

(continued)

[3] This was to manage the amount of work in collecting and pre-processing data, not the core analytics which can easily scale to huge volumes of data.

(continued)

Data source	Description	Number of articles
Total		2838

5.3 *Analytical Results*

With the approach we followed in Sect. 4.3, we wanted to test the hypothesis that our technique could find a 'needle in a haystack' −27 'fringe belief' texts where we hid the labeling of those texts from the unsupervised anomaly detection technique. Here, because the four news sources are known to us a priori, it may make more sense to order the heatmap by news source, as in Fig. 4.

From Fig. 4, we see a *lack* of the bright-yellow clusters that we saw in Fig. 1, which is expected given that we are now looking at Russian news articles, not Twitter. It would be surprising if there were near-matches between articles, as this would point

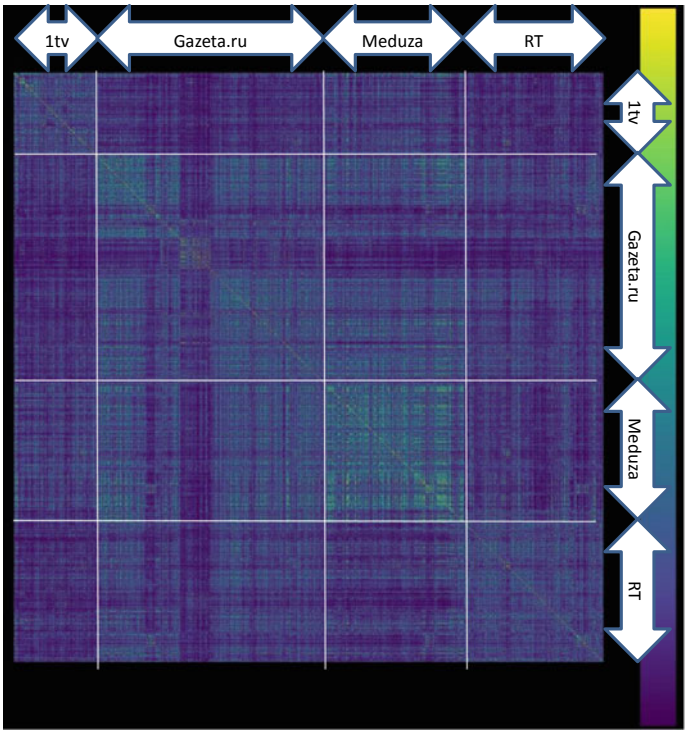


Fig. 4 Heatmap of inter-text similarities, Russian news

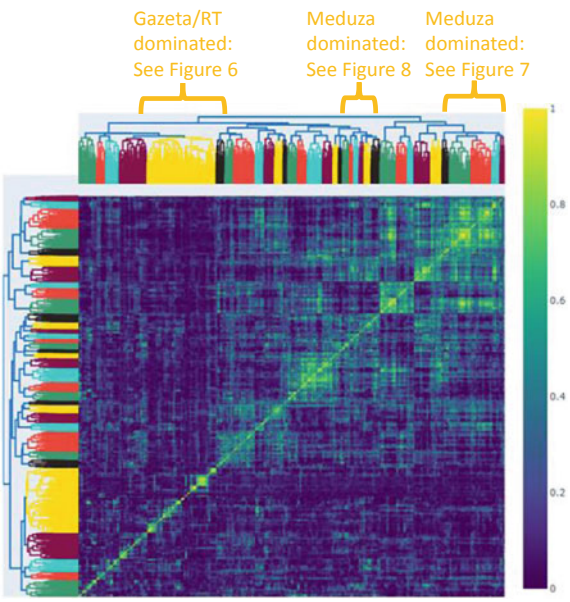


Fig. 5 Heatmap + dendrogram showing clusters and outliers

to plagiarism. In Fig. 4 we do see that Meduza is differentiated from the other news sources, in that the shading is a brighter blue in the Meduza/Meduza quadrant than in any other of the 16 quadrants. The heatmap also reveals that Meduza is more similar to Gazeta.ru than it is to either 1tv.ru or RT.com (a result which was surprising to us), and it also reveals inter-source similarities or relationships between RT.com and Gazeta.ru.

It is also possible in effect to randomize the ordering of articles, and let the ordering implied by the hierarchical clustering reveal its own patterns, which we do in Fig. 5. Here, again, the skilled reader of the output will be looking for areas of the plot which show evidence both of anomalous clustering, and anomalous patterns of similarity compared to the background noise. Clusters that stand out for their anomalous or homogeneous characteristics (respectively) are shown in Figs. 6, 7 and 8. Zooming in to the axis labels, we discover that some of the more homogeneous clusters (Figs. 6, 7 and 8) consist of a large proportion of RT.com and Gazeta.ru articles; the more anomalous (more variegated) clusters (Figs. 7 and 8) consist of a relatively large number of Meduza articles. This lends support to our hypothesis that the technique is successful in distinguishing between 'mainstream' and 'fringe', and, in line with expectations, that Meduza is often an outlier—part of the 'fringe'—in the Russian-language news landscape.[4]

[4] Though Meduza is based in Latvia, not Russia, regional dialect can safely be ruled out as a factor in its anomalousness. This is partly because the staff of Meduza relocated from Russia to Latvia to escape media suppression, and partly because Russian dialects around the former USSR in any case tend to vary much less than English dialects.

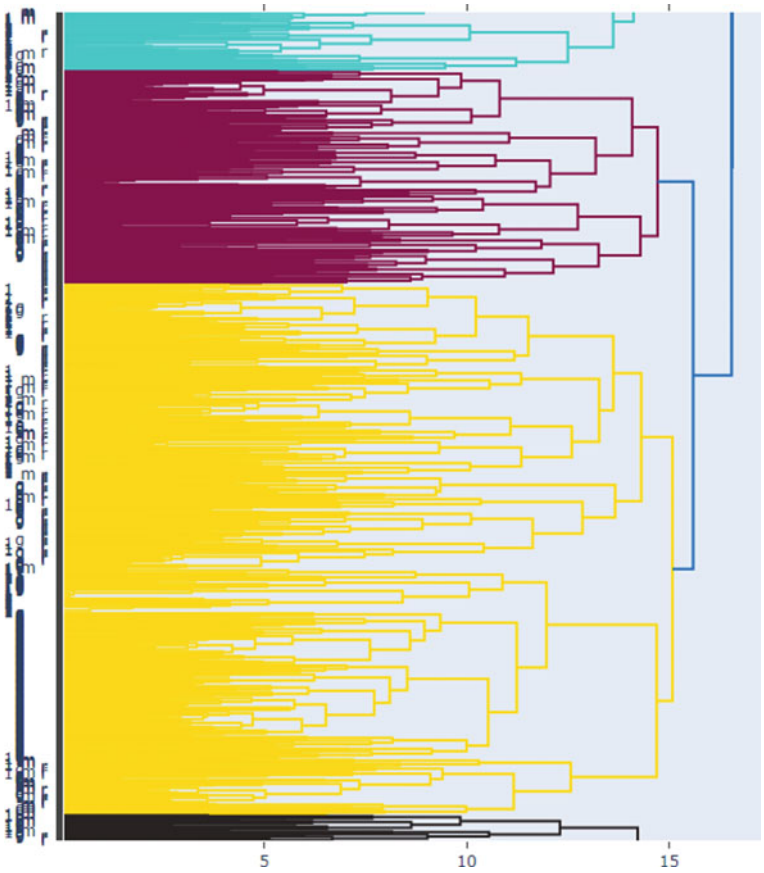


Fig. 6 Detail of homogeneous clusters, highly dominated by Gazeta (g) and RT (r)[5]

5.4 Digging Deeper

Setting aside SVD and its application to anomaly detection in this section, we can dig deeper and discover statistically what, specifically, makes Meduza different from the other Russian-language sources. A Keyness Relative Frequency analysis is one of the best ways to determine differences in relative vocabulary frequency between sets of 'target' and 'reference' documents. To compute the Keyness score for a given word, four numbers are needed: the frequency of the word in the target corpus, the frequency of the word in the reference corpus, and the number of other words in both sets of documents (see Table 1).

[5] Y-axis labels in this and subsequent figures are tabulated to aid in distinguishing which are '1' (1TV), 'g' (Gazeta), 'm' (Meduza) and 'r' (RT) where the labels overlap.

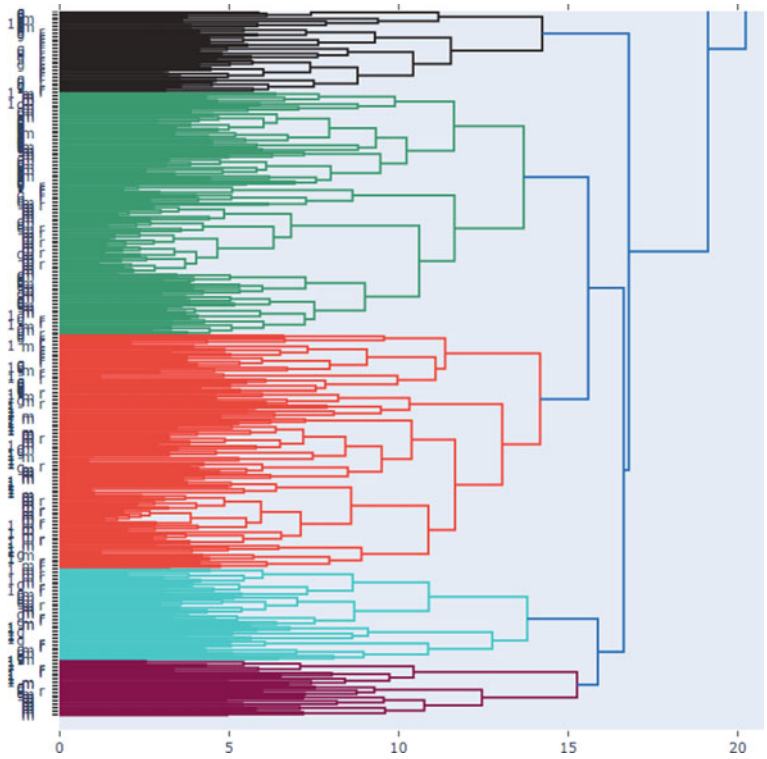


Fig. 7 Detail of anomalous cluster, more dominated by Meduza (m)

The Chi-squared Keyness value is then given by the following formula, where higher values represent increased relative frequency in the target corpus:

$$\chi = \frac{N(ad - bc)^2}{(a+b)(c+d)(a+c)(b+d)}$$

Here, we selected just Meduza and RT. Keyness analyses reveal clear differences between the independent and State-controlled sources (Fig. 9).

This entirely statistical analysis was fully in line with our expectations, given the Kremlin's restrictions on how the war in Ukraine is reported. Among word n-grams most associated with RT, the source based in Russia, we find forms of the words for 'special operation Ukraine' (спецоперация Украине), whereas for Meduza, forms of the word for 'war' (война) are in each of the top 5 n-grams. The statistics confirm very clearly that Kremlin control over the information space has tangible and measurable effects. The independent source Meduza is, in terms of Russian media, 'fringe', and it is fringe in exactly the way we would expect.

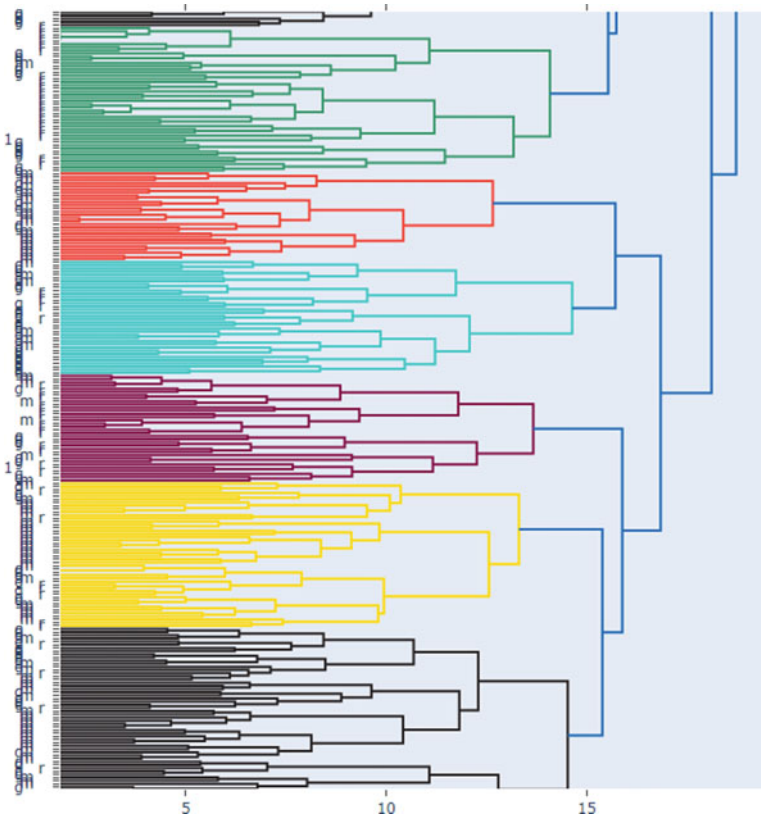


Fig. 8 Detail of anomalous clusters, many dominated by Meduza (m)

Table 1 Word frequencies used in Keyness analysis

	Target corpus	Reference corpus	Total
Word w	a	b	$a + b$
Not word w	c	d	$c + d$
Total	$a + c$	$b + d$	$N = a + b + c + d$

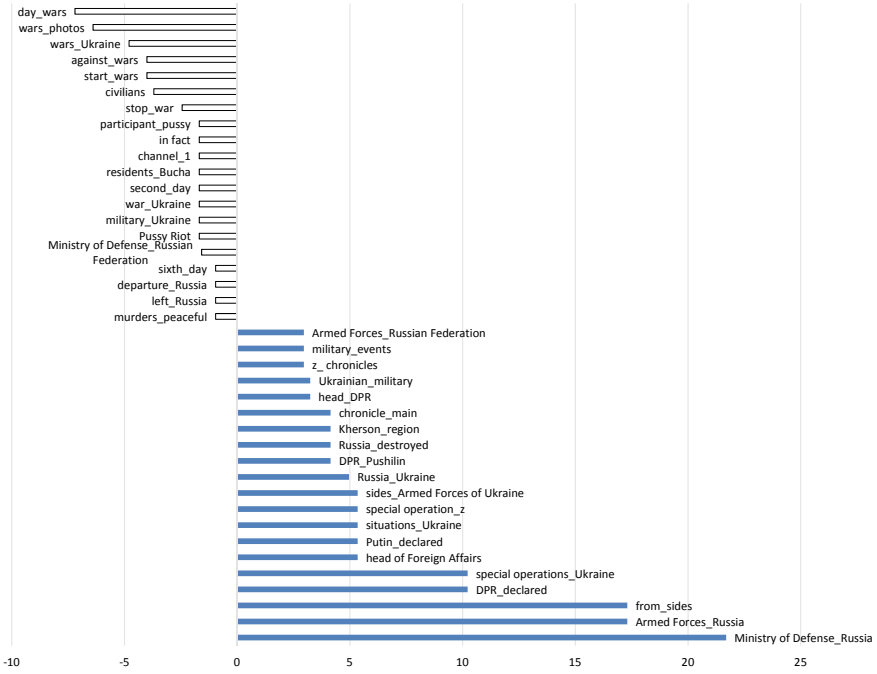


Fig. 9 Keyness analysis of RT (blue) versus Meduza (gray): Article Titles, Bigrams

6 Conclusion

In this paper we have demonstrated a novel application of unsupervised anomaly detection which we have shown can distinguish between 'mainstream' and 'fringe' beliefs, even when we do not know what particular needle we are looking for in the haystack of information. Our method draws upon the insight of Montell [1] that cults (understood broadly) tend to have their own distinctive vocabularies, and that is because language frames how people think about issues, and can even be used to manipulate thought patterns of large groups of people. We emphasized that the 'fringe' and the 'mainstream' are determined here by the data itself, not by extrinsic or a priori ideas about what should be accepted and what should not. By casting the problem in this way, we play to the strengths of unsupervised analysis; but we also appropriately leave it to the judgment of a human analyst whether the 'fringe' has it 'right' and the 'mainstream' does not—or vice versa.

We applied this technique to two datasets (both relating to Russia) and showed that, as expected, the technique is able to draw the analyst's attention to text representing a fringe belief about Russia; and that the technique also tends to confirm that in the Russian media landscape where the Kremlin exerts its sway, an independent Russian-language news source is indeed an outlier on the 'fringe'.

Nothing about the technique, fundamentally, is geared towards any particular dataset and it could be applied very generally to text data in different languages and from different domains. In this respect, we see this technique as being of general use and valuable to intelligence analysts seeking to make sense of unfamiliar information environments, looking for 'unknown unknowns', and trying to gain a top-down understanding of different information 'landscapes'. The technique focuses the analyst's attention on what the data itself tells us is 'mainstream' and 'fringe', and allows the analyst to proceed from there to dig deeper and understand what differentiates one from the other.

Acknowledgements Galisteo Consulting Group, Inc., gratefully acknowledges the funding and support provided by the Office of Naval Research for the development of this report. This material is based upon work supported by the Office of Naval Research under Contract No. N00014-21-P-2006. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Office of Naval Research.

References

1. Montell, A.: Cultish: The Language of Fanaticism. HarperCollins, New York, NY (2021)
2. Robinson, D.: Statistical Language Analysis for Automatic Exfiltration Event Detection. SAND Report, Albuquerque: Sandia National Laboratories (2010). https://www.osti.gov/servlets/purl/983675
3. Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J.B., Grout, J., Corlay, S., Ivanov, P., Avila, D., Abdalla, S., Willing, C.: Jupyter development team. In: Loizides, F., Schmidt, B (eds.) Jupyter Notebooks—A Publishing Format for Reproducible Computational Workflows, Players, Agents and Agendas, Positioning and Power in Academic Publishing, pp. 87–90
4. Grinstein, G., Scholtz, J., Whiting, M.A., Plaisant, C.: VAST 2009 Challenge: An Insider Threat. United States (2009). https://doi.org/10.1109/VAST.2009.5334454
5. LaHaye, T.: The Beginning of the End. Tyndale House Publishers, United States (1991)
6. Boyer, P.: When Time Shall be no More: Prophecy Belief in Modern American Culture. The Belknap Press of Harvard University Press, Cambridge, Massachusetts and London, UK (1992)
7. Lee, L.: The enemies within: Gog of Magog in Ezekiel 38–39. HTS Teologiese Studies (2017). Accessed 13 Sept 2021. https://www.gale.com/apps/doc/A520512732/AONE?u=nm_p_oweb&sid=googleScholar&xid=bd26297b
8. Haaretz.: FYI: Putin = Gog, Crimea = Magog, the Apocalypse Is Here and the Messiah Is Coming. Haaretz (2014). Accessed 13 Sept 2021. https://www.haaretz.com/.premium-fyi-putin-gog-crimea-magog-1.5341056
9. Lindsey, H.: The Late, Great Planet Earth. Zondervan, United States (1970)
10. Государственная Дума. 2022. Вводится ответственность за распространение фейков о действиях ВС РФ. Accessed from http://duma.gov.ru/news/53620/ on 20 June 2022
11. Сурганова, Елизавета.: "Галина Тимченко: «Никто из нас не мечтает делать «Колокол»." forbes.ru (2014). Accessed 14 June 2022. https://www.forbes.ru/kompanii/internet-telekom-i-media/267611-galina-timchenko-nikto-iz-nas-ne-mechtaet-delat-kolokol.
12. Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K., Harshman, R.: Indexing by latent semantic analysis. J. Am. Soc. Inf. Sci. **41**, 391–407 (1990)
13. Gallagher, R., Reagan, A.J., Danforth, C.M., Dodds, P.S.: Divergent discourse between protests and counter-protests: #BlackLivesMatter and #AllLivesMatter. PLoS ONE **13**, 4 (2018). https://doi.org/10.1371/journal.pone.0195644

14. Nichols, R., Slingerland, E., Nielbo, K., Bergeton, U., Logan, C., Kleinman, S.: Modeling the contested relationship between Analects, Mencius, and Xunzi: Preliminary evidence from a machine-learning approach. J. Asian Stud. **77**(1), 19–57 (2018). https://doi.org/10.1017/S0021911817000973
15. Hossein, N., Tran, T.T.T., Kautz, H.: Discovering Political Slang in Readers' Comments. In: Proceedings of the Twelfth International AAAI Conference on Web and Social Media (ICWSM 2018), pp. 612–615 (2018).

Evaluating a Crowd Logistics Network Using Agent-Based Modeling

Preetam Kulkarni and Caroline Krejci

Abstract Crowd logistics is a part of the sharing economy in which individual carriers offer to transport and deliver items for other individuals or businesses for a fee via an online platform. While crowd logistics platforms have the potential to offer more flexible and responsive delivery services for much lower rates than traditional logistics providers, it is challenging for platforms to achieve a critical mass of participants on both sides (senders and carriers) to allow the service to grow and thrive. This research uses agent-based modeling to explore the effects of participant behavior on the performance of a two-sided crowd-sourced logistics platform. Preliminary experimentation with the model tests the effects of heterogeneous agent decision logic on platform performance, including service level and network growth over time. Results demonstrate significant differences in performance between heterogeneous and homogeneous decision rule assignment and suggest that agent-based modeling is a particularly suitable method for studying the behavior of crowdsourced platforms.

1 Introduction and Background

Crowd logistics is a part of the sharing economy in which individuals (carriers) offer to transport and deliver items for other individuals or businesses (senders) for a fee. It is a novel way of providing logistics services that taps into the underutilized logistics resources and capabilities of individuals, using mobile applications and web-based platforms [3]. Carriers often use their personal vehicles for transport, and deliveries do not necessarily require an additional trip for them—they may instead leverage their existing travel patterns to earn additional income, such as dropping off a package while commuting to work [11]. For senders, the primary advantage of crowd logistics is the ability to get better service at a lower price than traditional logistics providers can offer, including same-day deliveries.

P. Kulkarni · C. Krejci (✉)
University of Texas at Arlington, Arlington, TX 76019, USA
e-mail: caroline.krejci@uta.edu

Z. Yang and S. Núñez-Corrales (eds.), *Proceedings of the 2022 Conference of The Computational Social Science Society of the Americas*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-031-37553-8_2

A platform is successful and self-sustaining only when its network of participants either grows in size or at least remains constant after a period of growth. This requires participation of a minimum number of participants on both sides, which is referred to as critical mass [14]. Based on the behavior of the participants, it can be very difficult to achieve critical mass. Thus, it is important to understand the motivations of both carriers and senders to ensure the success and growth of a crowd logistics platform. The more attractive the logistics value proposition is for users in terms of proximity, speed, adaptability, or accessibility, the stronger the growth of the initiative [3]. Furthermore, for a sender it is important to know when the goods are delivered and that the delivery process is safe [8]. However, research on sharing economy platforms indicates that there can be a wide variety of other motivations. [2] identified materialism, sociability, and volunteering as predictors of sharing motives in different sharing contexts. [13] identified desire for payment, competence/skill development, and social affiliation to be motivators for crowdsourcing platform participation. [1] found economic and social motivators (i.e., making/saving money; meeting new people) for both the customer and service provider, while hedonic value, reduced risk, and environmental benefits also motivated customer participation, and entrepreneurial freedom appealed to service providers.

While analytical and statistical models have been used to explain/predict participation of carriers in crowd logistics platforms (e.g., [6]), such models cannot capture the dynamics of two-sided participation, especially network effects, and they cannot discern the conditions that may lead to achieving (or failing to achieve) a critical mass of participants. Agent-based modeling (ABM) is useful for this purpose. For example, [15] modeled the relationships among properties of tasks, characteristics of workers, and performance metrics of a crowd-sourcing platform via ABM. The model was validated by running experiments with real human workers on Amazon Mechanical Turk. However, the authors did not use the model to examine macro-level metrics like platform growth over time. [5] modeled the effects of game theory experience on humans' strategic behavior by creating a two-sided dynamic interactive simulation in ABM and studied the macro level outcome i.e., coalition structure. ABM was also used by [4] to evaluate service levels and asset utilization in a crowd logistics platform, however, only the behavior of carriers who deliver the packages was modeled. By contrast, [12] created an ABM of a food rescue program that captures the behavior of both the senders (restaurants) and carriers (volunteers) who deliver food to people in need. Experimentation with the model demonstrated that one of the important factors for the success of a platform is to have sufficient participating carriers available when the platform is initially launched, as well as ensuring that sender and carrier participation remains balanced as the platform grows over time.

This paper describes an agent-based model that represents a stylized crowd logistics system in which heterogeneous carrier and sender agents decide on a daily basis if they will participate in the system, and if so, how many jobs they will participate in and which ones. Experimentation with the model demonstrates the effects of decision rule heterogeneity on carrier and sender participation in the system over time, as well as total number of successfully completed jobs. The next section provides a

description of the model, followed by a description of experimentation and results, and finally a discussion of results and a plan for future research.

2 Model Description

The ABM was developed using NetLogo 6.2.2 and is described using Overview, Design Concepts and Details (ODD) protocol [9]. The updated protocol which fixed the issues and ambiguities of the original protocol was followed [10].

2.1 Purpose

The objective of this model is to perform exploratory research on the factors that influence network effects in a two-sided crowd logistics platform with heterogeneous participants. The long-term intent of this model is to use it to gain a better understanding of how the features and affordances of a crowd logistics platform should be designed to encourage participation from carriers and senders, such that the platform grows and sustains itself. Of particular interest are systems that are not intended for purely commercial goals, e.g., systems with humanitarian or community-supporting aims.

2.2 Entities, State Variables and Scales

Entities include senders, carriers, and destinations. Senders participate in the platform to ship packages to specific destinations, while carriers are trying to earn income by delivering the packages on behalf of the senders. There are 250 agents that can take on a sender's role, 250 agents that can work as carriers, and 8 destinations. It is assumed that an agent can participate in the platform only in its assigned role—it can be a sender or a carrier, but not both. Table 1 summarizes the key attributes and state variables of all three agent classes.

2.3 Process Overview and Scheduling

At the beginning of each day (tick), all sender agents that are participating in the platform generate delivery requests. A participating carrier is then chosen randomly to bid on a randomly selected delivery request. If the sender accepts the bid, the carrier executes the delivery and then waits at that destination for its next opportunity to bid on a delivery request. Another carrier is then randomly selected to evaluate a

Table 1 Attributes of sender and carrier agents

Agent class	Attribute	Description	Possible values
Sender and carrier	Identification number (ID)	Unique identifier	[1–500]
Sender and carrier	Participation status	Participation status of an agent on a given day of the simulation	Y/N
Sender and carrier	Participation history	List of number of delivery requests fulfilled (senders) or number of deliveries completed (carriers) in each tick	Varies by tick
Sender and carrier	Number of friends	Length of the list of IDs of agents that are "friends"	Normal (7, 2)
Sender and carrier	Participation evaluation period	# of days after which a sender or a carrier agent evaluates its participation decision	Experimentally varied
Sender and carrier	Participation decision rule	# of friends that need to participate during evaluation for the agent to join the platform	Experimentally varied
Sender and carrier	Continued participation decision rule	Min. # of deliveries/trips per day needed for the agent to continue participation	2
Sender	No. of pickup requests	Number of pickup requests by sender agents each day	Uniform (1, 8)
Sender	Pickup request list	List of who # of destinations	Varies by tick
Sender	Assigned carriers ID	List of carrier IDs who matched with a sender on a day	Varies by tick
Carrier	Home coordinates	Carrier returns to initial coordinates every day after deliveries	X: [-33, 33] Y: [-33, 33]
Carrier	Target sender ID	List of sender IDs that matched with a carrier on a day	Varies by tick
Carrier	Target destination	List of who numbers of destinations	Varies by tick
Carrier	Total trip distance	List of distances traveled for each trip in the current tick	Varies by tick
Carrier	Total trip income	List of income earned for each trip in the current tick	Varies by tick
Carrier	Daily trip limit	Limit on # of trips that a carrier agent is willing to complete	Experimentally varied
Carrier	Trip profit percentage	Percentage of the trip cost that is charged by the carrier as profit	Normal (12, 2)
Destination	Visitors	List of IDs of carriers that visit a destination	Varies by tick

delivery request, and this process continues until all carriers have reached their daily trip limits and/or all requests have been fulfilled.

The participation history of each sender and carrier is then updated. Based on this history and the participation of other agents in their social network (i.e., friends), the agents make decisions about future participation: whether to continue participating in the platform, leave the platform, or join the platform if they are currently not participating.

2.4 Design Concepts

- *Basic principles*—Agents decide whether or not they will join the platform based on the participation levels of their social networks. Carriers and Senders bid and accept/reject bids, respectively, based on an evaluation of the estimated monetary value of the job, in terms of transportation cost and time. They decide whether to continue participating in the platform based on their ability to successfully match with a sender/carrier over time.
- *Emergence*—The number of participants on the platform, the number of successful deliveries, the percentage of carriers working at full capacity, and the percentage of senders whose delivery requests are fulfilled are all emergent performance metrics.
- *Adaptation*—The sender agents respond to the lack of carrier agents to send their packages by quitting the platform. Similarly, carrier agents respond to the lack of opportunities to earn income on the platform by quitting the platform. The participants also decide to rejoin the platform if enough of their friends say that they are participating and satisfied with the platform.
- *Objectives*—The objective of the senders is to be able to find an affordable alternative to send packages to a particular destination (i.e., rather than making the delivery itself), whereas the carrier's objective is to be able to earn an acceptable amount of profit by making trips to deliver sender agents packages.
- *Interaction*—Senders and carriers interact via the platform by placing and accepting bids for delivery jobs, respectively.
- *Stochasticity*—Some agent attributes are stochastic; see Tables 1 and 2.
- *Observation*—Crowd logistics system performance metrics include the number of successful deliveries, the number of participating carriers and senders, the percentage of carriers working at maximum capacity, and the senders who find a match for all of their request's help. These values are captured over the length of the simulation runs (i.e., in each tick).

2.5 Initialization

At the beginning of each simulation run, the home coordinates and IDs of all 250 senders and 250 carriers are read from a text file. The senders and carriers are then created and positioned on their home coordinates. *Participation status* of a subset of randomly selected senders and carriers is changed to "Yes"; the sizes of these subsets are experimentally varied (see Table 2). Each sender and carrier then build's its social network by selecting *friends* of their own kind, i.e., senders select other senders as friends and carriers select other carriers as friends. The initial values of all other agent attributes (Tables 1 and 2) are then assigned.

2.6 Input Data

The model does not use input data to represent time-varying processes [10].

2.7 Sub-Models

The model contains three sub-models, which are described below.

Sub-model 1—Join or leave the platform

After waiting for *Participation evaluation period* number of ticks, each sender and carrier that is not currently participating in the platform will evaluate whether or not they want to join. The decision to join the platform is taken by the participant if the number of its friends currently participating is at least equal to *Participation decision rule*. After participating for at least the *Participation evaluation period*, each participating sender and carrier evaluates whether it wants to continue participating. The agent will decide to leave the platform if the number of times its delivery request/bid was accepted in the most recent *Participation evaluation period* is less than *Continued participation decision rule* multiplied by *Participation evaluation period*. Otherwise, the agent will continue participating for another tick.

Sub-model 2—Delivery requests, bids and matching

First, all participating senders generate their lists of delivery requests for the day. A delivery request consists of the who number of the destination. Then, each participating carrier is selected in a random order to review a randomly selected delivery request. The carrier calculates the Euclidean distances between the carrier's current location, the sender's location (for pickup), and the delivery destination. The carrier then calculates the cost of making this trip and the amount it will charge to make its desired profit, and it shares this bid with the sender. The sender will only accept the bid if the amount the carrier charges for the trip is less than or equal to the sender's

cost of performing the delivery itself. If the sender accepts, the carrier executes the trip by moving from its current location to the sender's location and then to the destination to make the delivery. This process is repeated for each participating carrier. In each tick, the maximum number of deliveries that a carrier can make is equal to its *Daily trip limit*. Once this limit is reached, the carrier will no longer place bids on delivery requests. Furthermore, if a carrier is unable to find a match with any of the senders, the carrier is not picked randomly for the match process again during the remainder of the day. This process is summarized in the flowchart shown in Fig. 1.

Sub-model 3—Update participation history

After the bidding process is completed, each carrier records the number of successful trips it completed, and each sender records the number of its delivery requests that were fulfilled.

3 Experimentation and Results

A preliminary set of experiments was performed to determine the effects of agent heterogeneity on platform performance over time. First, a baseline scenario was run (Scenario 1), in which all agents' participation evaluation period was set to 7 ticks, all carriers' daily trip limits were set to 10 trips, and the participation decision rule for all agents was defined as having a minimum of 3 friends participating before joining. In reality, however, participants are likely to have heterogeneous participation criteria. To capture this, two additional experiments were performed, in which the values defining participation evaluation period, participation decision rule, and daily trip limits were assigned to agents randomly during model initialization, using a uniform distribution (Scenario 2) and a normal distribution (Scenario 3). For these three scenarios, the model was initialized with 22 participating carriers and 20 participating senders. To examine the impact of a platform's start-up conditions on its performance over time, the model was also run with 9 carriers and 8 senders initially participating (Scenario 4). Table 2 summarizes the experimental conditions for all four scenarios.

For each scenario, the model was run for 100 ticks, and the following platform performance measures were captured in each tick:

- Number of senders and carriers participating
- Number of senders and carriers that were able to find at least one match
- Percentage of carriers that were fully utilized (i.e., the number of completed deliveries was equal to the agent's daily trip limit)
- Percentage of senders with all delivery requests met
- Total number of successfully completed deliveries

Figure 2 shows an example of typical results over 100 ticks for Scenario 1. Results indicate that the platform is unable to attract new participants, and participation of both senders and carriers gradually but steadily declines and is unlikely to be sufficient

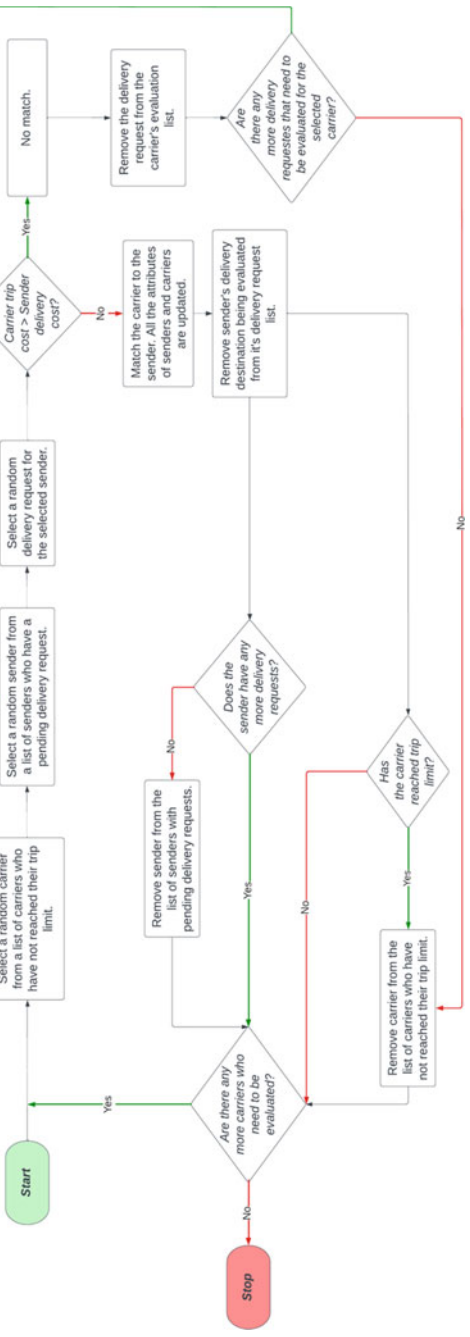


Fig. 1 Flowchart for the match process

Table 2 Experimental scenarios

Scenario #	# Of initial participating carriers	# Of initial participating senders	Participation evaluation period (ticks)	Carrier daily trip limit (trips)	Participation decision rule (min # friends)
1	22	20	7	10	3
2	22	20	Uniform (2, 7)	Uniform (1, 10)	Uniform (1, 3)
3	22	20	Normal (4.5, 2)	Normal (5, 2)	Normal (3, 2)
4	9	8	Normal (4.5, 2)	Normal (5, 2)	Normal (3, 2)

to maintain the platform in the long run. Participation never gains momentum, likely because the participation evaluation period of all agents is long (i.e., 7 ticks), and agents must see at least 3 of their friends participating before joining, such that early platform growth via network effects is stifled.

By contrast, in Scenario 2 there is a rapid increase in sender and carrier participation, with all agents continuing to participate for the remainder of the simulation run. The most likely reason for this early and rapid increase is because there are relatively many participants that have a short evaluation period (i.e., 2 or 3 ticks) and a low barrier to joining (i.e., only requiring one friend to be a participant). Once the platform has this critical mass of early participants, network effects allow it to continue growing very quickly via the agents' social networks and to continue to sustain itself, since participants can easily satisfy their continued participation rule (i.e., a minimum of 2 pickup requests filled/bids accepted per tick on average over the evaluation period). The participation levels off after 15 days with all or nearly all agents participating. As a result, Scenario 2 typically yields around 20 times as many successful deliveries per day as Scenario 1. Half of all carriers are being fully

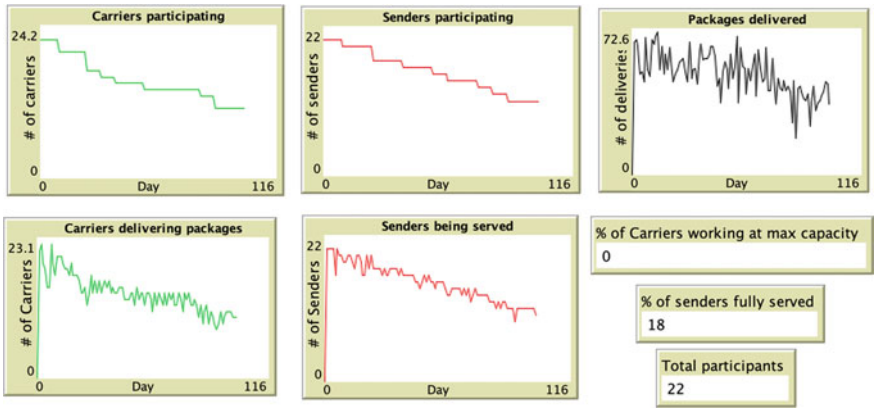


Fig. 2 Example output of Scenario 1

utilized at the end of the run, and more than a third of all senders have all of their delivery requests filled (Fig. 3).

In Scenario 3, participation also increases and eventually maintains itself, although the rate of increase is not nearly as rapid as in Scenario 2. As seen in Fig. 4, it took around 40 days for the platform to level off at 200 carriers/senders. This is probably because many agents' evaluation period is concentrated around the mean participation evaluation period of 5 ticks, such that it takes longer for most agents to decide to join. However, the platform does not decline as it does in Scenario 1, probably because there are enough agents with low participation evaluation period/decision rule values to ensure that there are sufficient participants in the early stages to attract and retain others.

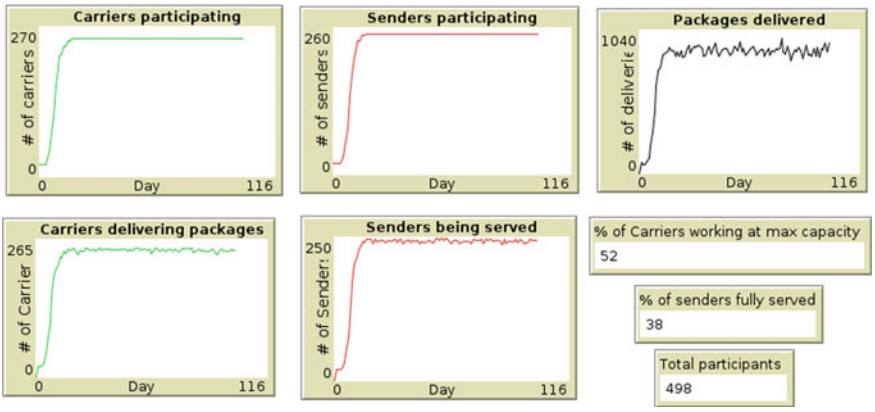


Fig. 3 Example output of Scenario 2

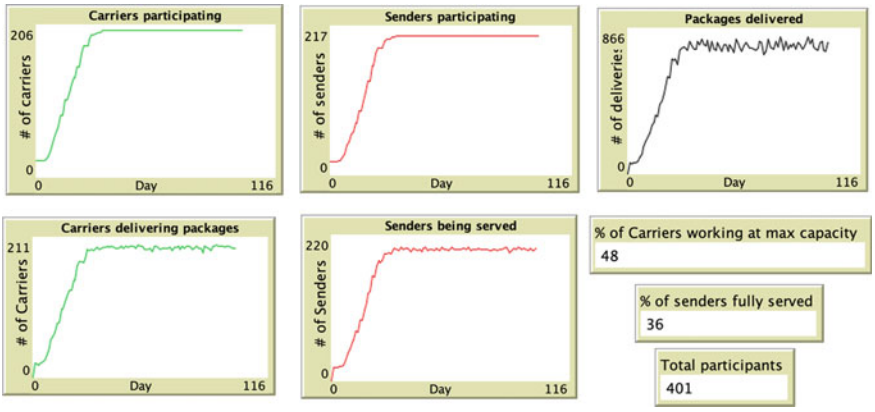


Fig. 4 Example output for Scenario 3

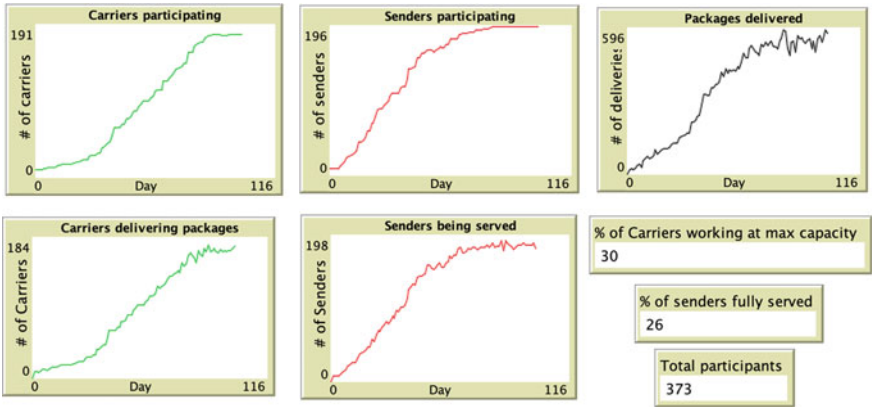


Fig. 5 Example output for Scenario 4

Figure 5 shows the results of Scenario 4, in which the participation rules are again assigned to agents via a normal distribution, but the numbers of initial senders and carriers participating are reduced from 20 and 22 to 8 and 9, respectively. As in Scenario 3, the number of participants increases over time; however, the rate of increase is much more gradual, with growth peaking and leveling off at around 75 days and the percentage of participants fully utilized/served drops to 30% for carriers and 26% for senders. This result indicates the importance of initial participation for platform growth over time, which was similarly observed in [12].

Figures 6 and 7 summarize the average values of the performance metrics at the end of the 100th tick for Scenarios 1, 2, and 3 over 30 replications. In Fig. 6, it can be seen that the participation and number of successfully completed deliveries are greater when the platform participation rules are randomly assigned to agents using a uniform distribution (Scenario 2) as compared with rule assignment via a normal distribution (Scenario 3). This result is likely due to the carrier daily trip limit in Scenario 2 being assigned to agents evenly across a range of 1 to 10 trips, whereas in Scenario 3 the carriers' daily trip limit is centered around a mean of 5 trips, such that carrier trip limit is limiting the system's capacity to make deliveries. While the performance of the platform is poor in Scenario 1, this result might not hold true if the platform had more initial participants from both sides and the evaluation period was less.

Figure 7 indicates that the percentage of fully-utilized carriers and fully-served senders is also highest when participation rules are distributed uniformly (Scenario 2), likely because overall participation in Scenario 2 is higher than in Scenarios 1 and 3. However, the values of these metrics are only slightly lower if participation rules are distributed normally (Scenario 3). The percentage of fully-utilized carriers is slightly higher than the percentage of fully-served senders in both Scenarios 2 and 3, probably because there are more jobs for the carriers to take than their trip limits. Hence, there are more fully utilized carriers than senders who are fully served. This

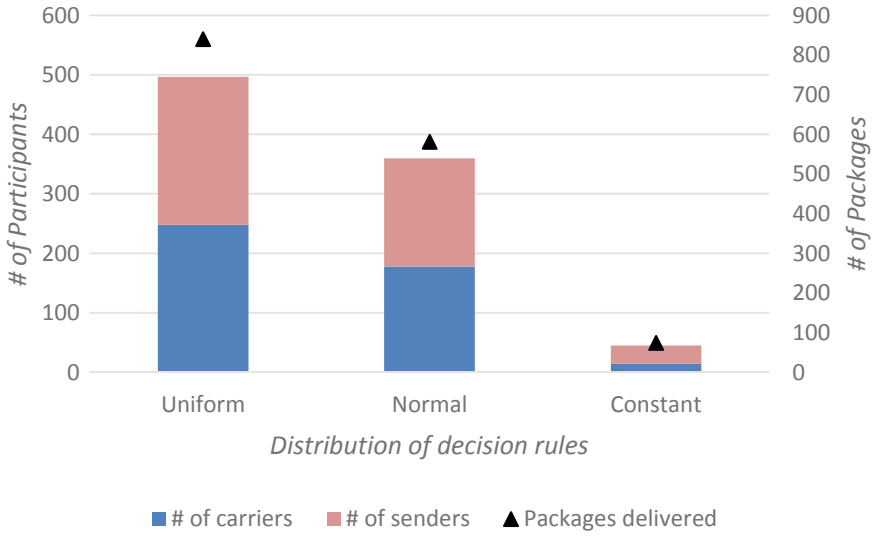


Fig. 6 Average number of participants and completed deliveries after 100 ticks for Scenarios 1, 2, and 3

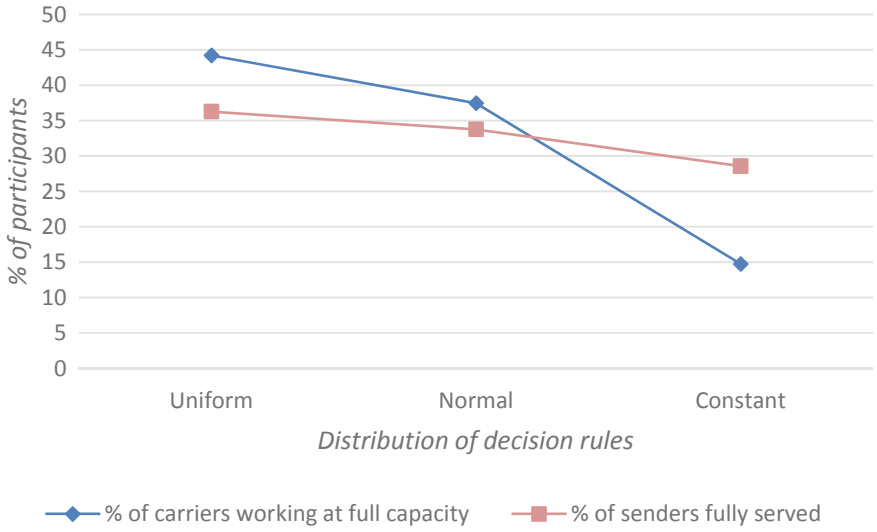


Fig. 7 Average percentage of carriers fully utilized and senders fully after 100 ticks for Scenarios 1, 2, and 3

trend is the opposite when the decision rules have a constant value. Owing to the higher participation of senders than carriers who have a high daily trip limit, carriers are not fully utilized whereas the senders find a match for their jobs more often and are fully served.

4 Conclusion

This paper described an agent-based model of a stylized crowd logistics system in which sender's delivery requests are matched with carrier agents on a daily basis. The model was used to perform preliminary experimentation to explore the effects of heterogeneous assignment of platform participation decision rules to agents on the platform's ability to grow and sustain itself over time. The relationship between the number of initial participants and the growth of platform participation is also examined.

Ongoing experimentation with this model explores platform behavior for a wide range of parameter settings for sensitivity analysis. The output of this model will also be compared with analytical models that predict critical mass and platform growth (e.g., the models described by [7]. Future research will involve empirical human behavior data collection from actual crowd logistics platform participants via participatory ABM techniques; this data will then be incorporated into the model in an effort to create a valid representation of a crowd logistics system.

Acknowledgements This material is based upon work supported by The National Science Foundation under Award No. 2046632.

References

1. Benoit, S., Baker, T.L., Bolton, R.N., Gruber, T., Kandampully, J.: A triadic framework for collaborative consumption (CC): Motives, activities and resources and capabilities of actors. J. Bus. Res. **79**, 219–227 (2017). https://doi.org/10.1016/j.jbusres.2017.05.004
2. Bucher, E., Fieseler, C., Lutz, C.: What's mine is yours (for a nominal fee)–Exploring the spectrum of utilitarian to altruistic motives for Internet-mediated sharing. Comput. Hum. Behav. **62**, 316–326 (2016). https://doi.org/10.1016/j.chb.2016.04.002
3. Carbone, V., Rouquet, A., Roussat, C.: The Rise of crowd logistics: a new way to co-create logistics value. J. Bus. Logist. **38**(4), 238–252 (2017). https://doi.org/10.1111/jbl.12164
4. Chen, P., Chankov, S.M.: Crowdsourced delivery for last-mile distribution: an agent-based modelling and simulation approach. In: 2017 IEEE international conference on industrial engineering and engineering management (IEEM), pp. 1271–1275 (2017). https://doi.org/10.1109/IEEM.2017.8290097
5. Collins A.J., Etemadidavan, S.: Interactive agent-based simulation for experimentation: a case study with cooperative game theory. Modelling **2**(4), Article 4 (2021). https://doi.org/10.3390/modelling2040023

6. Ermagun, A., Stathopoulos, A.: To bid or not to bid: an empirical study of the supply determinants of crowd-shipping. Transp. Res. Part A: Policy Pract. **116**, 468–483 (2018). https://doi.org/10.1016/j.tra.2018.06.019
7. Evans DS, Schmalensee R.: Failure to launch: critical mass in platform businesses. SSRN (2010). https://www.dspace.mit.edu/handle/1721.1/76685
8. Frehe, V., Mehmann, J., Teuteberg, F.: Understanding and assessing crowd logistics business models—using everyday people for last mile delivery. J. Bus. Ind. Mark. **32**(1), 75–97 (2017). https://doi.org/10.1108/JBIM-10-2015-0182
9. Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S. K., Huse, G., Huth, A., Jepsen, J.U., Jørgensen, C., Mooij, W.M., Müller, B., Pe'er, G., Piou, C., Railsback, S.F., Robbins, A.M., et al.: A standard protocol for describing individual-based and agent-based models. Ecol. Model. **198**(1), 115–126 (2006). https://doi.org/10.1016/j.ecolmodel.2006.04.023
10. Grimm, V., Berger, U., DeAngelis, D.L., Polhill, J.G., Giske, J., Railsback, S.F.: The ODD protocol: a review and first update. Ecol. Model. **221**(23), 2760–2768 (2010). https://doi.org/10.1016/j.ecolmodel.2010.08.019
11. Le, T.V., Ukkusuri, S.V.: Crowd-Shipping Services for Last Mile Delivery: Analysis from Survey Data in Two Countries (2018). arXiv:1810.02856, http://arxiv.org/abs/1810.02856
12. Mittal, A., Oran Gibson, N., Krejci, C.C., Marusak, A.A.: Crowd-shipping for urban food rescue logistics. Int. J. Phys. Distrib. Logist. Manag. **51**(5), 486–507 (2021). https://doi.org/10.1108/IJPDLM-01-2020-0001
13. Pee, L.G., Koh, E., Goh, M.: Trait motivations of crowdsourcing and task choice: a distal-proximal perspective. Int. J. Inf. Manage. **40**, 28–41 (2018). https://doi.org/10.1016/j.ijinfomgt.2018.01.008
14. Rougès, J.-F., Montreuil, B.: Crowdsourcing delivery: New interconnected business models to reinvent delivery (2014), p. 19
15. Zou, G., Gil, A., Tharayil, M.: An agent-based model for crowdsourcing systems. In: Proceedings of the Winter Simulation Conference (2014), pp. 407–418. https://doi.org/10.1109/WSC.2014.7019907

Introducing Land Constraints to Macroeconomic Agent-Based Models

Jacob Kelter, Uri Wilensky, and Joseph Potvin

Abstract We present a macroeconomic agent-based model that incorporates land to constrain maximum agricultural production. The model contains three types of agents: firms, households, and land plots. Firms employ households to produce consumer goods which household then buy from firms. The maximum production of firms is limited by the area of land available to them. Varying the availability of land and investigating the results on unemployment, wage rates, firm turnover rates, profits, and inequality among households and firms shows that the interests of firms and households conflict regarding the ideal amount of land available for production.

1 Introduction

Agent-based models (ABMs) can reproduce many stylized facts of the economy including business cycles, endogenous fluctuations and correlations in macro-level economic variables, firm-size distributions and more [1–3]. Increasingly over the past three decades, ABM literature has addressed the intersection of economics and the environment, including energy technology markets [4, 5], and climate-economy interactions [6, 7]. There have also been sector-oriented investigations of agriculture and land-use practices and their interactions with the economy [8, 9]. However, these models are oriented to a specific sector. Macroeconomic ABMs have not inte-

J. Kelter (✉) · U. Wilensky
Northwestern University, 633 Clark St, Evanston, IL 60208, USA
e-mail: jacobkelter@u.northwestern.edu
URL: https://xalgorithms.org

U. Wilensky
e-mail: uri@northwestern.edu

J. Potvin
Xalgorithms Foundation, Ottaw, ON, Canada
e-mail: jpotvin@xalgorithms.org

Z. Yang and S. Núñez-Corrales (eds.), *Proceedings of the 2022 Conference of The Computational Social Science Society of the Americas*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-031-37553-8_3

grated “Land” as a primary factor of production along with the other three factors emphasized by classical economics: “Labor”, “Capital” and “Organization” [10].

The ABM presented in this paper lays the groundwork for investigating feedback loops between the economy and ecological integrity of land. This is part of an ongoing modeling project of wider scope, but the aspect of the model presented here focuses only on how availability of productive Land impacts the economy.

2 The Model

2.1 Model Structure

The model,[1] implemented in the NetLogo agent-based modeling environment [11], contains three classes of agents: firms, households, and land. Figure 1 represents the relationships between the different types of agents. We take a pure agent-based approach—as opposed to representing labor and consumption at an aggregate level as in the Dosi et al. [3, 12] family of models—so that we can explore the effects of various parameters on wealth distribution and the effects of various frictions in the labor market on the overall economy.

Firms produce homogeneous consumer goods on plots of land and sell them to households. A firm’s output is constrained by the area of land available to it and the land’s productive capacity per unit area. In the experiments reported here, the productive capacity of land per unit area is held fixed across all experiments and the area of land available is held fixed within experiments but varied across them. In future work, the activity of a firm will increase or decrease the land’s productive capacity over time depending on the quality of its land management practices. This is depicted in Fig. 1 with a grey dotted line since it is part of the model design but not used in this preliminary paper.

The model is initialized with a set of households, firms, and plots of land. Each firm is associated with a single plot of land. We abstract from land markets, and therefore land is not bought or sold. Households begin with an initial amount of currency liquidity, they each supply their Labor to a single firm, and they have links to several firms as consumers. Firms raise initial start-up funds from one or more households as investors (described in more detail below). To keep the model simple, there is no banking sector, finance, or credit in the model. Firm entry is funded directly by households rather than through banks as in the Delli Gatti et al. [2] family of models. The present model holds money supply fixed, resulting in zero inflation; a subsequent version of the model will have a dynamic monetary system.

[1] Model code available at https://github.com/jzkelter/tabular-standards/wiki/How-do-download-and-use-the-model.

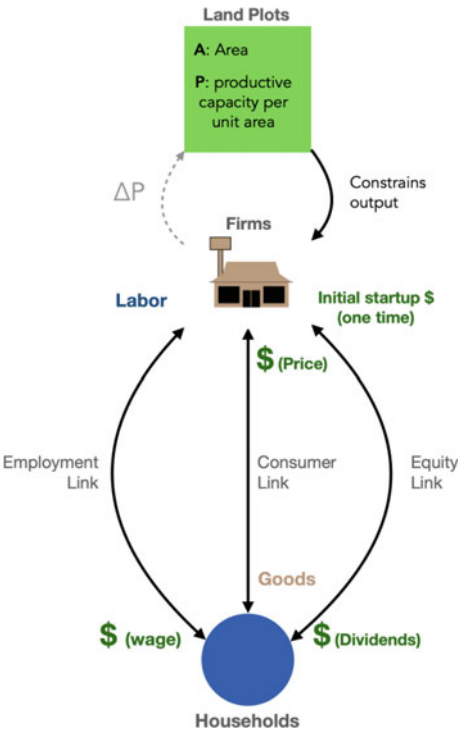


Fig. 1 The structure of the model. Firms each have a plot of land. Households and firms have three types of relationships: employment, consumption, and equity (investment)

2.2 Sequence of Events

The model uses discrete time steps, and the sequence of events in each period is as follows:

1. At the start of the period, firms without enough liquidity to pay a single worker go bankrupt and are replaced.
2. Firms, to plan the upcoming period:
   (a) Estimate demand in the upcoming period
   (b) Adjust wage rates based on success/failure in filling job positions
   (c) Plan output based on previous sales (constrained by land) and adjust price
   (d) Adjust desired labor (lay off workers or posts job openings)
3. Households:
   (a) Update consumer links
   (b) Adjust reservation wage
   (c) Search for employment

4. Firms:
   (a) Pay wages (happens at the beginning of the month to simplify firm planning and after households have searched for employment so new hires get paid)
   (b) Distribute profits from prior period (this can only happen after paying wages)
5. Households set consumption for the month
6. Firms produce output goods
7. Households buy and consume goods.

Each of these events is described in more detail below. In most cases, when describing procedures that involve parameters, we use the parameter name from the model's code directly instead of translating the parameter names from the code to mathematical symbols, which would then need to be translated back again by anyone who reads the code. To further ease reading, we usually describe the model using the concrete values for parameters used in simulations and footnote what the relevant parameter name is in the code. We do use mathematical notation in select cases when we think it facilitates communication. For full specification of the model, see the online supplemental documentation[2] based on the Dahlem guidelines [13].

2.3 Bankruptcy and Firm Replacement (Step 1)

At the beginning of each period, if a firm does not have enough liquidity to pay a single worker, it goes bankrupt. Any remaining liquidity it has is returned to shareholders. A new firm is then created which raises start-up capital from households.[3] The new firm asks households in a random order for funds. Households are willing to invest up to half of their current liquidity in the new firm, but their investment is only accepted if it represents at least 10% of the total value the firm is raising. In this way, there is an emergent class of capitalists based on wealth rather than a hard-coded class of capitalists as in [1]. Households own a fraction of the firm in proportion to what fraction of the startup funds they provided. The firm is initialized with 20 consumer-links,[4] otherwise, it would usually fail to sell anything the first period and immediately go out of business. These initial consumer links can be thought of as being due to an initial advertising campaign. The new firm is also automatically given the plot of the land that the bankrupt firm vacated.

[2] https://github.com/jzkelter/tabular-standards/blob/main/Main%20Model/Dahlem_Description.md.

[3] The amount is equal to the parameter STARTUP-LIQUIDITY.

[4] INITIAL-CONSUMER-LINKS.

2.4 Firms Estimate Demand (Step 2a)

Firms estimate demand based on a rolling average of past sales, $S_{ave}(t)$ which is updated by the sales of the previous period by the following equation:

$$S_{ave}(t) = mS_{ave}(t-1) + (1-m)s(t)$$

where $s(t)$ is sales in the time period, t, that just passed and m is a "memory" constant[5] between 0 and 1 that determines how much the firm remembers/weights previous average sales compared to the prior period's sales.

2.5 Firms Adjust Wage Rates (Step 2b)

Firms adjust wages based on their success or failure in hiring. If a firm wanted to hire a worker last month and failed, it increases its wage to attract workers. On the other hand, if a firm has had no vacancies for the past 12 months,[6] the firm decreases wages. In either case, the increase/decrease is by a random fraction chosen uniformly between 0 and 20%.[7]

2.6 Firms Adjust Planned Output and Price (Step 2c)

Ideally, firms want to fully satisfy their expected demand. Since demand may exceed expectations, firms try to keep a buffer stock of 50% of expected demand.[8] So, after production and prior to selling goods, firms aim to have 150% of expected demand in stock. Goods are non-perishable in the model. This means that if a firm has already built up its buffer, it rarely has to produce much more than expected demand. Firms may not be able to produce enough to have 150% of expected demand in stock before sales begin due to limited land, liquidity, or failure to hire adequate workers. As this is the planning stage, only the first two limitations come into play, and they determine how many workers the firm will aim to have this period. The number of workers a firm desires is equal to target production divided by the "tech-parameter" which determines labor productivity. This assumes that production is a linear function of labor (no changing returns to scale). For the purpose of this paper, tech-parameter is uniform across firms and held constant (i.e., there is no technological innovation). The pseudo-code in Algorithm 1 describes the process of firms to plan output and their desired number of workers.

[5] FIRM-MEMORY-CONSTANT.

[6] MONTHS-TO-LOWER-WAGE.

[7] MAX-WAGE-CHANGE.

[8] DESIRED-BUFFER-FRAC.

Algorithm 1 Firm process to plan output

```
set target_production = 1.5 * expected_demand - current_inventory
if target_production > total productive capacity of land then
    set target_production = total productive capacity of land
end if
set target_n_workers = target_production / tech_parameter
if liquidity < target_n_workers * wage-rate then
    reduce target_n_workers to maximum that can be afforded given liquidity
end if
```

After planned output has been decided, the firm adjusts its price. Following [2], a firm will not increase both output and price. A firm will raise prices only if all three of the following conditions are met:

1. Demand was higher than expected last period (which means expected demand this period is higher than last period)
2. The firm is unable to satisfy expected demand this period (either due to lack of liquidity to hire workers, or due to reaching the maximum productive capacity of the land).
3. The firm's price is less than the average price of other firms

The rationale for these conditions is that firms aim to increase market share before increasing unit profits. If condition 1 is met but not 2, this means the firm will try to meet the increased expected demand at the current price. If conditions 1 and 2 are both true but not 3, the firm will not risk losing market share by raising prices further above the average price of other firms.

A firm will decrease price if the following three conditions are met:

1. Demand was significantly less than expected last period, as measured by inventory being 120%[9] or more of the ideal buffer amount
2. The firm has enough liquidity to meet expected demand this period
3. The firm's price is more than the average price of other firms.

Condition 1 guarantees there is surplus. Condition 2 checks that the firm is able to fulfill expected demand, which suggests there will probably still be surplus. If this is true and the firm's price is above the current average, the firm decreases price to try to gain market share.

In the case of either raising or lowering price, the firm increases or decreases its price by a random percentage between 0 and 20%.[10]

[9] BUFFER%-TO-LOWER-PRICE.

[10] MAX-PRICE-CHANGE.

2.7 Firms Adjust Labor (Step 2d)

If a firm has fewer workers than desired, it automatically has a job opening(s) available. It is then left to households searching for jobs to find these firms. If a firm has more workers than desired, it will attempt to lay off workers. Rather than keep track of labor contract lengths for each worker, we instead allow firms to lay off workers probabilistically. The firm attempts to lay off each worker it does not want and succeeds with probability equal to the parameter LAYOFF-PROBABILITY. A low layoff probability is equivalent to long labor contracts and a high layoff probability is equivalent to short labor contracts.

In addition, a firm that cannot afford to pay its current number of workers lays off as many workers as needed so that it will be able to afford the wage bill.

2.8 Households Update Consumer Links (Step 3a)

Households have 7 consumer links.[11] If they have fewer than this (due to a firm going out of business) they create new trading links. If a household has more than 7 consumer links, it randomly deletes one.

After guaranteeing they have enough consumer links, households probabilistically search for more desirable trading links. With a 25% probability,[12] the household will pick a random firm and, if its price is cheaper than the household's most expensive current consumer link, will delete its most expensive consumer link and create one with the cheaper firm. If a household has a consumer link with a firm that failed to satisfy its demand last period, with a 50% probability[13] the household will replace it with the randomly selected firm. In both cases, randomly selected firms are chosen from the set of firms the household does not currently have a consumer link with and their chance of selection is weighted by the number of consumer links they have (larger firms are more likely to be selected).

2.9 Households Adjust Reservation Wage (Step 3b)

A household's reservation wage is the minimum wage it is willing to accept for employment. If a household is unemployed, it decreases its current reservation wage to 90%[14] of its current value. If the household is employed and its current wage is above its reservation wage, it increases its reservation wage to equal its current wage.

---

[11] N-TRADING-LINKS.

[12] PROB-REPLACE-FIRM-PRICE.

[13] PROB-REPLACE-FIRM-QUANT.

[14] RES-WAGE-CHANGE.

2.10 Households Search for Employment (Step 3c)

An unemployed household checks five[15] randomly chosen firms for job openings and takes a job with the first one that offers a wage above the household's reservation wage. An employed household will check one random firm for a better paying job if its wage is below its reservation wage or with probability equal to 10%.[16] If the randomly chosen firm has a job opening at a better wage, the household switches jobs.

2.11 Firms Pay Wages and Distribute Profits (Step 4)

At this point, all employment is set for the month and firms pay their workers at their current wages. In case sales are lower than expected, firms keep some liquidity in reserve equal to 30% of current labor costs. Whatever liquidity remains is distributed to households with equity in the firm in proportion to their equity.

2.12 Households Set Consumption for the Month (Step 5)

Households, having been paid, set their consumption for the month based on the equation:

$$C = L^{\alpha}$$

where C is planned consumption, L is the household's current liquidity, and α is a parameter[17] determining diminishing marginal utility of consumption. Based on this equation, consumption always increases with increased liquidity, but unless $\alpha = 1$, the increase is sub-linear. The simulations in this paper use $\alpha = 0.6$.

2.13 Firms Produce Output Goods (Step 6)

Firms produce output based on the number of workers they have and land constraints. Output is linear with the number of workers (output = n-workers * tech-parameter) up until the total productive capacity of the land which equals land area times productive capacity per area. This is equivalent to an assumption that each worker can work a certain area of land and there is no benefit from additional labor applied to the land.

---

[15] SEARCH-N.

[16] SEARCH-BETTER-JOB-PROB.

[17] DIMINISHING-UTILITY-CONSTANT.

We assume productive capacity per area of 1 such that the total productive capacity is equivalent to the area of land available.

2.14 Households Buy and Consume Goods (Step 7)

Households visit the firms they have consumer links with one at a time in a random order[18] and attempt to satisfy their demand by buying from that firm. If the firm runs out of inventory, the household visits the next consumer firm up until it either satisfies its demand or runs out of consumer links.

3 Simulations

3.1 Simulation Setup

The simulations were designed to explore the effect of altering the availability of land on the aggregate economy. The main variable varied is available land area per capita,[19] denoted here as A_{pc}. It is expressed per capita (per household) to make simulations with different numbers of households comparable. Each run can be thought of as a parallel universe with more or less available land. In the real world, available land per capita can vary due to population growth/decline, ecological degradation/restoration, or changes in land-use.

Productivity of land per unit area was assumed to be 1. The tech-parameter was chosen to be 1 as well—meaning that one unit of labor can work one unit of land and produce one unit of output per period. So, when $A_{pc} = 1$ all households could theoretically be productively employed. When $A_{pc} < 1$, the total amount of land is less than the overall labor capacity. A_{pc} was varied between 0.2 and 10 with varying increments[20] depending on how close the value of A_{pc} was to the critical value of 1. The simulations were run with 1000 households and 60 firms for a total of 500 periods after a 500 period "burn-in" time to allow the model to reach a steady state. Based on visual inspection, the model reaches a steady state after around 200 periods; a 500-period burn-in was chosen to safely reach the model's steady state. The number of households in the model is arbitrary but similar to prior macroeconomic ABM research [2, 14]. The ratio of firms to household is chosen to roughly match a typical firm in the United States. In the United States, 85% of employers have 19 or fewer employees [15], and the average firm in the model will have around 15 workers

[18] PICK-CHEAPEST-FIRM? = false. If households do sort firms based on price, results are similar in most parameter settings. A full discussion of the impact of households sorting firm by price is beyond the scope of this paper.

[19] LAND-AREA-PER-CAPITA.

[20] All values used are: [0.2 0.4 0.6 0.8 1 1.1 1.2 1.3 1.4 1.6 1.8 2 3 4 6 8 10].

when unemployment is low. Of course, in the real economy there is a small number of massive employers, but this is ignored in the model. The appendix lists all the parameters used in the simulations and their values.

3.2 Results

To give an initial sense of the output of the model, Fig. 2 shows time series of unemployment, mean wage, and mean profits for $A_{pc} = 1$ (blue) and $A_{pc} = 3$ (orange). Wages and profits are expressed as a fraction of average consumer good price because monetary units in the model are arbitrary. Dividing by the price of goods gives a measure of actual purchasing power. The time series show 500 periods of model execution starting after the 500 period "burn-in" to allow the model to stabilize. For both values of A_{pc}, all three variables fluctuate around a mean. When $A_{pc} = 1$, unemployment is high, wages are low, and profits are high. When $A_{pc} = 3$, unemployment is low, wages are high, and profits are low. We now turn to a systematic analysis of how changing A_{pc} affects households and firms according to several important variables.

Increasing A_{pc} benefits both household and firms up to around $A_{pc} = 1$. After that, increasing A_{pc} favors households over the interests of firms. When A_{pc} is low, unemployment is very high, as can be seen in Fig. 3a. High unemployment correlates with low wages as seen in Fig. 3b, and inequality between households is very high as seen in the graph of the Gini coefficient in Fig. 3e (blue line). The extremely high inequality when A_{pc} is low results from three factors: (1) many households are unemployed, (2) even among those who are employed, wages are very low, and (3) a very small number of households end up owning the majority of the equity in firms and therefore collect the vast majority of profits. As A_{pc} increases from below 1 to slightly above 1, unemployment drops rapidly and then slowly rises again with higher A_{pc}, leveling off at around 13%. This slow increase in unemployment with higher A_{pc} is likely due to more firms going out of business, as discussed in the next paragraph. Wage rate increases rapidly with increasing A_{pc}, leveling off at around 0.75. This means that under conditions of high A_{pc}, households are paid about 75% of what they produce, the remaining 25% being kept as profits. Due to relatively low unemployment and higher wages, the Gini coefficient for households drops with increasing A_{pc} and levels off around 0.42. To summarize, as A_{pc} increases, unemployment decreases rapidly and then slowly increases, wages increase, and household inequality decreases.

The story is different for firms. They also benefit as A_{pc} increases from below 1 to slightly above 1 as seen in the decreasing turnover rate (Fig. 3c) and increasing profits (Fig. 3d). However, as A_{pc} continues to increase, turnover rates increase and profits decrease. As a result, inequality among firms increases with increasing A_{pc} (Fig. 3e). The reason for these patterns is that when A_{pc} is low, even if demand increases, a firm cannot produce more due to the limited amount of land. So, a firm will not hire workers away from other firms, and competition between firms remains low. As A_{pc} increases, firms can produce more and therefore compete more for market

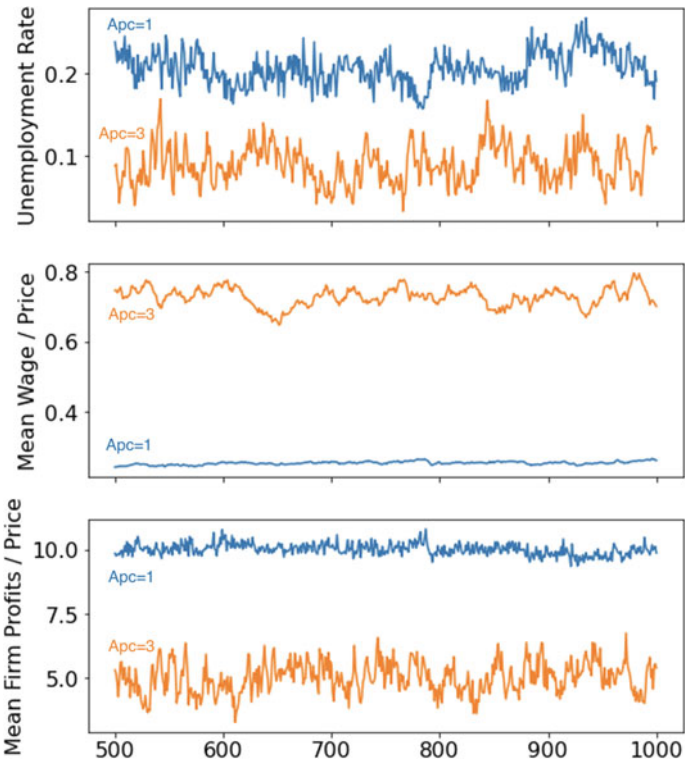


Fig. 2 Time series of unemployment, mean wage rate, and mean profits for two values of A_{pc}

share. Such competition decreases their rates of profit and increases turnover rate. To summarize, firm profits and turnover rate improve as A_{pc} increases to around 1 when all households can be productively employed but as A_{pc} increases further, profits sag and turnover increases.

4 Conclusion

The results of the model show a basic conflict between the interests of households and firms regarding the availability of productive land. When there is a surplus of labor compared to available land, wages are suppressed, and unemployment is high. When labor and land availability approximately balance, unemployment is low, but wages are low as well. This is because firms have no incentive to compete for workers since they do not have extra land with which to increase production. Such a situation favors firms, resulting in low turnover rates and high profits. This results in low inequality between firms but extreme inequality among households, because over

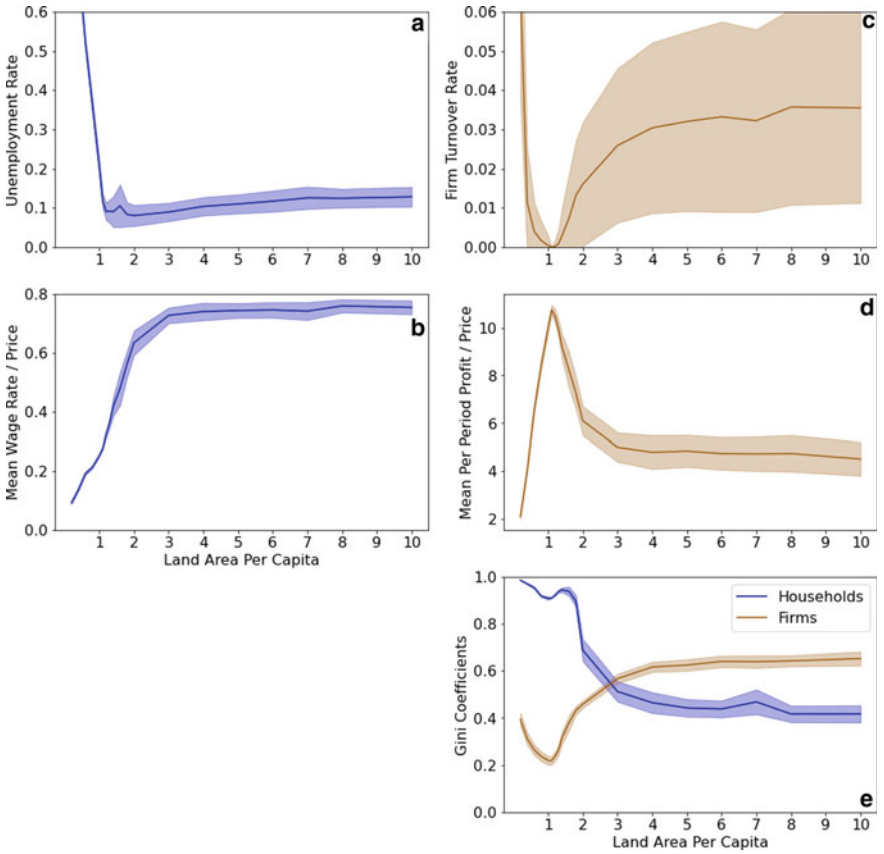


Fig. 3 The effect of changing A_{pc} on both households and firms. The solid lines represent the mean and the shaded area one standard deviation. Households can each work one unit of land per period, so $A_{pc} = 1$ means that, in theory, all households could be productively employed. Wages and Profits are expressed as a fraction of average price of goods to give a measure of actual purchasing power

time a small number of households come to own the large majority of firm equity and wages for remaining households are low. As the available land increases, firms have an incentive to increase production to compete for market share, leading them to compete for workers which drives up wages. Inequality among households decreases while inequality and turnover among firms increases.

The model presented here fills a gap in extant macroeconomic ABMs by connecting the productive capacity of the economy to the underlying productive capacity of the Earth. As such, it lays the foundation for a number of future dynamic modeling interactions between the economy and ecosystems. The immediate next step is to model firms impacting the productivity of the land due to either degenerative or regenerative farming practices as discussed in Sect. 2.1. Beyond this, we plan to model more complex economies with a supply network of multiple types of firms and

introduce indexed pricing schemes designed to stabilize economic fluctuations and align free market incentives with sustained ecological integrity as a passive emergent effect.

Acknowledgements Thank you to Jacob Wit and William Conboy for their contributions to the code of the ABM presented in this paper.

Appendix—Simulation Parameters

The following is a complete list of the parameters used in the simulations reported in this paper as they would appear in a NetLogo Behavior Space experiment:

```
["setup-structure" "Single-PG&CG-TC=1.json"]
["LAND-AREA-PER-CAPITA" 0.2 0.4 0.6 0.8 1 1.1 1.2 1.3 1.4 1.6 1.8 2 3 4 6 8 10]
["MIN-WAGE-RATE" 2.5]
["DIMINISHING-UTILITY-CONSTANT" 0.6]
["pick-cheapest-firm?" false]
["delli-gatti-consumer-search?" false]
["N-TRADING-LINKS" 7]
["MONTHS-TO-LOWER-WAGE" 12]
["layoff-probability" 0.5]
["n-households" 1000]
["n-firms" 60]
["SEARCH-N" 5]
["fix-n-framework-agreements?" false]
["PROB-REPLACE-FIRM-PRICE" 0.25]
["framework-duration" 24]
["MAX-PRICE-CHANGE" 0.2]
["index-in-use" "no index"]
["SEARCH-BETTER-JOB-PROB" 0.1]
["mean-new-agreements-per-month" 2]
["firm-memory-constant" 0.8]
["min-wage-80\%-of-tech-param?" false]
["STARTUP-LIQUIDITY" 100]
["primary-good-prod-function" "linear"]
["alpha" 1]
["transactions-per-month" 1]
["DESIRED-BUFFER-FRAC" 0.5]
["RES-WAGE-CHANGE" 0.9]
["N-FRAMEWORK-AGREEMENTS" 7]
["s" 0.1]
["BUFFER-LABOR-FRACTION" 0.3]
["firm-competency" 0]
["MAX-WAGE-CHANGE" 0.2]
["PROB-REPLACE-FIRM-QUANT" 0.5]
["BACKGROUND-IMPROVEMENT" "10"]
```

References

1. Assenza, T., Delli Gatti, D., Grazzini, J.: Emergent dynamics of a macroeconomic agent based model with capital and credit. J. Econ. Dyn. Control. **50**, 5–28. issn: 01651889 (2015)
2. Delli Gatti, D., Desiderio, S., Gaffeo, E., Cirillo, P., Gallegati, M.: Macroeconomics from the Bottom-Up. Springer Science & Business Media (2011). isbn: 978-88-470-1971-3
3. Dosi, G., Fagiolo, G., Roventini, A.: An evolutionary model of endogenous business cycles. Comput Econ **27**, 3–34 (2006). issn: 0927-7099, 1572-9974
4. Hesselink, L.X.W., Chappin, E.J.L.: Adoption of energy efficient technologies by households—barriers, policies and agent-based modelling studies. Renew. Sustain. Energy Rev. **99**, 29–41 (2019). issn: 1364- 0321
5. Rai, V., Robinson, S.A.: Agent-based modeling of energy technology adoption: empirical integration of social, behavioral, economic, and environmental factors. Environ. Model. Softw. **70**, 163–177 (2015). issn: 1364-8152
6. Balint, T. et al.: Complexity and the economics of climate change: a survey and a look forward. Ecol. Econ. **138**, 252–265 (2017). issn: 0921-8009
7. Ciarli, T., Savona, M.: Modelling the evolution of economic structure and climate change: a review. Ecol. Econ. **158**, 51–64 (2019). issn: 0921-8009
8. Coronese, M., Occelli, M., Lamperti, F., Roventini, A.: AgriLOVE: Agriculture, Land-Use and Technical Change in an Evolutionary. Agent-Based Model SSRN Scholarly Paper. Rochester, NY (2021)
9. Shahpari, S., Allison, J., Harrison, M.T., Stanley, R.: An integrated economic, environmental and social approach to agricultural land-use planning. Land **10**, 364 (2021). issn: 2073-445X
10. Marshall, A.: Principles of Economics Eighth. Macmillan, London, [1979] (1920). isbn: 978-1-343-83441-5
11. Wilensky, U.: NetLogo Center for Connected Learning and Computer-Based Modeling. Northwestern University, Evanston, IL (1999)
12. Dosi, G., Fagiolo, G., Roventini, A.: The microfoundations of business cycles: an evolutionary, multi-agent model. J. Evol. Econ. **18**, 413–432 (2008). issn: 1432-1386
13. Wolf, S., et al.: Describing economic agent-based models – dahlem abm documentation guidelines. Complex. Econ. **2**, 63–74 (2013). issn: 2210- 4275, 2210-4283
14. Lengnick, M.: Agent-based macroeconomics: a baseline model. J. Econ. Behav. Organ. **86**, 102–120 (2013). issn: 01672681
15. Hait, A.W.: The Majority of U.S. Businesses Have Fewer Than Five Employees (2021). https://www.census.gov/library/stories/2021/01/what-is-a-small-business.html

Interactions, Model Mechanisms and Behavioral Attractors in Complex Social Systems

H Van Dyke Parunak and Santiago Núñez-Corrales

Abstract In social modeling, a *computational environment* runs a *model* that represents the *world*. The states the model explores (its *behavioral attractor*) are typically fewer than its description suggests. The mapping between model and attractor depends not only on its *parameters* (exploring variants of the world) and its *conventions* (imposed by the computing environment), but also its *mechanisms* (components of the model representing selected dimensions of the world). This paper equates *mechanisms* with sets of coupled *interaction classes*, thus connecting the relative richness of possible choices of agent behaviors to the size of the state space sampled by computational procedures. We illustrate the impact of different mechanisms on the attractor with a specific simulation platform, SCAMP. In our case, in general, the more mechanisms one implements, the smaller the attractor, but with unexpected twists. We discuss the implications of the richness of the corresponding repertoire of interactions available to agents during simulation for the apparent combinatorial explosion of future possible states in agent collectives. We finally observe how some of these twists appear to correspond with the existence of constraints, hinting at underlying conservation laws in silico and ideally in real systems these intend to portray.

1 Introduction

The user of an agent-based social model (ABM) is largely occupied with imputing observed behaviors to underlying *social mechanisms*, either internal to individual agents or collective, and interrogating social models drives most of the research pragmatics. How many distinct behaviors can agents manifest? How does their spatial

H. V. D. Parunak (✉)
Parallax Advanced Research, Beavercreek, OH 45431, USA
e-mail: van.parunak@parallaxresearch.org

S. Núñez-Corrales
National Center for Supercomputing Applications, University of Illinois, Champaign, IL 61801, USA

Z. Yang and S. Núñez-Corrales (eds.), *Proceedings of the 2022 Conference of The Computational Social Science Society of the Americas*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-031-37553-8_4

distribution vary over time? Can we relate changes in the number and classes of behaviors to corresponding changes in the resulting macroscale patterns of organization? These questions translate to the underlying search for general mechanisms connecting behaviors to classes of interactions, and then for more fundamental principles that constrain their action. This article discusses this relation explicitly using a concrete example.

2 Behaviors, Mechanisms and Interactions

The primary purpose of an ABM is enacting mechanisms and observing micro- and macroscale consequences across scales, with an emphasis on systems observed or expected to display emergent properties. Mechanisms that operate at the microscale often relate to intensive quantities in a system, those pertaining to individual agents, while extensive quantities characterize the entire system. Traces of these quantities reflect behaviors, which can be described as modes of operation of a system, or of its parts, whose composition and aggregation manifest as temporal patterns. For instance, in organization theory, organizational routines play the role of behaviors that explain flexibility and change within organizations [8]. Combinatorially, the more behaviors per actor an ABM captures, the more possible worlds it formally describes. However, the range of possible worlds *actually* visited (the model's *behavioral attractor*) is usually much smaller than the static model suggests.

The mapping between model and attractor depends on *parameters*, *conventions*, and *mechanisms*. Each of these describes a different component of the modeling enterprise, in which a *computational environment* runs a *model* that represents the *world* (Fig. 1). *Parameters* capture how varying the architecture of the world being simulated changes its static and dynamical shape, *conventions* dictate how the abstract model executes on a computer to generate an attractor, and *mechanisms* are the abstract building blocks driving the time evolution resulting in behaviors that leave traces. Therefore, behaviors are simultaneously *composite* and *composable*, suggesting that the size of the behavioral attractor depends on some small set of primitives at or below the level of mechanisms. An example is the attempt to connect sequence data *efficiently* to the structure, function and history of proteins to a simple amino acid alphabet simultaneously approximate and informative, and with low computational cost compared to protein folding calculations [9].

ABM models similarly attempt to reflect and explain reality with a limited repertoire of mechanistic building blocks. This agenda assumes that a model with fewer mechanisms than the world's facets can still give useful information. Most modeling frameworks offer few alternative mechanisms, seducing modelers to ignore the impact of mechanism choice at the expense of either hiding or constraining both simulation relevance and fidelity in favor of intellectual tractability of the resulting model traces. New perspectives, methods and tools are needed in the ABM community to overcome difficulties originating in the intrinsic combinatorics involved in

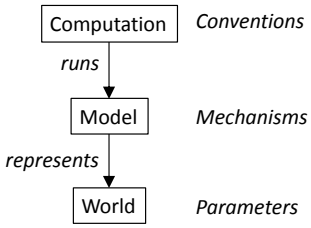


Fig. 1 Sources of variability in an ABM

the analysis of complex systems, particularly when agents have several mechanisms at their disposal.

Agents can be endowed with multiple mechanisms. SCAMP (Social Causality using Agents with Multiple Perspectives) [20], for example, is an ABM framework offering agents multiple independent mechanisms. ABM frameworks generally describe mechanisms in three main ways: interactions of an agent with its internal state, interactions between agents and their environment, and interactions among agents themselves. Furthermore, interaction patterns yield recognizable spatial, temporal or causal structures. We define an interaction as a localized event where two agents exchange information, degrees of freedom, in ways that modify their internal states, spanning a joint ensemble of possible future states depending on the agents' local contexts, and for which knowledge about exact details is constrained collectively and individually by intrinsic and combinatorially originating uncertainties [17].

We describe mechanisms as collections of similar interactions regardless of where and when they occur, suggesting the existence of *interaction classes*. Thus, the dynamical description of a model is the collection of interaction classes present at any given moment. As a system evolves in time, the number of its interaction classes determines the expected number of possible mechanisms and behaviors, which we call the system's *abstract state potential*. However, not all possible interactions are realized in any given system, whether in the real world or in a model.

Interactions at one scale emerge from laws governing more fundamental scales, posing effective dynamical constraints that reshape emergent properties generated by interactions under special circumstances. Thus, the combinatorial explosion produced by introducing more mechanisms, i.e., more tightly coupled sets of interaction classes, is simultaneously pruned by the hard boundaries provided by governing laws, some of which can be stated in the form of conservation laws.

For example, ABM simulations are often constrained by an abstract form of conservation of mass. Consider the minimum number of copper atoms required for conductivity to arise ($\sim 10^4$) versus the far fewer individuals required for proto-institutions to emerge [14]. Conductivity, produced by the cooperative effect of electron holes in crystalline lattices, is constrained by conservation laws via Maxwell's equations. Proto-institutions, produced by the cooperative effect of actors aligning their goals and needs, are constrained by social norms and individual human behaviors. Recent work [15] found an (exponentially) inverse relation between the number

of entities in a system and the number of interaction classes needed for new emergent laws to appear at the next thermodynamic scale of the system.

Conservation laws can illumine our understanding of the behavioral attractor of a model. They are tied to analogues of energy for a given ABM, suggesting some form of the principle of least action. In the presence of stochastic events within complex systems, this entails the existence of a multiplicity of paths to reach an arbitrary state, with at least one path that is optimal in some sense (e.g., minimal distance to move to another location, greatest proximity to satisfying a goal). The latter means that not only is the behavioral attractor itself determined by the number and variety of mechanisms, but the trajectories to reach the attractor are also dynamically constrained. That is, in the language of path integrals, the probability of following a certain path is determined by the form of the interactions, privileging some over others.

To summarize: in general, the more mechanisms in an ABM, the smaller the attractor ("the more you model, the less you see"), but interactions among mechanisms lead to anomalies, or generative effects [1]. Also, by the law of requisite variety [2], we expect the computational machinery required to be significantly more complex. For instance, a more constrained attractor may lie partly outside less constrained ones with the same conventions and parameters. Adding mechanisms can not only sharpen the model's focus, but also shift its location and alter the requirements for emergent behavior to appear in terms of the number of necessary entities. *Parameters* control how interaction classes operate between actors and their environment, *conventions* match interaction classes present in real systems to interaction classes available in the computational machinery, and *mechanisms* link interactions and behaviors.

3 Related and Prior Work

Behavior varies with model *parameters*, widely studied in agent-based systems (e.g., [4, 5, 7, 28]), including studies of tipping points (parameter values where behavior changes discontinuously, leading to a phase shift) and lever points (parameters whose change has a lasting, directed effect) [3, 23]. Wolfram [27] identified four distinct classes of one-dimensional 0–1 nearest-neighbor cellular automata, varying only the update rule, the key model parameter. Verification methods such as sensitivity analysis [10] (p. 24) or comparison of agent trajectories with observed data also explore behavioral changes when parameters change, but not the impact of changing conventions or mechanisms.

Studies of different computational *conventions* are less common, but revealing. For example, a differential equation model and an agent-based model can yield qualitatively different results for the same parameters [24, 25]. Among agent-based models, different scheduling disciplines for entities that in reality execute concurrently lead to different results [11, 13]. An extensive literature discusses scheduler synchrony [16].

This study focuses neither on the *parameters* that vary the world explored by a model nor on the *conventions* imposed by computation, but on differing sets of *mechanisms* that the model uses to represent facets of the world and on their interpretation in the broader context of interactions. Naively, one hopes that even a primitive model will be useful, and that adding more mechanisms will add more detail to the results of the initial model. Unexpectedly, such refinements can also move the focus, and cause other anomalies. This effect has not been explored previously because most modeling frameworks do not offer multiple mechanisms that can be activated independently of one another.

4 Experimental Methodology

SCAMP is a causal language and simulator for social scenarios with multiple mechanisms that can be activated independently of one another. Events and goals in the simulation are represented by networks. A *causal event graph* (CEG) is a directed graph whose nodes represent types of events in which agents can participate, and whose agency edges show allowed agent movement from one event type to another. A *hierarchical goal network* (HGN) is a directed acyclic graph that models the goals of a group of agents and how those goals are related to participation on events in the CEG. Leaf nodes in the HGN are linked, or *zipped,* to event nodes that either support or block them [6, 20, 21] give further details.

SCAMP uses polyagents [18], representing each domain entity by a single *avatar* that deploys a swarm of *ghosts*. The ghosts explore their avatar's possible next choices by looking ahead a fixed distance. At each step, they choose probabilistically among the nodes in the CEG that are immediate successors to their current node, and increment a variable on the node proportional to the value of the position reach. The avatar chooses its next step by choosing probabilistically based on the features deposited by its ghosts. This mechanism simulates the well-documented psychological process of evaluating actions by mental simulation of possible outcomes [12].

From the perspective of a social scientist, SCAMP models high-level mechanisms such as choice influenced by tactical preferences and strategic goals. These mechanisms are composed from lower-level interactions (Fig. 2), and these interactions make SCAMP an ideal platform for exploring the application of a generalized theory of interactions to behavioral dynamics.

We base our experiments on a model of civil strife inspired by recent history in Syria. The CEG in this model includes 460 event nodes with 1106 agency edges and 400 influence edges. The six HGNs, one for each group, include 122 goals or subgoals 77 leaf goals are zipped to 177 event nodes. Our methodology has three parts.

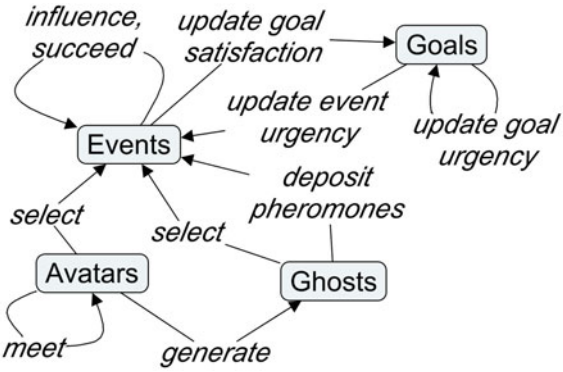


Fig. 2 Every SCAMP *mechanism* is a subnetwork of entity *interactions*

4.1 Define Behavior Space

We interrogate the resulting SCAMP model for *node coverage* and *overlap*. *Node coverage* has three forms: the number of nodes visited by avatars or ghosts, and the number of *successors* considered by ghosts. We measure these values for at least six runs of each configuration. Let Q and R be the sets of nodes for two runs of the same configuration. Then the *overlap* between Q and R is $|Q \cap R|/(|Q| + |R| - |Q \cap R|)$.

We hypothesize that as we add mechanisms, and by extension interaction classes, node coverage in each category will drop (the attractors will shrink) while overlaps will increase, because the system will be attracted into the same region of state space. In terms of the theoretical framework of interactions [15], thanks to repertoire sufficiency fewer agents are needed for anomalies to arise.

4.2 Select Active Mechanisms

SCAMP offers a variety of mechanisms.

Structure of the CEG. The CEG constraints agents' behavioral trajectories. Even for random walks, the branching factors differ along different paths, so that nodes only accessible along highly branched paths will have a lower probability of being sampled in a run of a given length than those with less ramified approaches. In our example, the average node degree in the CEG is 4.74, close to an infinite square lattice, yet highly variable. The kurtosis of node degree is 8.7, reflecting a tail of nodes with high degree.

For comparison, we do a random walk over a rectangular directed lattice of 21 * 22 = 462 nodes, with both ghost and avatar determinism set to 0. A random walk on a regular lattice with restart will visit every node if it runs long enough. We expect the CEG to perform similarly. We also do a random walk over the CEG model itself, augmented with a single START and a single STOP node.

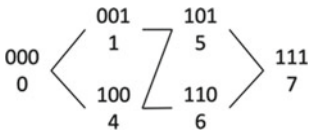


Fig. 3 Configuration lattice: 001 = influences, 010 = HGN, 100 = preferences

Psychological preference. A *feature space* defines agent preferences and event features. Without preferences, ghosts perform a random walk in laying down the presence features that guide avatars. With preferences, ghosts stochastically favor some nodes over others, using a roulette constructed from the features of accessible nodes. We expect (a) agents using preferences will explore fewer nodes than those walking randomly, (b) overlap across runs will be greater with preferences than without, and (c) the longer the model runs, the more nodes will be visited.

Strategic reasoning. Each HGN monitors the recent participation on event types to which it is zipped to assess its current *satisfaction*, then computes the *urgency* feature of each of these events. Agents respond to urgency according to their preferences. If an agent is running without preferences, the HGN is irrelevant. But if preferences are active, we expect HGNs to focus the agents' attention, reducing the number of nodes explored and increasing their overlap.

Influence edges model causal influences among event types between which agents do not move directly, modulating the probability of destination nodes dynamically based on participation levels on source nodes. Again, this mechanism should reduce the number of nodes visited and increase their overlap.

A *configuration* is a binary string indicating active mechanisms. The first position shows whether (1) or not (0) preferences are active. The second position shows HGNs, and the third, the use of influence edges. Thus in 000, the only mechanism is the structure of the CEG, 100 indicates the use of preferences alone, 110 adds HGNs, and 001 is the use of influence edges alone. The decimal values of these strings identify configurations 0 (no mechanisms active) to 7 (all mechanisms active). Configurations 2 and 3 (HGNs without preferences) violate the model and are not included. Configurations 4–7 include preferences, configurations 6 and 7 include HGNs, and odd configurations include influence edges. Our configurations thus form a partial lattice (Fig. 3). All configurations use the same parameters and run with the same conventions.

4.3 Establish Random Baseline

In addition to a space in which the attractor is defined and mechanisms that might impact it, we provide two baselines: L (the 21 * 22 lattice) and R (the CEG), with both ghosts and avatars ignoring the roulette entirely. In configuration 0, unlike R, avatars follow their (randomly moving) ghosts.

5 Results

Our experiments [21] illustrate how studying the behavioral attractor as a function of model mechanisms can confirm or correct our intuitions and highlight behaviors, mechanisms and interactions that invite further study.

5.1 Adding Mechanisms Increases Ghost Selectivity at the Expense of Overlaps

Figure 4 compares the coverage and overlap of avatar visits (av), ghost visits (gv), and successors considered by ghosts (sc) for the baseline configuration (000 ~ 0) and the most constrained (111 ~ 7). Ghosts visit fewer nodes than they consider, and avatars visit fewer than those visited by ghosts. Added mechanisms focus ghosts' attention, as expected, but the number of nodes visited by avatars is unchanged. However broadly or narrowly the ghosts explore, an avatar chooses one path, and in a run of fixed length visits only a limited number of nodes. The avatar nodes are not the same in the two configurations, but the coverage is the same size. We expect overlap to increase with mechanisms, as agents focus their attention on fewer nodes. Figure 4 confirms this intuition for avatar visits, but overlaps for gv and sc actually *decrease*, as discussed in Sect. 5.3.

In a regular directed lattice, coverage would increase with run length. Figure 5 shows the effect of increasing run length from 1000 to 2000 Repast ticks, comparing configuration 0 with 4. In x-axis labels, the first digit (0, 4) is configuration, and the second (1, 2) is run length in k-ticks. Coverage increases for avatar visits, and for gv and sc in configuration 0. But for configuration 4, preferences lead the system to converge, and longer runs do not increase gv or sc. Convergence results from new constraints over interaction classes that preferentially select some behaviors above others.

5.2 Mechanisms Modulate Attractor Size and Location

In Fig. 4, sc and gv are fewer with all mechanisms than with the CEG alone. Figure 6 shows sc for intermediate configurations. Gv shows the same pattern.

In the baselines, random walk on a lattice (configuration L) offers fewer successors to consider (and thus lower gv) than on the CEG (configuration R), reflecting the long tail in the CEG's degree distribution.

Both sc and gv tend to decrease as we add mechanisms. The difference between configurations 0–1 and 4–7 suggests preferences have more effect than HGNs or influence edges. Also, configuration 6 appears to be *lower* than the more highly constrained 7, revealing a realistic interaction between two HGNs and influence

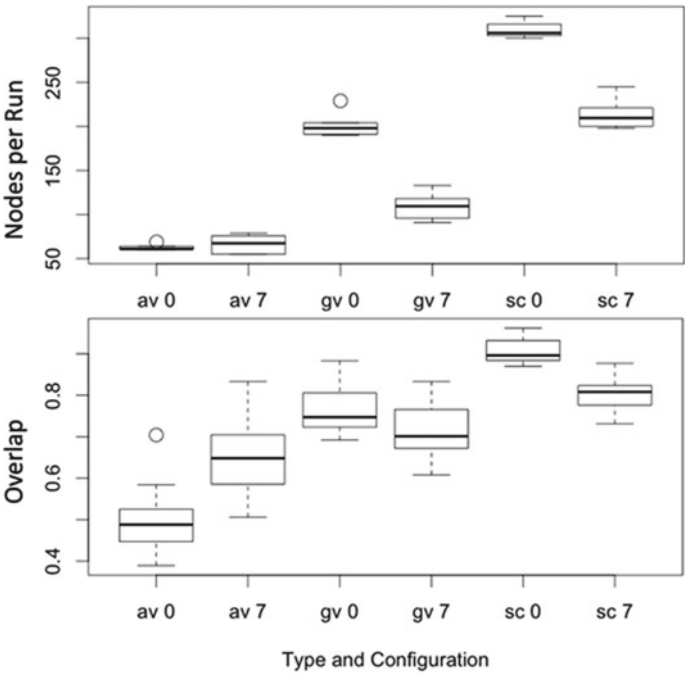


Fig. 4 Nodes visited (Top) and overlaps (Bottom) by types and configurations

Fig. 5 Effect of run length on coverage

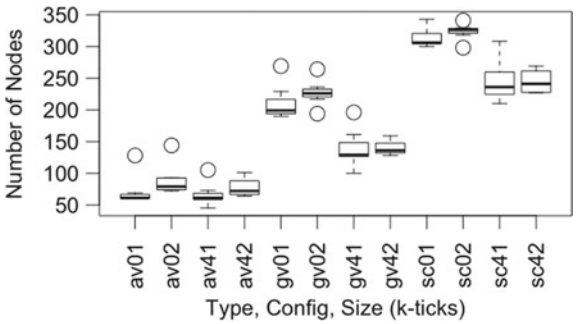
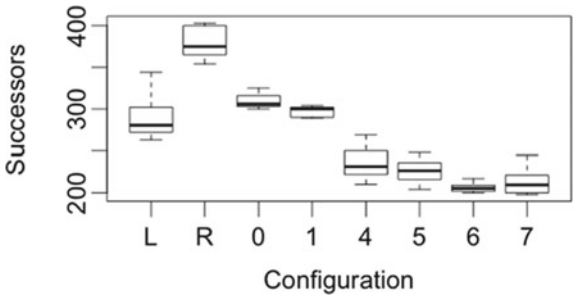


Fig. 6 Successors by configuration



edges. An agent's goals guide its actions by identifying high-priority events in which the agent should participate, and the usefulness of goals will decrease if influence edges block access to those urgent events.

Adding constraints not only decreases attractor size (for gv and sc), but also shifts its location. Define the *alignment* of two configurations as the percent of nodes in the attractor of the more constrained configuration that are also in less constrained configuration. Figure 7 shows the alignment and size ratio for six pairs of configurations, two for each metric. In each case, one pair compares the attractor for configuration 0 with that for configuration 7 (0 → 7), while the other compares configuration 7 with a less constrained configuration (6 → 7 or 1 → 7), but still more constrained than 0. Note:

- As expected, the size ratio is smallest for the greatest increase in constraints (0 → 7). The more constrained the system becomes, the smaller the accessible state space, and the less the attractor shrinks.
- Contrary to expectation, increasing constraints *reduces* the alignment between attractors. This effect is greatest not in moving from 0 to 7, but from intermediate configurations to 7. Importantly, it is greatest for avatars, whose trajectories model physical entities and are most likely to be used for policy recommendations. This result challenges the common assumption that ignoring facets of the real world gives a fuzzier but still essentially correct outcome. In fact, adding mechanisms for these facets can shift the model's output.

In terms of our theoretical framework, under the hood, interaction classes with matching signatures compose instead of aggregate, leading to significantly different behavioral landscapes.

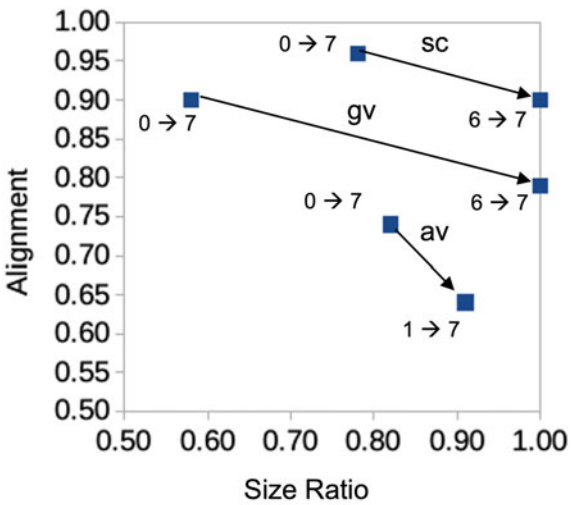


Fig. 7 Alignment versus Size ratio of attractors as constraints increase

5.3 *Increasing Mechanisms Generates Causal Entropic Forces*

In addition to monitoring node coverage (estimating a configuration's attractor), it is also useful to study variation among the nodes visited in different runs of the same configuration. Intuitively, we expect overlap to increase with number of mechanisms. This intuition must be qualified.

With significance p = 2E–16, sc has the most overlap, followed by gv and then av.. This difference probably reflects the fact (Fig. 4.) that for a fixed number of nodes, av < gv < sc. Higher coverage of the CEG leaves fewer nodes on which runs can differ with each other.

Figure 8 shows how av and sc overlaps vary with configuration. Av overlap satisfies our intuition that more mechanisms guide agents into similar regions of the CEG, increasing overlap. Consistent with this dynamic, configurations L and R, where both ghosts and avatars execute random walks, have the lowest overlaps. Configuration 6 yields the highest overlap. Adding influence edges in configuration 7 reduces overlap, reflecting their interaction with HGNs.

Overlaps in sc are more complex. Setting aside L and R, sc overlaps *decrease* with added mechanisms! As with sc and gv coverage, there is a sharp drop with

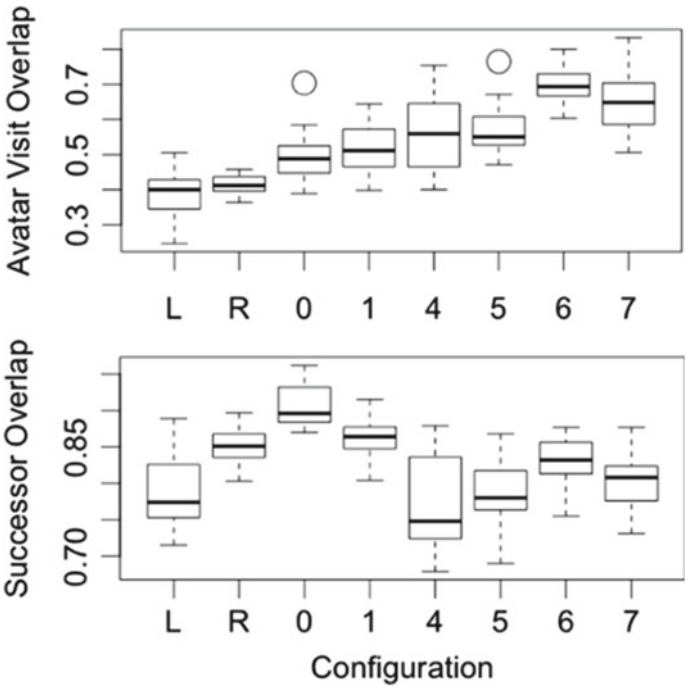


Fig. 8 Overlaps by configuration

configuration 4, when preferences become active. Again, the power of HGNs in drawing agents together is clear in the increased overlap in configuration 6, but faces interference from influence edges in configuration 7.

The overall negative correlation between sc overlap and number of mechanisms is surprising. Perhaps the mechanisms lead the agents into parts of the CEG that they otherwise would not visit. Preferences in particular can lead agents to prefer highly branching regions that otherwise would be relatively inaccessible. In such a region with high node degree, SCAMP's roulette selection can push different runs in different directions, increasing sc coverage and thus reducing overlap. Modelers who assign favorable features to some events may focus more attention on them and ramify the paths to which they lead more than for other events, a form of modeling bias.

Let us reinterpret these results in terms of interaction classes. Avatars outsource future state discovery to ghosts, which explore the feasibility of successors. We propose that interactions between avatars and ghosts, and interactions between ghosts and successors are fundamentally different, and that such difference explains the counterintuitive outcomes observed here. Low overlap for highly constrained configurations relative to unconstrained ones reflects the information gained by increasing the number of mechanisms available to agents, leading in turn to sampling a more diverse behavior space with the same resources (ghosts), since having a larger repertoire of interaction classes lets agents collectively explore new, otherwise inaccessible regions of the behavioral attractor. Interactions between avatars and ghosts appear to be more constrained since ghosts reduce uncertainty and random variation. These interactions appear more information-like in the sense that they abstract details about future paths to keep the computational cost of avatars relatively constant, and their coverage sufficiently high.

The stochastic exploration of ghosts has the signature of an entropic force, a force that arises out of thermodynamic (hence stochastic) systems as they tend to maximize their entropy [22]. Ghosts then filter information to avatars by selecting paths that maximize future freedom of action. This is the definition of causal entropic forces [26], which have been theorized to have a significant role in how we define intelligence.

6 Discussion and Future Work

While our specific results are of great interest to users of SCAMP, our message is important for the responsible use of any ABM framework, in two ways. First, modelers have a sense of the range of possibilities covered by their models, based on the static structure of those models. The attractor visited by the running model may be significantly smaller. Users need to understand a model's coverage under different conditions, and modelers need to understand how adding mechanisms may impact that coverage. Sometimes users will want to increase coverage to consider more possible outcomes; in other cases they will want to decrease it to focus on the most

plausible outcomes. Second, adding mechanisms to capture more dimensions of the real world can not only provide a more focused result, but also shift its location in state space. Third, translating fundamental aspects of the ABM world into the language of interactions and interaction classes clarifies how parameters, conventions and mechanisms intersect at deeper levels of analysis.

This translation leads in several directions.

1. We need a more rigorous exploration of the counterintuitive behaviors described above. This exploration will also help validate the hypotheses we suggest, and is particularly challenging at the start-up of the simulation, identifiable by plotting the entropy of each agent's roulette over time [19].
2. Two theoretical avenues are apparent.

    a. Translating the current model into the generalized theory of interactions can reveal connections between behaviors, mechanisms and interactions with higher intellectual efficiency both in formal and graphical manners.
    b. Since interaction classes for a model are related to the structure of CEGs and HGNs, we are exploring spectral graph methods to untangle interaction motifs connected to the mechanisms described here to make the underlying complexity of SCAMP's results tractable.

References

1. Adam, E.M.: Systems, generativity and interactional effects. Thesis at Massachusetts Institute of Technology, Department of EECS (2017)
2. Ashby, W.R.: Requisite variety and its implications for the control of complex systems. Cursos Congr. Univ. Santiago de Compostela **1**(2), 83–99 (1958)
3. Brueckner, S., Parunak, H.V.D.: Information-driven phase changes in multi-agent coordination. In Proceedings of Workshop on Engineering Self-Organizing Systems (ESOA, at AAMAS 2005), pp. 104–119, Springer, Berlin (2005)
4. Butner, J.E., Wiltshire, T.J., Munion, A.K.: Modeling multi-agent self-organization through the lens of higher order attractor dynamics. Front. Psychol. **8**, 380 (2017)
5. Cenek, M., Dahl, S.K.: Geometry of behavioral spaces: A computational approach to analysis and understanding of agent based models and agent behaviors. Chaos: Interdiscip. J. Nonlinear Sci. **26**(11) (2016)
6. Cox, M.T., Parunak, H.V.D.: Hierarchical goal representations in social causality simulations. In: Parallax Advanced Research. Beavercreek, OH (2020)
7. Falandays, J.B., Smaldino, P.: The emergence of cultural attractors: an agent-based model of collective perceptual alignment. In: Annual Meeting of the Cognitive Science Society, vol. 43 (2021)
8. Feldman, M.S., Pentland, B.T.: Reconceptualizing organizational routines as a source of flexibility and change. Adm. Sci. Q. **48**(1), 94–118 (2003)
9. Ferrada, E.: The amino acid alphabet and the architecture of the protein sequence-structure map. I. Binary alphabets. Plos Comput. Biol. **10**(12):e1003946 (2014)
10. Gilbert, N., Troitzsch, K.G.: Simulation for the Social Scientist, 2nd edn. United Kingdom, Open University Press, Buckingham (2005)

11. Huberman, B.A., Glance, N.S.: Evolutionary games and computer simulations. Proc. Natl. Acad. Sci. USA **90**(16), 7716–7718 (1993)
12. Kahneman, D., Tversky, A.: The Simulation Heuristic. In: Kahneman, D., Slovic, P., Tversky, A. (eds.) Judgment Under Uncertainty: Heuristics and Biases, pp. 201–208. Cambridge University Press, Cambridge, UK (1982)
13. Mudigonda, S., Núñez-Corrales, S., Venkatachalapathy, R., Graham, J.: Scheduler dependencies in agent-based models: a case-study using a contagion model. Computational Social Science Society of the Americas, Springer, Santa Fe, NM (2021)
14. Mudigonda, S.P., Friesen, M.J.: Social primitives: exploring spark of life collective behavior in agent-based models. In Proceedings of Conference of the Computational Social Science Society of the Americas, pp. 315–340, Springer, Berlin (2020)
15. Núñez-Corrales, S.: Toward a unified view of complex multiscale stochastic systems: a generalized theory of interactions and its computational infrastructure for their universal and efficient investigation. Thesis at University of Illinois at Urbana-Champaign, Department of Informatics (2020)
16. Núñez-Corrales, S., Friesen, M., Srikanth, M., Venkatachalapathy, R., Graham, J.: In-Silico models with greater fidelity to social processes: towards ABM platforms with realistic concurrency. Computational Social Science Society of the Americas, Springer, Santa Fe, NM (2020)
17. Núñez-Corrales, S., Jakobsson, E.: A generalized theory of interactions for complex multiscale stochastic systems with thermodynamic irreversibility. Bull. Am. Phys. Soc. **65** (2020)
18. Parunak, H.V.D., Brueckner, S.: Concurrent modeling of alternative worlds with polyagents. In: The Seventh International Workshop on Multi-Agent-Based Simulation (MABS06, at AAMAS06), pp. 128–141, Springer, Hakodate, Japan (2006)
19. Parunak, H.V.D.: Learning Actor Preferences by Evolution. Computational Social Science (CSS21), CSSSA, Santa Fe, NM (2021)
20. Parunak, H.V.D., Greanya, J., McCarthy, M., Morell, J.A., Nadella, S., Sappelsa, L.: SCAMP's stigmergic model of social conflict. Comput. Math. Organ. Theory (2021)
21. Parunak, H.V.D.: Model mechanisms and behavioral attractors. social simulation conference (SSC2022), pp. (in press), University of Milan, Milan, Italy (2022)
22. Roos, N.: Entropic forces in brownian motion. Am. J. Phys. **82**(12), 1161–1166 (2014)
23. Savit, R., Brueckner, S.A., Parunak, H.V.D., Sauter, J.: General Structure of Resource Allocation Games. Altarum, Ann Arbor, MI (2002). https://www.abcresearch.org/abc/papers/RAGpaper.pdf
24. Shnerb, N.M., Louzoun, Y., Bettelheim, E., Solomon, S.: The importance of being discrete: Life always wins on the surface. Proc. Natl. Acad. Sci. USA, **97**(19), 10322–10324 (2000)
25. Wilson, W.G.: Resolving discrepancies between deterministic population models and individual-based simulations. Am. Nat. **151**(2), 116–134 (1998)
26. Wissner-Gross, A.D., Freer, C.E.: Causal entropic forces. Phys. Rev. Lett. **110**(16), 168702 (2013)
27. Wolfram, S.: Cellular Automata and Complexity: Collected Papers. Reading, MA, Addison-Wesley (1994)
28. Zia, A., Koliba, C.: The emergence of attractors under multi-level institutional designs: agent-based modeling of intergovernmental decision making for funding transportation projects. AI Soc. **30**(3), 315–331 (2015)

Modeling Farmers' Adoption Potential to New Bioenergy Crops: An Agent-Based Approach

Kazi Ullah and Andrew Crooks

Abstract The use of fossil fuels is the primary source of greenhouse gas emissions but there are alternatives to these especially in the form of biofuels, fuels derived from bioenergy crops. This paper aims to determine farmers' potential adoption rates of newly introduced bioenergy crops with a specific example of carinata in the state of Georgia. The determination is done using an agent-based modeling technique with two principal assumptions—farmers are profit maximizer and they are influenced by neighboring farmers. Two diffusion parameters (traditional and expansion) are followed along with two willingness (high and low) scenarios to switch at varying production economics to carinata and other prominent traditional field crops (cotton, peanuts, corn) in the study region. We find that a contract prices around $9, $8 and $7 can be a viable option for encouraging farmers to adopt carinata in low, average, and high profit conditions, respectively. Expansion diffusion (that diffuses all over the geographical area), rather than centered to the few places like traditional diffusion at the early stage of adoption in conjunction with higher willingness conditions influences higher adoption rates in the short-term. As such, the model can be used to understand the behavioral economics of carinata in Georgia and beyond, as well as offering a potential tool to study similar bioenergy crops.

K. Ullah (✉) · A. Crooks
University at Buffalo, Buffalo, NY 14261, USA
e-mail: ullahkm@ornl.gov; ullahkazimasel@gmail.com

A. Crooks
e-mail: atcrooks@buffalo.edu

K. Ullah
Oak Ridge National Laboratory, Oak Ridge, TN 37830, USA

Z. Yang and S. Núñez-Corrales (eds.), *Proceedings of the 2022 Conference of The Computational Social Science Society of the Americas*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-031-37553-8_5

1 Introduction

The uses of fossil fuels such as petroleum, natural gas, and coal is the primary source of greenhouse gas (GHG) emissions. Globally, about 65% of GHG emissions in 2010 occurred due to burning fossil fuels and currently, commercial aviation is responsible for 2.6% of annual global CO_2 emissions [19]. Therefore, the mitigation strategies to combat climate change impacts are gaining attention from all sorts of transportation sectors, including that of the aviation industry. Utilizing biofuels can be the prime strategy in the goal to reduce GHG emissions. For example, it has been suggested that advanced biofuels produced from energy crops could reduce GHG emissions by as much as 50% when compared to fossil fuels [11]. In addition to this, bioenergy crops usually have high yield potentials; they can be grown productively on low-quality, fallow and marginal lands; they can increase soil carbon and reduce soil erosion (e.g., [35]), and are promising alternatives for rural economic development [29]. However, there is a great deal of economic, behavioral, and environmental challenges associated with adopting bioenergy crops, such as price and yield risks [24], lack of an established market [12], inexperience with new management practices, and cost of new crop-specific equipment [23]. For the last two decades several studies have tried to capture these challenges using several farm-scale modeling techniques ranging from agent-based models (e.g., [18]); choice experiment models (e.g., [23]); and mathematical programming models (e.g., [8]) to analyze different policy scenarios. Among those available modeling techniques, agent-based models have been argued to be an elegant tool for farm-scale modeling. The rationale for this is that farmers are heterogenous in their attitudes towards adopting bioenergy crops [30]. This heterogenous behavior is easily captured in such style of models [9], and it has been shown that agent-based modeling can capture farmers' heterogenous behaviors and their interaction with the biophysical environment and with other farmers [25]. These characteristics of agent-based models overcome some of the limitations of traditional econometric-based or theoretical microeconomic models which struggled incorporate heterogeneous behavior and spatial interactions [5]. It has also been argued that agent-based models can imitate the reality of farming and can be used more closely to understand the adoption pattern of a newly introduced bioenergy crops, where farmers' individual attitudes have large impact on overall adoption rates (e.g., [2, 20]).

However, to date there has only been a handful agent-based models that have explored bioenergy crop adoption (e.g., [2, 10, 15, 18, 20]). More recent work by Ullah and Dwivedi [30] tried to address two gaps from previous studies. The first was the joint determination of farmers' profitability, neighborhood influences and risk preferences to build a case in the field of computational social and behavioral science to study the perspective of farmers' attitude toward bioenergy crop adoption. Secondly, their work moved away from a hypothetical grid space and created a realistic but simple environmental landscape utilizing actual biophysical information to build farmers' interaction with the environment. Furthermore, Ullah and Dwivedi's [30] study applied three sub-modeling techniques under three principal

assumptions:—(1) farmers are profit maximizers [3], (2) farmers are influenced by their neighboring farmers [18], and (3) farmers are risk averse [24]. However, Ullah and Dwivedi's [30] study was limited to a small-scale watershed level (e.g., 650 farmer agents with 19,622 acre of farmland). Another major limitation in that study was how neighborhood influences where assigned. Their study represented only one neighborhood which was not connected to the surrounding neighborhoods. Thus, the adoption of the crop diffused over the simulation period monotonically by considering the previous adoption rates same for all the farmers over the study area. Therefore, in that study, farmers were influenced by only other neighboring farmers in the same community. Information sharing from other neighboring communities was missing, which is something that has been witnessed in reality (e.g., [4]). As a result, the diffusion of crop adoption could not be simulated to a greater geographical area.

This paper significantly extends the work Ullah and Dwivedi [30] by incorporating a greater regional study area, making it geographically explicit and assigning neighborhood influences on the farmers within their communities along with the surrounding communities. Thus, this study aims to model the adoption potential of bioenergy crops, where the case study is adopting carinata (an oilseed crop) as a newly introduced energy crop in context of Georgia, United States (US). In what follows, we first present the methodology and rationale for choosing the study area (Sect. 2) before presenting the results of the model in Sect. 3 and finally in Sect. 4 we provide a summary of the paper and areas of further work.

2 Methodology

While we present our model in the following sections, we also provide a more detailed *Overview, Design concepts and Details* (ODD) protocol [14] along with the model and the data needed to run the model at https://www.comses.net/codebase-release/5c2c06f0-3f6d-4f8d-b198-ce24b55feb2f/. This additional material allows for a more in-depth description of the model, as well as facilitates the replication of results or extension of the model.

2.1 Study Area and Purpose

Brassica carinata, or simply carinata, is a promising annual oilseed crop for the commercial production of sustainable aviation fuel (SAF [28]). As a cover crop, carinata can provide several ecosystem services by reducing soil erosion, nutrient leaching, increasing soil organic matter, and retaining moisture [17]. In the Southeast (SE) US, carinata can be potentially cultivated on about 3.4 million acres of fallow agricultural land during the winter season [1]. The SE is also home to the world's busiest airport, the Hartsfield-Jackson Atlanta International Airport, located in Georgia which consumes around 3.9 million tons of conventional aviation fuel

(CAF) per year, which is around 5.2% of the total CAF consumption in the US [31]. Therefore, the large supply of carinata feedstock could meet the immediate demand of SAF in Georgia and beyond. However, much of the contemporary research on promoting carinata in the US South has been done only at the experimental level [13]. The challenges in adopting carinata at a regional level is yet unknown. To address this challenge one of the first steps is to explore the farmers' attitudes towards adopting energy crops while still ensuring feedstock availability. Therefore, this study determines the potential adoption rates of carinata under different crop economics, behavioral and diffusion scenarios.

Building upon how Atlanta airport could utilize carinata for SAF, for a case study region we chose Georgia. Georgia is a state in the SE region of the US having an area of 95,635.24 sq. miles which agriculture/pasture make up 20% (US Department of Agriculture (USDA)/National Agricultural Statistics Service (NASS) [32]). With respect to agricultural lands, there are three major field crops: cotton, peanut and corn [22], however the vast majority of agricultural land occupied by these crops remains fallow in the winter season [1]. This provides opportunities for alternative crops in the winter. For example, it is estimated that around 1.9 million acres of these fallow lands could be utilized for cultivating carinata [1]. Figure 1 shows the county-wise potential land availability for producing carinata seeds [13] and the purpose of our study to understand how the diffusion of carinata adoption would take place. Our aim is to determine the future adoption rates of carinata as a newly introduced bioenergy crop in Georgia for producing SAF.

2.2 Entities, State Variables, and Scales

The agents in our model represent farmers. Each farmer is an agent who owns 247 acre of crop land, which is average farm size in Georgia [33] and follow either one of the three most popular three years crop rotations in Georgia: Cotton-Cotton-Cotton; Cotton-Cotton-Peanuts; Cotton-Cotton-Corn. These three rotations represent at least 95% of field crops in Georgia [22]. In each county, three types of farmer agents are randomly created according to their respective crop rotation's ratio of total farmlands estimated from the total crop areas divided by the average area of crop land. We would argue that the creation of farmer agents in this manner is a good approximation for giving farmers aggregated information at county level, while at the same time preserving privacy of farmers which is often done in other agricultural agent-based models like AgriPoliS [16].

Agents' attitudes towards cultivation of certain row crops (e.g., corn, cotton, peanuts and carinata) on farmlands are defined by several profit maximizing variables and neighborhood influences' parameters. The profit maximization variables (i.e., yield, production cost and price) determine the crop production economics at the discounted value to evaluate the profitability of integration of carinata into the major traditional crop rotations. While the production costs involve operating costs for producing a crop, including seeds, fertilizer, irrigation, fuels, and other similar

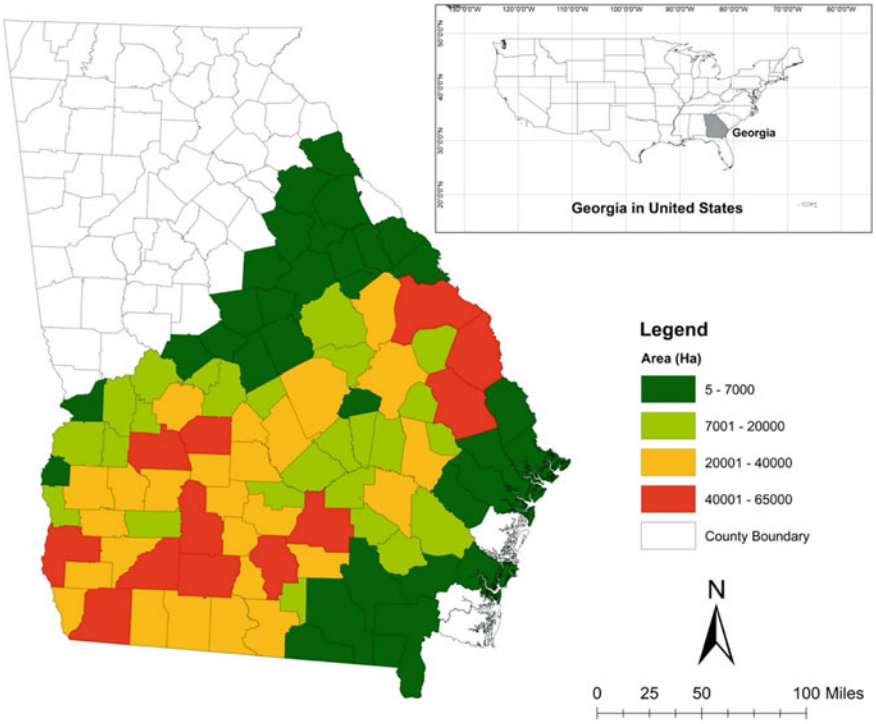


Fig. 1 County-wise land availability for carinata production (adopted from [13])

services. The allotted overhead costs, such as costs of labor, machinery and equipment, taxes, insurance, and other general farming overheads, are not included in net return estimations. As carinata is a row crop, it does not require any new machinery and equipment when being introduced into a new area. Therefore, considering only operating costs have no biasness for profit comparisons. Dedicated energy crops, which have no food value, are cultivated by the farmers as a cash crop in aspiration of making profit [23]. Therefore, without making profit farmers are not motivated for cultivating energy crops. The neighborhood influence parameters ascertain the adoption rate of rotations with carinata in the neighborhood that includes the county that a farmer belong and the surrounding adjacent counties. Each farmer has his own adoption threshold, which reflects how positive a farmer is for adopting carinata compared to the adoption rate [2, 4]. The adoption threshold parameters with different standard deviations shows whether the initial willingness on adopting carinata among the farmers are high or low.

Typically, it is suggested that carinata should be produced as double crop, once in every three years with two-years rotation gap [27]. To account for this in this model, the rotational period is three years for farm-scale modeling, which is extended up to 33 years (2018–2050) for long-term planning so that biorefinery investors can observe the feasibility of supply in the long run. Each tick or time step (t) of the

model therefore represents three years (due to crop rotations). The model is spatially explicit in terms of aggregating individual agents' decisions to the county level. However, decisions are made at the farm level.

2.3 Process Overview and Scheduling

Farmer agents' adoption decisions of carinata are reflected in two sub-models of profit modeling, and diffusion modeling (see supplementary material at https://www.comses.net/codebase-release/5c2c06f0-3f6d-4f8d-b198-ce24b55feb2f for more details). The profit modeling evaluates farmers' profits of row crop rotations with and without carinata. While the diffusion modeling determines farmers' attitudes towards adopting carinata under neighborhood influences. Farmers decide to adopt carinata for the current period only when they find their profit with carinata rotation is greater than without carinata rotation in the previous period, and the neighborhood influences from the same previous period build a positive outlook for adoption. For each farmer, the adoption behavior of the current period is updated and feeds into the next time step. Thus, the model works in a recursive manner until the end of the simulation period. Figure 2 shows the flow of farmers' decision-making framework spread across two sub-models. How farmers decisions are reflected in both sub-models is discussed at detail in the supplementary material.

2.4 Initialization

At the farm level, three categories of farmer agents are created in each county with specific crop rotations:—(1) cotton-cotton-cotton farmers; (2) cotton-cotton-peanut farmers; and (3) cotton-cotton-corn farmers. The number of farmer agent in each category under a particular county was created according to the ratio of those three major crop rotations among the total field crop area of the county for the year of 2015–2017. The total farmland and the ratios of major rotations of each county were captured from the Crop Data Layer (CDL) [32]. The creation of farmer agents from actual crop rotation histories enabled us to build a more spatially and temporarily informed agent-based model compared to existing models in the context of energy crop adoption (e.g., [7, 20, 26]). By utilizing this temporally and spatially explicit crop distribution attribute, we can more realistically estimate farmers' profits and the potential integration of carinata into traditional rotations according to the agronomic conditions (e.g., the herbicide effect for cultivating carinata after peanuts) [27]. The adoption thresholds of the farmers are set using two normal distributions, which ultimately create a high and a low initial willingness scenario. Initial adoption rate (AR) is assigned zero at the start of the simulation period. The AR value is a neighborhood level value (which is discussed further below), however, all the farmers

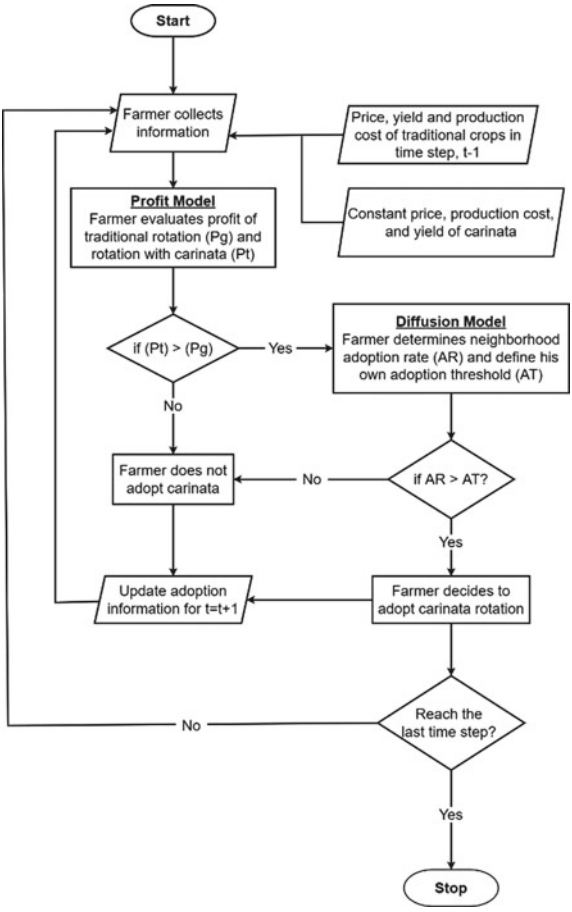


Fig. 2 Process, overview and scheduling

in their respective neighborhoods are equally informed about that value, hence, the parameter is assigned as a farmers' attribute for simulation purposes.

At the global level, crop yields, prices, and farming costs are set at the initialization of the model. The crop economics data is acquired from USDA, Economic Research Service at the Southern Seaboard regional level [34]. The initial contract price of carinata is fixed by analyzing historical crop rotations and by comparing with the best profitable scenarios of traditional crop rotations in previous three years period from base year (see [30] for more details).

Two diffusion types are selected—(1) Traditional and (2) Expansion diffusion [21] as shown in Fig. 3. The rationale for exploring these different diffusion processes is to explore how carinata might defuse over the area. These could be considered as two different policy options—one in which a pilot study is focused on a small geographical location, and one in which farmers are selected from across the state (i.e., the entire Georgia for this study). Traditional diffusion, which can also be

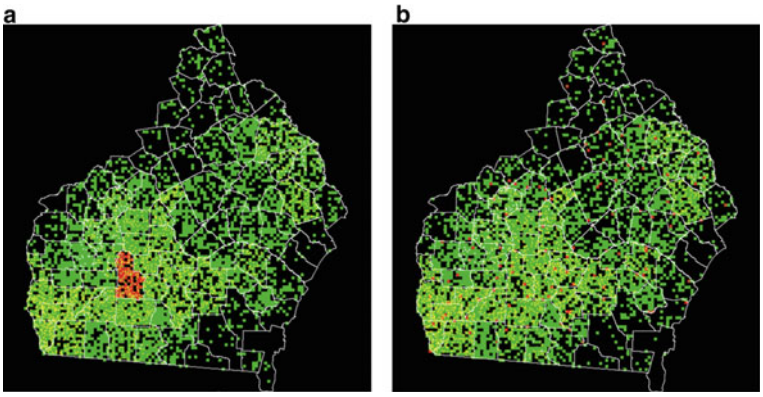


Fig. 3 Adoption scenarios at time step 1 where Red agents (i.e., farmers) are the early adopters: **a** traditional **b** expansion diffusion exam

considered as contiguous diffusion, (e.g., [21]) starts from pilot site in a single or few adjacent counties located at Experimental Little River Watershed, which contains around 2.5–5% farmers of the whole Georgia. This part of Little River Watershed is the most prominent place in Georgia for doing experimental and field research led by universities and relevant agricultural extension departments [6]. Starting at this watershed, the adoption behavior is diffused from early adopters to the neighboring farmers and subsequently, it spreads all over Georgia throughout the simulation period. In case of expansion diffusion, the early adopters are spread all over Georgia at the initial stage rather than located within a single small geographical area; and then, other neighboring farmers learn from their experiences, thus, the adoption behavior is diffused over the study area.

3 Results

Before presenting the results of the model, we applied several verification processes to ensure that the model matches its design. We achieved model verification using iterative design review (i.e., code walkthroughs), visual debugging and parameter testing via sensitivity analysis. Once we were satisfied the model was verified, we then moved onto scenarios exploration. The model is simulated under three profits (low, average, high), two diffusions and two willingness scenarios on adoption with the contract prices of carinata at $7, $8, $9 and $ 10. The low profit scenario is defined with lowest yield (40 bu/acre) and highest production cost ($280/acre) of carinata. The average profit is determined with average yield (50 bu/acre) and production cost ($270/acre). The high profit is calculated using highest yield (60 bu/acre) and lowest production cost ($260). Each scenario of the model is run for 10 times and the average results are presented in this paper.

According to Fig. 4, there is virtually no profit outcomes at price, $7, hence, there are very low adoption rates in the long run (e.g., only 3.3% by 2050 on average, and less than 5% for any scenarios). Adoption rates go considerably higher from $8 to $9 and $10. For instances, 7,408 (41.2%), 16,585 (92.2%), and 17,232 (95.8%) farmers adopted carinata by 2032 at a price of $8, $9 and $10, respectively under traditional and low initial willingness scenario (Fig. 4a). However, the adoption rates fluctuate at the price of $9 and $10, and there is no significant difference between them for long term adoptions. For example, in 2044, the adoption rate under expansion diffusion and high willingness scenario was 80% at a price of $10 and that adoption rate under the similar scenario was 94.5% at the price of $9 (Fig. 4d). By 2050, the adoption rate at $10 was higher than that of $9. From investors perspective, they would rather offer a contract price of $9 to the farmers rather than fixing it more than that, because investors will get almost equal adoption rates at this price. The overall adoption rates in the long run remain almost similar for all the scenarios, but adoption rates are higher at the year of 2026 for high initial willingness scenarios, and for expansion diffusion with high willingness scenarios, that figure is highest at any given contract price.

Carinata could be a profitable enterprise even at a price of $7 in an average yield and production cost as shown in Fig. 5. A high adoption rate appeared at $7 under any condition of average profit scenarios. However, the adoption rates get more stable at $8. Therefore, investors may look for getting carinata seeds at that contract price. Similar to Figs. 4 and 5 also shows the higher adoption rates (e.g., 60.7% and 65.6% at price $8) in short term (before 2026) in higher initial willingness scenarios (Fig. 5b, d) and highest in case of expansion diffusion (Fig. 5d). The same findings are also reflected with Fig. 6 in the high profit scenarios. However, in high profit scenarios, a contract price of $7 can be enough to maintain the desirable adoption rates in the

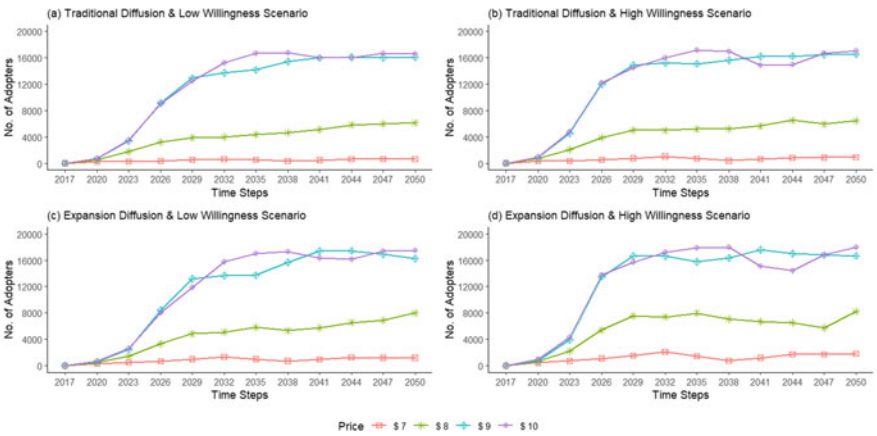


Fig. 4 Number of farmers who adopt carinata in specific rotation years with low profit condition (carinata yield = 40 bu/acre, carinata production cost = $280/acre)

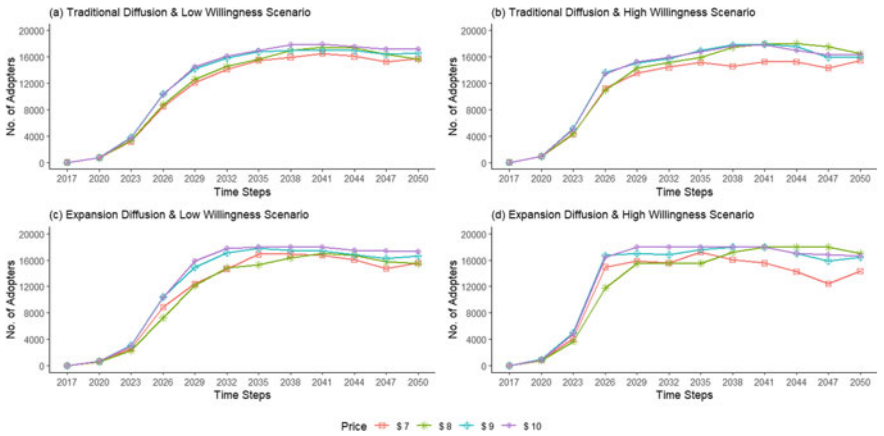


Fig. 5 Number of farmers who adopt carinata in the rotation years with average profit condition (carinata yield = 50 bu/acre, carinata production cost = $270/acre)

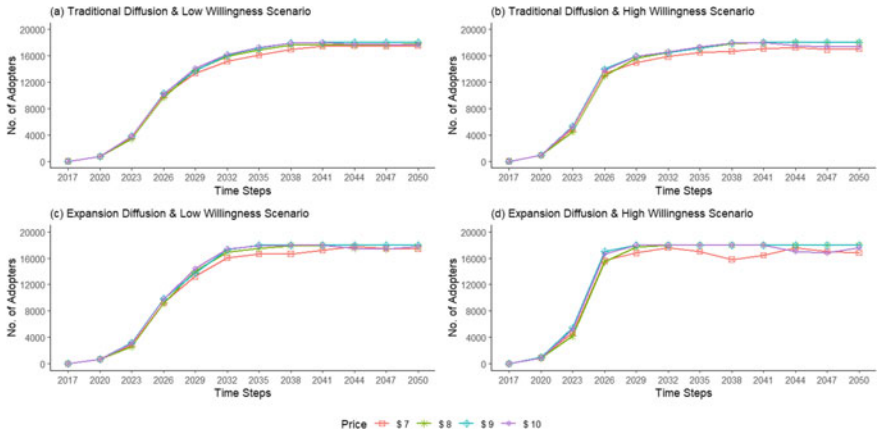


Fig. 6 Number of farmers who adopt carinata in the rotation years with high profit condition (carinata yield = 60 bu/acre, carinata production cost = $260/acre)

long run, such as the adoption rates ranged between 93.2 and 97.3% by 2050 for all the scenarios at that price.

4 Summary

The aim of this paper was to explore how farmers might adopt bioenergy crops across a large geographical area. Results from our model suggest that a viable contract price made by investors can persuade farmers to adopt carinata. Similar to what one might

observe in the real world (e.g., [4]), the adoption rate for this newly introduced crop remains low at first but changing the dynamics of initial willingness on adoption can speed up the rate of this new bioenergy crop, especially under the expansion diffusion scenario. Therefore, it could be suggested that if there was a campaign to promote the adoption of this bioenergy crop, policy makers should consider its initial dispersion of pilot sites over a large geographical region and provide a reasonably high contract price.

Looking towards future work, all models have their limitations, and this model is no different. One such area is that of the farmers' risk aversion, which is currently not accounted for in this model, but it might impact their land allocation decisions (e.g., [2]). Therefore, one might want to explore the usefulness of risk portfolio estimation methods (e.g., mean–variance optimization, statistical dominance analysis), which could be embedded within the farmers individual decision making. In addition to this, the current model does not consider dynamic environmental factors such as weather conditions that can affect crop yield or how yield variations across the counties due to different soil conditions and the environmental benefits or loss that can accrue for cultivating bioenergy crops. To fulfill this gap, our future modeling work will consider frost event frequencies based on past events along with potential yield variations and net soil organic carbon stocks which can be estimated from other models such as the DayCent model [13]. Even with these limitations and areas of further work this paper offers a new way to explore how farmers might adopt bioenergy fuels and through the provision of the source code and data allows others to extend or adapt the model to their own biomass and bioenergy study fields which could be used as a tool to reduce greenhouse gas emissions from fossil fuels.

References

1. Alam, A., Dwivedi, P.: Modeling site suitability and production potential of carinata-based sustainable jet fuel in the southeastern United States. J. Clean. Prod. **239**, 117817 (2019)
2. Alexander, P., Moran, D., Rounsevell, M.D.A., Smith, P.: Modelling the perennial energy crop market: the role of spatial diffusion. J. R. Soc. Interface **10**(88), 20130656 (2013)
3. Anand, M., Miao, R., Khanna, M.: Adopting bioenergy crops: does farmers' attitude toward loss matter? Agric. Econ. **50**(4), 435–450 (2019)
4. Berger, T.: Agent-based spatial models applied to agriculture: a simulation tool for technology diffusion, resource use changes and policy analysis. Agric. Econ. **25**(2–3), 245–260 (2001)
5. Berger, T., Troost, C.: Agent-based modelling of climate adaptation and mitigation options in agriculture. J. Agric. Econ. **65**(2), 323–348 (2014)
6. Bosch, D.D., Sheridan, J.M., Lowrance, R.R., Hubbard, R.K., Strickland, T.C., Feyereisen, G.W., Sullivan, D.G.: Little River experimental watershed database. Water Resour. Res. **43**(9), W09470 (2007)
7. Brown, C., Bakam, I., Smith, P., Matthews, R.: An agent-based modelling approach to evaluate factors influencing bioenergy crop adoption in north-east Scotland. GCB Bioenergy **8**(1), 226–244 (2016)
8. Chen, X., Önal, H.: Modeling agricultural supply response using mathematical programming and crop mixes. Am. J. Agr. Econ. **94**(3), 674–686 (2012)
9. Crooks, A.T., Malleson, N., Manley, E., Heppenstall, A.: Agent-based modelling and geographical information systems: a practical primer. SAGE Publications, London, UK (2019)

10. Ding, D., Bennett, D., Secchi, S.: Investigating impacts of alternative crop market scenarios on land use change with an agent-based model. Land **4**(4), 1110–1137 (2015)
11. El Akkari, M., Réchauchère, O., Bispo, A., Gabrielle, B., Makowski, D.: A meta-analysis of the greenhouse gas abatement of bioenergy factoring in land use changes. Sci. Rep. **8**(1), 8563 (2018)
12. Fewell, J.E., Bergtold, J.S., Williams, J.R.: Farmers' willingness to contract switchgrass as a cellulosic bioenergy crop in Kansas. Energy Econ. **55**, 292–302 (2016)
13. Field, J.L., Zhang, Y., Marx, E., Boote, K.J., Easter, M., George, S., Hoghooghi, N., Johnston, G., Masum, F.H., Mulvaney, M.J., Paustian, K., Seepaul, R., Swan, A., Williams, S., Wright, D., Dwivedi, P.: Modeling yield, biogenic emissions, and carbon sequestration in southeastern cropping systems with winter carinata. Front. Energy Res. **323** (2022)
14. Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S., Huse, G., Huth, A., Jepsen, J., Jorgensen, C., Mooij, W., Muller, B., Pe'er, G., Piou, C., Railsback, S., Robbins, A., Robbins, M., Rossmanith, E., Ruger, N., Strand, E., Souissi, S., Stillman, R., Vabo, R., Visser, U., Deangelis, D.: A standard protocol for describing individual-based and agent-based models. Ecol. Model. **198**(1–2), 115–126 (2006)
15. Guillem, E.E., Murray-Rust, D., Robinson, D.T., Barnes, A., Rounsevell, M.D.A.: Modelling farmer decision-making to anticipate tradeoffs between provisioning ecosystem services and biodiversity. Agric. Syst. **137**, 12–23 (2015)
16. Happe, K., Kellermann, K., Balmann, A.: Agent-based analysis of agricultural policies: an illustration of the agricultural policy simulator AgriPoliS, its adaptation and behavior. Ecol. Soc. **11**(1), 49 (2006)
17. Hoghooghi, N., Bosch, D.D., Bledsoe, B.P.: Assessing hydrologic and water quality effects of land use conversion to Brassica carinata as a winter biofuel crop in the southeastern coastal plain of Georgia, USA using the SWAT model. GCB Bioenergy **13**(3), 473–492 (2021)
18. Huang, S., Hu, G., Chennault, C., Su, L., Brandes, E., Heaton, E., Schulte, L., Wang, L., Tyndall, J.: Agent-based modeling of bioenergy crop adoption and farmer decision-making. Energy **115**, 1188–1201 (2016)
19. ICAO Environmental Report 2016, https://doi.org/10.1017/CBO9781107415324.004. Last acceded 30 Oct 2022
20. Jin, E., Mendis, G.P., Sutherland, J.W.: Spatial agent-based modeling for dedicated energy crop adoption and cellulosic biofuel commercialization. Biofuels, Bioprod. Biorefin. **13**(3), 618–634 (2019)
21. Jordan-Bychkov, T.G., Domosh, M., Neumann, R.P., Price, P.L.: The Human Mosaic: A Thematic Introduction to Cultural Geography, 2nd edn. Macmillan, London, UK (2012)
22. Karami, O., Dwivedi, P., Lamb, M., Field, J.L.: Economics of crop rotations with and without carinata for sustainable aviation fuel production in the SE United States. Frontiers Energy Res. **10**, 461 (2022)
23. Khanna, M., Louviere, J., Yang, X.: Motivations to grow energy crops: the role of crop and contract attributes. Agric. Econ. **48**(3), 263–277 (2017)
24. Miao, R., Khanna, M.: Are bioenergy crops riskier than corn? Implications for biomass price. Choices: Mag. Food, Farm, Resour. Issues **29**(1), 1–6 (2014).
25. Nolan, J., Parker, D., van Kooten, G.C., Berger, T.: An overview of computational modeling in agricultural and resource economics. Can. J. Agric. Econ. **57**(4), 417–429 (2009)
26. Schulze, J., Gawel, E., Nolzen, H., Weise, H., Frank, K.: The expansion of short rotation forestry: characterization of determinants with an agent-based land use model. GCB Bioenergy **9**(6), 1042–1056 (2017)
27. Seepaul, R., George, S., Small, I., Marois, J., Wright, D.: Best Management Practices for Carinata Production in the Southeast. https://sparc-cap.org/wp-content/uploads/2018/03/Carinata-best-management-practices.pdf. Last acceded 30 Oct 2022
28. Seepaul, R., Kumar, S., Iboyi, J.E., Bashyal, M., Stansly, T.L., Bennett, R., Boote, K.J., Mulvaney, M.J., Small, I.M., George, S., Wright, D.L.: Brassica carinata: biology and agronomy as a biofuel crop. GCB Bioenergy **13**, 582–599 (2021)

29. Singh, B.P.: Biofuel crop sustainability baradigm. In: Singh, B.P. (ed) Biofuel Crop Sustainability, pp. 3–29, Wiley, London, England (2013)
30. Ullah, K.M., Dwivedi, P.: Ascertaining land allocation decisions of farmers about the adoption of carinata as a potential crop for sustainable aviation fuel production in the southern United States. GCB Bioenergy **14**(7), 824–839 (2022)
31. US Department of Transportation, Airline fuel cost and consumption, https://www.transtats.bts.gov/fuel.asp. Last acceded 30 Oct 2022
32. USDA/NASS, Cropland Data Layer, https://nassgeodata.gmu.edu/CropScape/. Last acceded 30 Oct 2022
33. USDA/NASS, State agriculture overview for Georgia. Accessed 12 July 2022, from https://www.nass.usda.gov/Quick_Stats/Ag_Overview/stateOverview.php?state=GEORGIA. Last acceded 30 Oct 2022
34. USDA Economic Research Service, Commodity Costs and Returns, https://www.ers.usda.gov/data-products/commodity-costs-and-returns/. Last acceded 30 Oct 2022
35. Yadav, P., Priyanka, P., Kumar, D., Yadav, A., Yadav, K.: Bioenergy crops: Recent advances and future outlook. In: Rastegari, A.A., Yadav, A.N., Gupta, A. (eds.) Prospects of Renewable Bioprocessing in Future Energy Systems, pp. 315–335. Springer, New York, NY. (2019)

Investigating Emergency Responders' Roles in a Dirty Bomb Event with an Agent-Based Model

Ellie Q. Chen and William G. Kennedy

Abstract Dirty bombs, formally known as radiological dispersal devices (RDDs), pose a new and potentially significant threat to the populace. They can be used by terrorists to cause explosions that would expose nearby civilians to radiation, resulting in injuries of varying severities, including death. The assistance of emergency responders may mitigate the negative consequences of dirty bomb events. We used an agent-based model to simulate what happens when emergency responders respond to a dirty bomb exploding in a public metropolitan place. Through our study, we confirmed the positive impact of emergency responders in a dirty bomb event and raised possible concerns for the safety of emergency responders.

1 Introduction

When the Russia-Ukraine war began this year, the fear of a nuclear confrontation arose in the world. In July of this year, New York even released a public service announcement in the case of a nuclear attack [3]. A news article from Science, the journal of the American Association for the Advancement of Science, states that in March, ingredients needed to construct a dirty bomb were lost, presumably stolen, from a Chornobyl Nuclear Power Plant monitoring lab and that "Chornobyl is not the only Ukrainian nuclear installation at risk in the war" [17], indicating concern regarding the loss of control of nuclear materials. The consternation caused by this incident shows the danger of RDDs and the necessity to study the inner workings and outcomes of dirty bomb events.

E. Q. Chen (✉)
Thomas Jefferson High School for Science and Technology, Alexandria, VA 22312, USA
e-mail: ellienova@gmail.com

W. G. Kennedy
Department of Computational and Data Sciences, George Mason University, Fairfax, VA 22030, USA
e-mail: wkennedy@gmu.edu

Z. Yang and S. Núñez-Corrales (eds.), *Proceedings of the 2022 Conference of The Computational Social Science Society of the Americas*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-031-37553-8_6

Emergency responders, also known interchangeably as first responders, are licensed and trained personnel including but not limited to law enforcement officers, emergency medical technicians (EMTs), and firefighters and are vital to dirty bomb events. According to the Occupational Safety and Health Administration (OSHA), emergency responders are responsible in the case of radiation emergencies [13]. The International Atomic Energy Agency (IAEA) provides a Manual for First Responders to a Radiological Emergency that defines a radiological emergency as "emergencies involving radioactive material that can occur anywhere and include... malicious threats/acts" [8]. By this definition, the detonation of a dirty bomb is a radiological emergency and would call for the assistance of first responders before resources from national organizations, such as the Department of Energy (DOE), Department of Homeland Security (DHS), and the FBI, respond. In such an emergency, first responders would save lives and guide civilians to safe areas away from radiation, and their role is therefore especially important.

The presence of first responders can thus lessen the negative consequences of dirty bomb explosions. Because first responders are a valuable and potent resource, it is essential that we study their role in dirty bomb events and similarly perilous situations. First responders are especially exposed to radiation and other hazards while undertaking their tasks [13], so it is imperative that we confirm their safety in radiological emergencies. Our study aims to investigate and report the impact of the presence of first responders on civilians' wellbeing as well as examine the wellbeing of first responders themselves in a dirty bomb event.

Our main contributions: (i) built an agent-based model whose framework can be used to create other models of dirty bomb events in metropolitan areas; (ii) provided potential to extend the model to include other types of agents and agent behaviors, especially interactions between agents (e.g. family members); (iii) conducted research through an experiment of the impact of on-site first responders in dirty bomb events, through which we obtained and discussed results for the sake of better preparedness in a real-life dirty bomb event.

2 Background

The sun provides all of us some exposure to natural radiation. It is a very large radioactive source at an average distance of 93 million miles away. With too much exposure, we can get a sunburn, usually mild but potentially a serious second-degree burn. Radiation from a nearby radiological source is worse. The amount of radiation received is dependent on the strength of the source, the distance from the source, whether there is any shielding between us and the source, and finally, the length of time we are exposed. That last factor can be under our control. Generally, radiation spreads out in all directions and its effects decrease with the square of the distance from the source and whether there is any shielding [12].

A basic thumb rule for determining the amount of radiation received is the Curie-Meter-REM rule [4]. It is that 1 Curie of radioactive material at 1 m results in a

dose rate of 1 REM per hour. A REM (Roentgen Equivalent in Man) is a unit of the radiation effects on a person. The dose can be scaled directly based on different source strengths, inversely by the square of the distance in meters, and directly based on the duration of the exposure. Therefore, a basic part of the guidance on minimizing the dose received is to get away from the source as quickly as possible, thereby increasing the distance and decreasing the exposure time.

The pocket guide issued by the Center for Disease Control and Prevention (CDC) for clinicians [2] describes dangerous doses, i.e., more than about 800 REM as a lethal dose with vomiting expected within ten minutes and death within a day or two. People trained to work with radioactive sources are allowed to receive a maximum of 5 REM per year and the general public protected such that they should not receive more than half a REM per year [12]. However, there have been some radiological accidents with exposures above these limits [9].

One of the worst accidents or incidents was in 1987 in Goiânia, Brazil. A radiotherapy institute's machine was in the process of being decommissioned when the process was held up by a legal dispute. During the delay, the radioactive Cesium source used in a therapy unit, approximately 1,400 Curies (Ci), was left unprotected and two thieves scavenged it and exposed themselves and others to lethal doses over a few days of close contact with the radioactive material. The incident was reported to the IAEA and they issued a report [7].

A major study of radiological dispersal devices (RDDs) (the formal name of a dirty bomb) was conducted in 2007 [16]. They considered the risks and economic impact of three dozen scenarios involving up to spent fuel rods from nuclear reactors and industrial irradiators for sterilization and food preservation can be upwards of 2-4 million Curies and disrupting the seaport of Los Angeles. The handling of such "hot" sources requires special equipment and processes to protect those handling such sources. The source in the Brazilian incident was at the upper end of the sources used in radiotherapy.

There are many studies on modeling dirty bomb events. Some, however, do not consider metropolitan tourist areas, e.g. Manley et al. [10] concerns airport settings, Pereira and Delgado [15] examines the Olympia village, Duan et al. [6] regards subway settings. Some do not model individual-level behaviors [1, 14, 18]. In most existing literature, including the aforementioned, first responders are not the focus of the research. As commented in [19], "emergency responders have not been adequately studied." Our paper models individual-level behaviors in a dirty bomb event in a metropolitan area and focuses on first responders.

Last year, a study considered the effects of a Brazilian-scale source in a dirty bomb on the National Mall in Washington, DC. That study was reported at the Annual Meeting of the Computational Social Science Society of the Americas [23] and discussed the impacts of a dirty bomb at a crowded public gathering such as the 4th of July. The model discovered that only those injured by the bomb's explosion and unable to get away from the source on their own received significant doses. In this extension of that study, we have included modeling the roles of first responders.

3 Methodology

3.1 About Emergency Responders

The Department of Homeland Security (DHS) provides a document titled "Radiological Dispersal Device Response Guidance: Planning for the First 100 Minute" [5], which supplies a timeline (Annex 7) for actions that emergency responders take after the dirty bomb explosion. According to the timeline, two minutes after the explosion, emergency responders would begin "lifesaving rescue operations"; by five minutes, they would "confirm the presence of radiation" and subsequently establish a Hot Zone, first defined as areas within 250 m of the explosion, and Shelter-In-Place Zone, first defined as areas within 250–500 m of the explosion [5]. Civilians are expected to leave the Hot Zone and shelter in buildings in the Shelter-In-Place Zone.

Our model makes two assumptions and deviations from the information provided by official sources: (i) we assumed that exit points for civilian evacuation, which could also provide a quick screening and decontamination, would be set up at five minutes after the explosion; (ii) we assumed that civilians will not be sheltering in place. The second assumption is based on the fact that buildings in the National Mall are closed to visitor entry. Thus, all modeled victims avoid entering buildings as they move. Otherwise, we have followed the aforementioned sequence of events in our simulation.

3.2 About the Previous Model

Our model is built on a previous model [23]. That model simulated civilians fleeing from a dirty bomb explosion in the National Mall (an area in Washington, D.C.) and recorded the numbers of those injured or killed and each person's amount of radiation received, considering the number of civilians present and the power of the bomb. The previous and new models were created using NetLogo [22].

In the previous model, civilians were represented by moving agents, the layout of the National Mall was represented by patches with a resolution of five meters, and the passing of time was represented by ticks representing ten seconds each. The previous model utilized a map of the National Mall, which can be seen in Fig. 1, 1000 m long and 500 m wide, with light green patches representing grass, dark green patches representing the picnic area, black patches representing buildings, and white patches representing roads. The model set up a number of civilians randomly spread out on the picnic area. After a dirty bomb detonates on a set coordinate in the picnic area, civilians flee on foot away from the explosion. Depending on their distance from the explosion, they have a delay in detecting the danger. Some civilians hesitate before running for zero to six ticks (zero to sixty seconds), discretely; the duration of their hesitance varies. The model continues running for another sixty ticks after the

Fig. 1 Map of the National Mall used in our simulation (courtesy of Xiong and Kennedy)

detonation. To this model, we added both a variation of speed among the population and the assistance of emergency responders.

3.3 About Our Model

Our model used various components from the previous model, with a few important distinctions:

1. Modeled different agent types using different classes.
2. Used a set of scaling factors to regulate variations in speed between individuals and among different conditions.
3. Distributed civilians on grass patches instead of limiting them to the picnic area.
4. Improved the direction-finding algorithm.
5. Allowed civilians to flee on or off the roads.
6. Wrote detailed data during the simulation to files for analysis.

3.3.1 Modeling Victims and First Responders

There are four types of agents within the model, the first of which is civilians. These are pedestrians walking around the area. The other three types, denoted rescuers, ushers, and exit locations, are different types of emergency responders conducting different tasks. Rescuers are responsible for lifesaving rescue operations; ushers move along in the area and tell mobile, uninjured civilians which direction to head; exit locations are stationary points to which civilians are guided to undergo screening and decontamination-they are also the only type of agents that doesn't represent humans.

3.3.2 Simulation Timeline and First Responder Behavior

According to Wikipedia, the preferred walking speed for humans is 1.42 m/s [20]. The National Mall park receives around 32 million visitors each year [11], so we estimated that around 1000 civilians would be present at any moment on a normal day. Before the explosion, a population of 1000 civilians are modeled as walking on the grass (the dark and light green patches) at a base walking speed of 1.5 m/s, which is three 5 m patches each 10 s tick. To account for heterogeneity in mobility among the population due to age, health, and other factors, we added an individual-level scaling factor on the base speed, sampled from a uniform distribution $U(0.5, 1.5)$.

Suddenly, the detonation, with a source of 13800 Curies, occurs. Civilians within one patch (5 m) of the center of the blast are killed instantaneously and civilians within five patches (25 m) are severely injured. Both of these groups are consequently modeled as unable to move. The rest, after realizing the explosion and hesitating, flee at a base running speed of 2.0 m/s (4.0 patches/tick) [21], and the same mobility-based scaling factor applies to the base running speed. We added three other scaling factors to account for differences in speed among victims due to injury, moving on different surfaces (grass versus roads), and congestion. In the previous model, civilians would stay on the road once they had fled onto it. In our model, they were not restricted to this behavior.

Two minutes (12 ticks) after the explosion, in accordance to the timeline provided by the DHS, rescuers arrive in the area, each one entering through one of four exit locations: those four locations are places where exits are to be set up. While ignoring the victims who have already died, the rescuers pair up with each other to carry out injured victims to the exit from which they set out. Their initial speed is set to the product of the highest scaling factor (1.5) and the base running speed (4.0 patches/tick); they are effectively the fastest agents at the site, moving at 3.0 m/s (6.0 patches/tick), since we assume that first responders performing search-and-rescue operations would be more physically fit than average. When the rescuers pick up a victim and carry them, however, they move at the base rescue speed, which is 1.0 patches/tick, or 0.5 m/s. Rescuers are affected by congestion-based and surface-based scaling factors.

Five minutes (30 ticks) after the explosion, the exits are established and ushers are dispatched. Four exits are set up at fixed coordinates around the area near the outer edges, one in each corner of a rectangle. Civilians queue at these exits while arriving and go through a quick screening and decontamination at a processing rate of one person per tick. Ushers each enter from a random exit and run throughout the area, and when they see fleeing, mobile civilians, they tell them to head toward the nearest exit. That civilian would then head toward that exit while maneuvering around buildings. Ushers are affected by congestion-based and surface-based scaling factors.

The simulation ends 100 min (600 ticks) after the explosion. This is substantially longer than the previous model [23]. According to the DHS, during this time, “it is unlikely that federal and state support is on scene” [5], so assistance of first responders is especially crucial.

3.3.3 Direction-Finding Algorithm for Agents

After the explosion, each agent has a target direction, either toward or away from the explosion. A civilian always runs away from the explosion site. A rescuer first runs towards the Hot Zone and, after picking up the injured victim, runs towards the exit location from which they set out. An usher, however, runs at a random direction throughout the simulation. When an agent encounters a building, they adjust their direction temporarily so that they can keep moving while maintaining their target direction.

To model this behavior, we used the following algorithm: connecting the agent with the explosion location, finding the two possible directions after having turned the minimum turn that allows them to keep moving, computing the two angles formulated by the two directions and the direction of the explosion location, and choosing the one that maintains their target direction. If the target direction is away from explosion, they choose the direction with a larger angle, and vice versa.

4 Results

4.1 Experimental Design

To investigate the role of first responders in a dirty bomb event and to consider potential policy changes, we have designed the following experiment, in which we studied these two key variables:

- *Presence of first responders*. We examined two scenarios, one with first responders and one without, while setting all other parameters constant. We reported the effect of the assistance of first responders on the average and maximum doses of radiation received by civilians. We hypothesized that the presence of first responders would lower those values.
- *Arrival time of exits and ushers*. We swept the arrival time of exits and ushers between 2 and 10 min, discretely, after the explosion and reported the average and maximum doses of radiation received by civilians. We hypothesized that if the exits were set up sooner, the average and maximum doses of radiation received by civilians would be lower than otherwise. While there would be a cost to quicken the dispatching of first responders, it might be beneficial for the wellbeing of the populace.

In all cases in which first responders are present, we also reported the maximum and average doses of radiation that first responders receive.

4.2 Observations and Explanations

We ran our model in two different ways: (i) running 100 trials per scenario, each trial with a different seed, and reporting the median values of them; (ii) setting a fixed random seed and running that once. Overall, the results demonstrated the benefits of the presence of first responders.

Note: After 25 min, the vast majority of civilians have evacuated the area, with the exception of a few people injured in the initial blast. Therefore, some of the following figures only depict data for up to 25 min after the explosions, as the graphs plateau afterward.

4.2.1 Benefits of First Responders

From Fig. 2, we find that with first responders, the average dose a civilian receives over time increases sublinearly after ushers arrive and exits are set up. In fact, the average dose reaches an upper bound of about 0.18 REM 20 min after the explosion. On the contrary, without first responders, the average dose a civilian receives increases linearly. This is likely because immobile civilians in the Hot Zone are rescued and taken away at around 20 min after the explosion. The average dose received by civilians thus plateaus afterward.

Without rescuers, the civilians injured by the blast stay near the source and receive constant rates of exposure, which dominate the decreasing doses received by the fleeing civilians. Rescuers move these civilians away from radioactive material. Since the marginal dose received by all civilians decreases as they move away from the radioactive source, the blue curve in Fig. 2 shows a decreasing slope that approaches zero when all living civilians are far enough from the source of radiation. This suggests that rescuers have a particularly crucial role to keep civilian radiation exposure under a safe threshold.

Figure 3 supports our findings from Fig. 2, but from the perspective of the maximum dose of radiation. When first responders provide assistance, living civilians receive a dose of around 60 REM at most, which is considered a "mild" dose. Without the help of first responders, some civilians receive a steadily-growing dose of radiation, and eventually, they could receive a much more severe or even "lethal" dose. Therefore, first responders are important to avoid unnecessary casualties.

Figures 4, 5, 6 and 7 display results of a simulation under a single run with a fixed seed. Figures 4 and 5, comparing scenarios with and without the assistance of first responders, affirm the findings from Figs. 2 and 3. On the other hand, Figs. 6 and 7 measure the risk undertaken and exposure received by first responders.

Figure 6 features a few cusps in the curve, as labeled. After the rescuers are dispatched at two minutes after the explosion, their distance to the source decreases over time. While observing the simulation, we saw that the rescuers arrived at those injured civilians made immobile by the blast in three different "batches." The slope of the curve in Fig. 6 is the average rate at which first responders receive radiation.

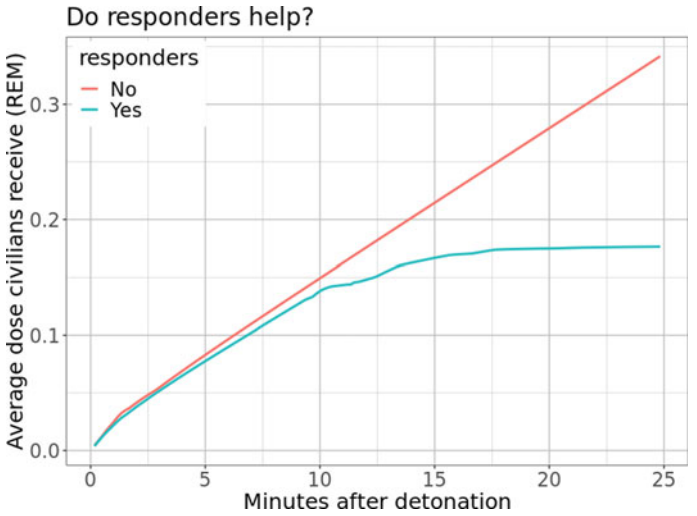


Fig. 2 Average dose of radiation received by civilians over time from the explosion to 25 min after, with and without the presence of first responders. The values plotted are the median of 100 trials

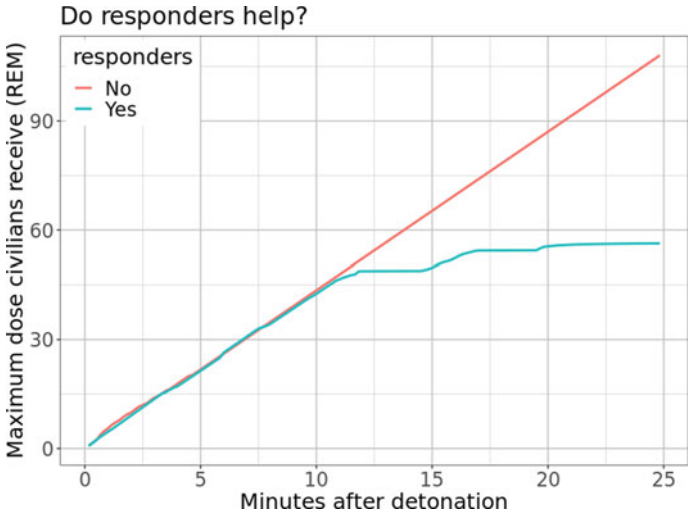


Fig. 3 Maximum dose of radiation received by civilians over time from the explosion to 25 min after with and without the presence of first responders. The values plotted are the median of 100 trials

We found that the sharp cusps where the curve suddenly becomes steep represent the moments at which each batch of rescuers approached the injured civilians; the rescuers' dose rate increases as they get closer to the source. As the rescuers retrieve injured civilians and bring them out, they continue to receive radiation but move at a

Fig. 4 Average dose radiation received by civilians over time for 100 min after the explosion with and without the presence of first responders. The values plotted are from the fixed seed run

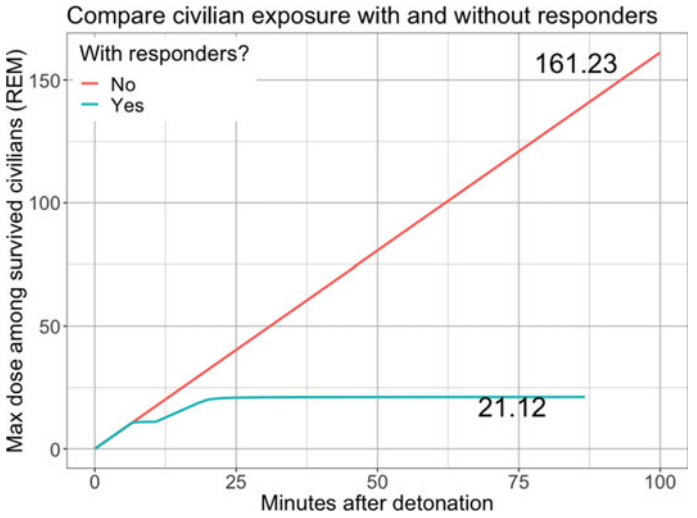


Fig. 5 Maximum dose of radiation received by civilians over time for 100 min after the explosion with and without the presence of first responders. The values plotted are from the fixed seed run

substantially slower speed due to the civilians they are carrying. If there was a way for rescuers to leave the area more quickly after having rescued injured civilians, that would be beneficial.

As shown in Fig. 7, the maximum dose first responders receive went up to 6.37 REM, more than the dose that people working with radioactive sources are allowed

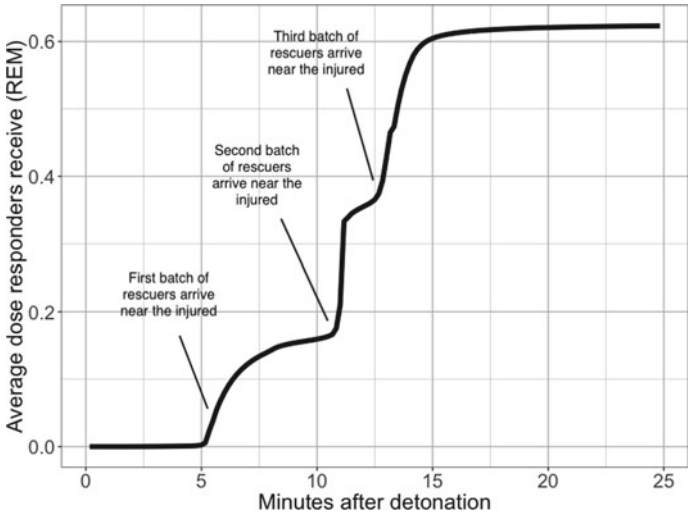


Fig. 6 Average dose of radiation received by first responders over time for 25 min after the explosion. The values plotted are from the fixed seed run

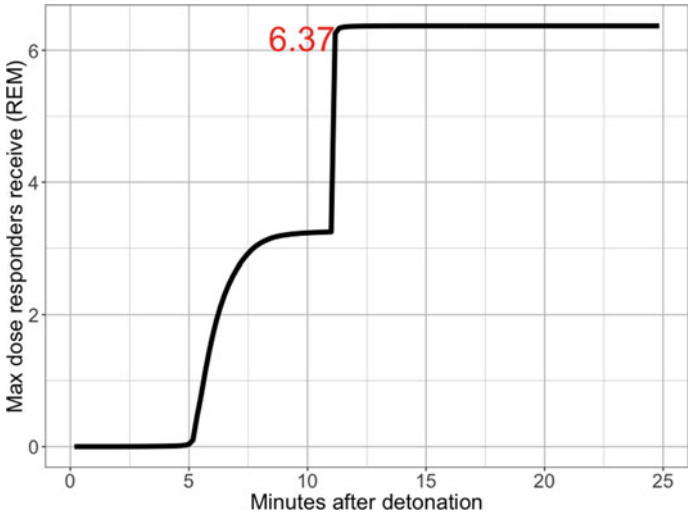


Fig. 7 Maximum dose of radiation received by first responders over time for 25 min after the explosion. The values plotted are from the fixed seed run

to receive in a year. While observing the simulation, we found that the emergency responder receiving the most radiation was an usher who happened to cross through the Hot Zone-at one point, they were less than 1 patch away from the source-while moving in the area. This result is a limitation of the model, but it also emphasizes the importance of thorough training for first responders and prompt communication to first responders of the location of the explosion in a real life dirty bomb event.

4.2.2 Arrival Time of First Responders

Figures 8 and 9 both suggest that it would be best if first responders were able to arrive and begin providing assistance four minutes after the explosion, since that is the scenario in which both civilians and first responders receive the least average dose of radiation.

The disadvantage of first responders arriving sooner is that more civilians are queued up at exits, and the process of screening and decontamination takes time, subjecting civilians to more exposure within the first 100 min. The process of screening and decontamination is crucial, since without it, civilians may suffer continued radiation from radioactive dust lingering on their clothing and skin. Arriving too late would put civilians made immobile by the blast at risk of severe injury and death. To avoid subsequent continuing exposure, we should screen and decontaminate as many civilians at site as possible. For this purpose, it helps to set up exit locations and dispatch ushers as soon as possible.

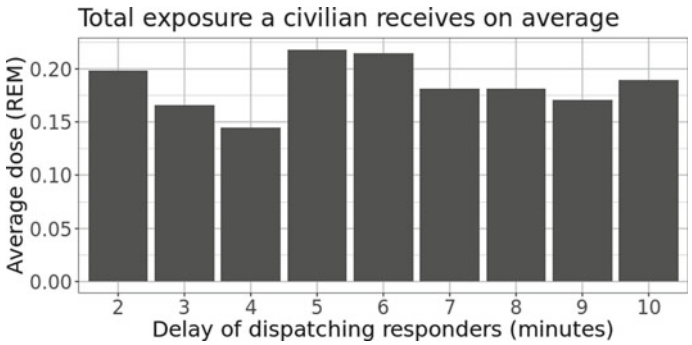


Fig. 8 Average dose of civilians accumulated up to 100 min after the explosion when varying the time of setting up exits and dispatching ushers. The median value over 100 runs is plotted

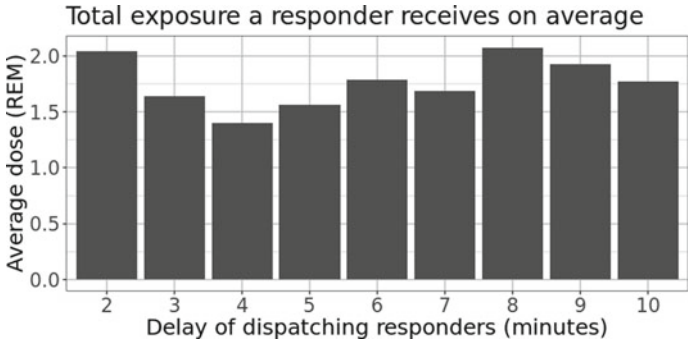


Fig. 9 Average dose of first responders accumulated up to 100 min after the explosion when varying the time of setting up exits and dispatching ushers. The median value over 100 runs is graphed

Interestingly, Fig. 9 shows that first responders receive a greater dose of radiation when arriving before than at four minutes after the explosion. First responders are impeded by congestion, as when they arrive sooner, there are more civilians at the scene. Therefore, the time they spend executing their tasks is longer than in situations where they arrive later.

Figures 8 and 9 demonstrate a tradeoff between the wellbeing of first responders and the wellbeing of civilians, a potent subject for consideration.

5 Conclusion

The presence of first responders substantially decreases both the average and the maximum doses of radiation civilians receive. Without first responders, the average dose civilians receive can go up to half a REM, which is the borderline maximum amount of radiation civilians are allowed to receive per year [12]. With first responders at the scene, however, the average dose civilians receive is significantly lower. As seen in Figs. 3 and 5, the maximum dose civilians receive is also significantly lower when first responders provide assistance.

In the fixed seed run, the maximum dose received by first responders went up to over 6 REM, which is greater than the amount that people working with radioactive sources are permitted to receive per year [12]. If more than one dirty bomb event were to happen in a year, one emergency responder would not be able to respond to multiple events. This suggests a need for concern regarding the safety of emergency responders in dirty bomb events.

Subjects for consideration include: (i) the tradeoff between the wellbeing of civilians and that of emergency responders; (ii) the potency of training and on-site communication.

Although these results pertain specifically to the National Mall area in Washington, D.C., our model can be easily applied to other metropolitan areas.

5.1 Limitations and Further Research

Many pathways are available to expand our model:

Victim volunteers. A variable this model did not consider is victim volunteers, i.e., "civilians" in the explosion who may be off-duty professionals or may possess medical training and will try to help the people around them. It could be interesting to study whether the presence of victim volunteers would impact the severity of injuries or rate of evacuation.

Allocation of resources. The proportion of search-and-rescue responders and exit-monitoring responders is an intriguing topic, as not every combination of responders would have the same level of success. With first responders being an essential

resource, it is important to know how to assign tasks among that resource for the most benefit.

Hiding behind buildings. Civilians attempting to create a solid barrier between themselves and the explosion may flee into the "shadow" of the Washington Monument. Modeling the fleeing of civilians behind buildings and the protection buildings offer from radiation would be challenging.

Size of area. The area in the National Mall that this model is based on is only one thousand meters by five hundred meters, which is rather small. Civilians may flee beyond this area, and decontamination stations may also be set up outside of this area. An excellent route for further study is expanding the size of the area covered by our model.

Effect of communication. Having some form of communication among first responders or between first responders and civilians might reduce their work or amplify their impact.

Acknowledgements We would like to thank George Mason University's Center for Social Complexity as well as the Aspiring Scientists Summer Internship Program for their assistance.

References

1. Biancotto, S., Malizia, A., Pinto, M., Contessa, G., Coniglio, A., D'Arienzo, M.: Analysis of a dirty bomb attack in a large metropolitan area: simulate the dispersion of radioactive materials. J. Instrum. **15**(02) (2020)
2. Center for Disease Control and Prevention (CDC).: Emergency management pocket guide for clinicians (n.d.). Accessed 12 July 2022 from https://www.cdc.gov/nceh/radiation/emergencies/pocket.htm
3. Davis, W.: New York put out a PSA in case of a nuclear attack, leaving many residents confused. National Public Radio (2022). Accessed 19 Aug 2022 from https://www.npr.org/2022/07/13/1111281998/nyc-nuclear-attack-video-psa-eric-adams
4. Department of Energy (DOE).: Radiological control technician training. DOE-HDBK-1122-99 (1999). http://www.eh.doe.gov/techstds/standard/hdbk1122/rad.html
5. Department of Homeland Security (DHS).: Radiological dispersal device (RDD) response guidance: Planning for the first 100 min (2017). https://www.dhs.gov/sites/default/files/publications/nustl_rdd-responseplanningguidance-public_28oct2021-508-revised.pdf
6. Duan, Z., Feng, X., Qin, B., Zeng, C., Yuan, T.: Simulation study on diffusion of smoke cluster from dirty bomb explosion in subway station. Radiat. Prot. (Taiyuan) **36**(5), 279–284 (2016)
7. International Atomic Energy Agency (IAEA): The Radiological Accident in Goiânia (1988). STI.PUB.815 ISBN 92-0-129088-8
8. International Atomic Energy Agency (IAEA).: Manual for first responders to a radiological emergency (2006). https://www-pub.iaea.org/MTCD/publications/PDF/epr_Firstresponder_web.pdf
9. Luckey, T.: Nuclear triage and the dirty bomb. Radiat. Prot. Manag. **20**(1), 11–17 (2003)
10. Manley, M., Kim, Y.S., Christensen, K., Chen, A.: Airport emergency evacuation planning: an agent-based simulation study of dirty bomb scenarios. IEEE Trans. Syst., Man, Cybern.: Syst. **46**(10), 1390–1403 (2015)
11. National Park Service.: Frequently asked questions—national mall and memorial parks (2020). Accessed 22 July 2022 from https://www.nps.gov/nama/faqs.htm

12. Nuclear Regulatory Commission (NRC).: Backgrounder on biological effects of radiation (2020). Accessed 12 July 2022 from: https://www.nrc.gov/reading-rm/doc-collections/fact-sheets/bio-effects-radiation.html
13. Occupational Safety and Health Administration (OSHA).: Radiation emergency preparedness and response (n.d.). https://www.osha.gov/emergency-preparedness/radiation/response
14. Park, J.: The economic impacts of dirty bomb attacks on the Los Angeles and Long Beach ports: applying the supply-driven NIEMO (National Interstate Economic Model). J. Homel. Secur. Emerg. Manag. **5**(1) (2008)
15. Pereira, J.F., Delgado, J.U.: Dirty bomb radiological simulations: two explosion scenarios using the Rio 2016 Olympic games athletes' village as a model. Braz. J. Radiat. Sci. **6**(2) (2018)
16. Rosoff, H., Von Winterfeldt, D.: A risk and economic analysis of dirty bomb attacks on the ports of Los Angeles and Long Beach. Risk Anal.: Int. J. **27**(3), 533–546 (2007)
17. Stone, R.: Dirty bomb ingredients go missing from Chornobyl monitoring lab. Science **376**(6588), 12–13 (2022)
18. Tang, Z., Li, Y., Hu, X., Wu, H.: Risk analysis of urban dirty bomb attacking based on Bayesian network. Sustainability **11**(2), 306 (2019)
19. Weinhold, B.: Emergency responder health: what have we learned from past disasters? (2010). https://doi.org/10.1289/ehp.118-a346
20. Wikipedia: Preferred walking speed (n.d.). https://en.wikipedia.org/wiki/Preferred_walking_speed
21. Wikipedia: Transition from walking to running (n.d.). https://en.wikipedia.org/wiki/Transition_from_walking_to_running
22. Wilensky, U.: Netlogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL (1999)
23. Xiong, K., Kennedy, W.G.: Modeling exposures from a "dirty bomb". Poster presentation at the Conference of The Computational Social Science Society of the Americas (2021)

Replacing Diamond-Dybvig

John S. Schuler

Abstract Diamond and Dybvig 1983 is a now classic model of banking failure. This model and the considerable ancilliary literature studies two equilibria: the "good" equilibrium of bank stability and the "bad" equilibrium of bank failure. A major limitation of these models is that while they acknowledge the fact of these two equilibria, they are silent on how a system in the desired equilibrium suddenly moves into the run equilibrium. Agentization refers to the process of taking usually classic models, economic or otherwise, and representing them in agent-based simulations that hopefully reproduce those model's central features. I consider an agentized Diamond-Dybvig model that reveals some major conceptual limitations in Diamond-Dybvig that limits its utility as the foundation of agent-based studies of bank runs. Then I present an alternative bank run model that may provide such a basis not only for the study of bank runs but also for broader models of financial contagion.

1 Introduction

Diamond and Dybvig [2] is a classic model of bank runs and often cited as a justification for government deposit insurance. Bank runs are an obvious example of the economic problem of self-fulfilling prophecies and present major external costs. Further, the Diamond-Dybvig model is often used as an argument for "inherent bank fragility". One challenge in the study of bank runs is that it has often been suggested they have a psychological elements; that is, so long as depositors believe they can get their money, the don't want it. This fits uncomfortably with traditional economic conceptions of agents. On the other hand, this can be taken as a simple fact of agent behavior and so modeled. We will see that perhaps techniques based on Von-Neumann- Morgenstern expected utility are not the best modeling tool for this purpose.

J. S. Schuler (✉)
Kukun Home Investment Intelligence, Menlo Park, CA, USA
e-mail: jsschuler@gmail.com

Z. Yang and S. Núñez-Corrales (eds.), *Proceedings of the 2022 Conference of The Computational Social Science Society of the Americas*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-031-37553-8_7

2 A Brief Description of the Diamond-Dybvig Model

It is challenging to fit the concept of a bank run into a traditional economic framework. Diamond-Dybvig represents an important attempt to do this. Simplified presentations of the Diamond-Dybvig model can be found in [1] or [3]. My brief exposition will be based on [3]. The model has two periods of time. Agents initially have a certain endowment that they wish to invest in some production technology with a known return at $t = 2$. Agents come in two types, 1 and 2 according to whether they will live until $t = 2$. Neither the agents nor the bank know their type. Thus, the "bank" in this model also serves as an insurance agency in the sense that the agents that will not live until $t = 2$ may withdraw their deposit plus some agreed upon premium. However, this comes at the expense of the return the agents receive at $t = 2$. Since the agents do not know how long they will live, they accept this lower return as an insurance premium. The difficulty comes when too many agents withdraw in at $t = 1$. If agents believe that there will be nothing left at $t = 2$, they will also withdraw early. Thus, the model contains two equilibria. If all other agents are withdrawing, it is always rational for the agent to also withdraw and gamble on a favorable place in line. On the other hand, if only the short-lived agents withdraw, all other agents stay the course.

3 Price Theoretic Criticisms of the Model

White criticizes this model in [3] from a price theoretic point of view. Firstly, the investment is an odd hybrid of debt and equity. I will return to this. The DD bank has no separate class of equity holders which can insulate depositors from losses. Its total debts always exceed its equity. This is relevant as there are actual historical examples of bank failures where all depositors were paid in full. Further, real world banks can suspend note redemption. In Diamond-Dybvig this interferes with consumption. In the real world it may or may not. Note though that these are criticisms of the model's *policy implications*. They do not address the aptness of Diamond-Dybvig as a description of bank-run dynamics.

4 Limitations of Price Theoretic Banking Models

In the real world, assuming the bank run is a self-fulflling prophecy, the bank is sound until it ceases to be and this is a sudden phase change. Price theoretic models can describe multiple equilibria but the relevant modeling question is *how banks move from one equilibrium to the other*. Within an equilibrium setting, one can distinguish between an illiquid bank and an insolvent one. In a broader financial crisis, the bank's assets no longer have a well-defined value. To study banking and financial crises, it is necessary to move beyond equilibrium models. Agentization is a way to proceed.

5 The Agentized Diamond-Dybvig Model

Since the agentized Diamong-Dybvig model is not the core topic of this short piece, it is sufficient to discuss what the agentization revealed. The Diamong-Dybvig ABM shows that the insurance aspects of the model are what actual drive the behavior. This extends the previously-mentioned debt-equity hybrid. While it is true that agents prefer to wait for the investment to mature, once some agents have withdrawn, agents often prefer to gamble on their place in line rather than the investment return. In real life banking, interest acrues and there is no bonus for withdrawing. Thus, this model artifact is driving the behavior. More fundamentally, agents are not buying debt or an equity; rather they are buying an *option*.

6 A Replacement Bank Run Model

In economics, it takes a model to beat a model. I have argued both the necessity of bank run ABM's and that the classical Diamond-Dybvig Model is not a good basis for the same. Thus, it is necessary to provide another model, ideally similar to the Diamond-Dybvig model but without the artifact driving the behavior. Before launching into the model description, it is useful to reconsider the process of banking.

Depositors want to borrow long-term and lend short-term. This is only possible where another entity is willing to take the other side of that deal. The result is an entity with short-term liabilities and long term assets. Let us attempt to adhere as closely as possible for the Diamond-Dybvig framework; accepting the investment aspect as well as endowing agents with a Von-Neumann-Morgenstern utility function. On the other hand, for the sake of simplicity, it is useful to ignore cash flow around the economy. Thus, agents will save for retirement and earn interest but will not borrow.

6.1 Brief Model Description

The replacement model has two types of agents. Both agents have a target endowment. Some agents are savers aiming to retire at 70 ticks. While their spending behavior displays some randomness, they have a target endowment calculated on the basis of the ability to consume from age 70–100. There is no uncertainty in agent life-spans. Other agents are spenders who do not save for retirement and have a target endowment that is constant over their lifetime. They earn their income their entire life. In both cases, agents earn interest on deposits. The savers build weath over time. The spenders do not. This is intended to capture a realistic aspect of depositors. Not all depositors have the same time horizon.

In each tick, agents run simulations to decide whether they are better off, in expectation, staying in the bank or withdrawing. There is no initial period where they decide

how much to deposit as there is no distinction between deposits and endowments. Whenever an agent withdraws, all other agents reconsider their position. The tick ends either with a bank failure or with the remaining agents acruing and interest payment. At the next tick, all agents resume banking. Thus, the cost for those who withdrew is one tick worth of interest payment.

6.2 Model Pseudo-Code

```
Initialize Agents
For 1 to Agent Count
      initialize Agent with endowment, risk preference, and
      probability parameter

Initialize Model
continue = TRUE
while continue
      Random Sort Agents
      continue = FALSE
      agents withdraw exogenously
      for agent in remaining agents
            agent decides whether or not to withdraw
            if withdraw continue=TRUE
for agent in remaining agents
      pay interest
for agent in all agents
      if agent is spender then pay income
      if agent is saver and age <=70 then pay income
      spend money
      age agent one year
      if agent age is 100, remove agent and generate new agent
```

6.3 Model Behavior and Lessons

In this particular model, there is almost always a bank run within a few ticks. The spending agents in particular face a very consequential ruin probability. This raises an interesting question: to the extent agents do not use banks to build wealth, there must be a greater downside to banking. This is consistent with the fact that poorer people are much more likely to not use banks. It also suggests that a follow up model should be even simpler than this model. Such a model would not involve incomes or interest payments but simply model the marginal decision of each agent to withdraw

or not. Then, the background parameters associated with various equilibria can be studied as part of nesting this model in a richer model.

7 Conclusion

These models both feature a single isolated bank but the techniques developed could be extended to networks of banks. The main requirement to extending this to a broader financial system is to explicitly model cash flow and lending behavior; perhaps borrowing tools from percolation theory. However, keeping such rich models in a kind of equilirbium is a challenge. The way forward is to associate desireable equilbria with a set of initial parameters and then make use of machine learning methods to calibrate the richer model toward a model realization consistent with the initial parameter space of the simpler model.

References

1. Diamond, D.W.: Banks and liquidity creation: a simple exposition of the diamond-dybvig model. FRB Richmond Econ. Q. **93**(2), 189–200 (2007)
2. Diamond, D.W., Dybvig, P.H.: Bank runs, deposit insurance, and liquidity. J. Polit. Econ. **91**(3), 401–419 (1983)
3. White, L.: The Theory of Monetary Institutions. Wiley-Blackwell (1999)

Higher-Order Interactions in ABM: A Case Study Using Topologically-Perturbed Voter Models

Santiago Núñez-Corrales, Rajesh Venkatachalapathy, Jeffrey Graham, and Srikanth Mudigonda

Abstract Using variants of the voter model, and inspired by simulations of such models on networks, we studied a variety of ABM implementations using a *random activation* scheduler incorporating dyadic and higher-order interactions. Our results provide evidence about the dependency of various observables on whether state updates are simultaneous or staggered per model step (i.e., matrix vs. ABM), if interactions are pairwise or higher-order, or if the underlying topology changes even when the abstract specification of the voter model is the same: simulation features usually thought of as computational—even intuitively innocuous- prove to be phenomenologically impactful. We found that average magnetization is largely modulated by the initial state in dyadic voter models, that exit probability is controlled by network and simulation types, and that interaction types divide consensus times except for 2D regular lattices, which exhibit surprising sensibility to these perturbations. In addition, regular lattices appear to contain spatio-temporal alternating motifs once certain magnetization values are reached, similar to gliders in Conway's Game of Life. Our findings suggest that ABM simulation workflows must incorporate multiple interaction types and spatial configurations in order to tease our robust findings from either implementation-dependent artifacts or misspecified models, guided by robust statistical physics principles.

SPEC Collaborative.

S. Núñez-Corrales (✉)
NCSA, University of Illinois Urbana-Champaign, Urbana, IL, USA
e-mail: nunezco2@illinois.edu

R. Venkatachalapathy
Systems Science Graduate Program, Portland State University, Portland, OR, USA
e-mail: venkatr@pdx.edu

J. Graham
Department of Mathematical Sciences, Susquehanna University, Selinsgrove, PA, USA
e-mail: graham@susqu.edu

S. Mudigonda
School for Professional Studies, Saint Louis University, St. Louis, MO, USA
e-mail: srikanth.mudigonda@slu.edu

Z. Yang and S. Núñez-Corrales (eds.), *Proceedings of the 2022 Conference of The Computational Social Science Society of the Americas*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-031-37553-8_8

1 Introduction

Simulations of dynamical systems rely on the availability of numerical recipes. Such recipes are either constructed by applied mathematicians using mathematically rigorous schemes built upon foundations in approximation theory, or by scientists and engineers using ad-hoc ansatzes and schemes that are not mathematically accurate but meet empirical rigor. Typically, the two approaches refine each other's techniques, with scientists and engineers developing more easy-to-use versions of formal methods, and applied mathematicians developing rigorous proofs and justifications of heuristic procedures. Our current work takes a first step to do something similar for agent-based simulations.

Specifically, we focus in this article on how different types of interactions and different implementations of state update policies per time step modify the trajectories and final states of the voter model. To tease out differences, we perturb these variants of voter models by changing the initial state of the system and the underlying topology connecting the agents. We compare this dependency on choice of interaction and update types in ABM simulations to choice of numerical approximation schemes used in simulations of stochastic and deterministic dynamical systems, and contend that ABM simulation workflows consciously involving variation of multiple features will make the simulations more robust, and increasing the chances of retaining fidelity to social phenomena under consideration.

In the next section, we provide motivating background and our main conceptual argument. Following this, we present a variant of *majority rule*-based voter model, and embed it in both regular lattices as well as heterogeneous networks. We introduce alternative voter model implementations that update agent states synchronously, as well as higher-order action based on three-way interactions among a focal agent and its neighbors; all variants of the voter model utilize *random activation* scheduling. Later, we present the simulation protocol and present our results in terms of *average magnetization*, *exit probability* and *consensus time*. The discussion section provides our current hypotheses for the observed differences. Finally, our conclusions explore future experiments, implications of our findings, and the necessity of incorporating multiple computational and phenomenological features into ABM simulation workflows guided by robust experimental designs.

2 Background and Rationale

The computational roots of ABM [32] and its wide user-base make what we set ourselves to do difficult. On the foundational side, unlike models based on ordinary differential equations (ODEs) and stochastic differential equations (SDEs), ABMs are computational beasts. In such models, achieving rigor entails a coherent and integrated approximation theory that is not only mathematically sound, but also conforms to theories of programming languages and computation. On the pragmatic side, the

interdisciplinary community relies on an improvisational approach that defies systematization, but is broadly anchored around whether the models are simulated on a particular platform (for example, NetLogo [31]) or are custom-built in a general purpose programming language. Here, we focus on one particular aspect of simulation workflows: the choice of features thought of as computational, or usually safe to vary. Schedulers, activation rules, interaction types and the topology of the space exemplify some of the possible choices.

The particular arguments made here build upon past work (see references cited in [28]) and are part of our on-going project to build a new ABM platform [21]. The main thrust of our project is to build an elastic and scalable ABM platform that is capable of maintaining fidelity to social processes underlying observed macroscopic phenomena. Of relevance here are issues surrounding fidelity to social reality, which demands that the depth of detail of real-world social processes is matched in simulations, being enabled by the platforms running those simulations. At the modeling level, simulation platforms must allow specifying various aspects of social processes like synchronicity, asynchronicity, parallelism, concurrency, etc.—some parts or all of which are handled by the simulation scheduler. This fact is typically acknowledged by developers of ABM platforms (see our discussion in [21]), and more prominently for NetLogo in [32]; however, it is not acknowledged that schedulers can indirectly specify interactions.

For example, most practitioners use default *random activation*-based schedulers to run their simulations. In such schedulers, the platform randomly picks a focal agent, updates its states according to the model specified rules, and moves on to the next randomly chosen agent. Typically, updates to a focal agent are dependent on another agent's state and possibly on the environment and other agents in the agent's neighborhood. The scheduler the moves onto the next randomly chosen agent. Once all the agents' have been activated and acted upon, the scheduler repeats another cycle - these cycle continue until a pre-specified condition (e.g., certain number of time steps have been completed, a certain state of equilibrium in the agent-environment has been reached, etc. For a scheduling protocol to work, several properties of agent-agent interactions and agent-environment interactions have to hold: the initial randomly-chosen order of agents does not matter; the interactions are simple enough to remain homogeneous in time; simultaneous updates of more than one agent is disallowed by the interactions; multiple agents don't participate in higher-order interactions; and so on.

To put it in concrete sociological terms, this scheduler assumes that the social interaction protocol is of a certain kind, in contrast to documented, i.e., real-world biological and social processes (See references cited in [21]), where some agents update their states concurrently, while some others update their states asynchronously. Some platforms offer options for staggered activation schedules [18]. Usually, the default provided by a particular platform, where the simulation is implemented, is chosen and alternative schedulers are not tested for whether they generate any meaningful differences in the end results. Users of ABM platforms, in using a default scheduler, may not realize that that their models may be approximated or misspecified, rather than faithfully implemented by the platform.

In other dynamical systems simulation literature, analogs of such problems and resolutions of brittleness of numerical recipes exist (See references cited in [21]). For example, for certain classes of *stiff* differential equations, special numerical recipes are used instead of Runge-Kutta family of techniques; for certain SDEs, depending on the specific nature of the problem, either Ito-based or Stratanovich-based calculus is used for developing numerical algorithms; and also for certain SDEs, the precise nature of the noise term decides the specific algorithms used; and finally, in spatio-temporally extended stochastic dynamic systems, update rules are provided by following the rules of the system's underlying graphical models [19].

As evidence of the non-triviality of the problem, and partly continuing past work [28], we present a simulation study using variants of voter models on regular and heterogeneous lattices (networks). As we report in the subsequent sections, even when the state space and the dynamics are simple, interaction and implementation choices under a *random activation* scheduler lead to dynamics that are unpredictable not only for the same lattice structure, but also for different parameters within the same lattice type.

3 Dyadic and Higher-Order Voter Models

The classic voter model [24] is an interacting particle system [2] where N agents (voters) reside on a complete graph—agents can communicate and interact with the entire population of voters. Each agent has one of two states $\{-1, +1\}$, an *opinion*, which they update by: (a) picking an agent uniformly at random, (b) adopting the state of this random neighbor, and (c) repeating the steps until consensus is reached. One such variant is the majority rule, where the voter adopts the state of the majority opinions among its neighbors. We adopt a variant of majority-rule-based voter model by restricting ourselves to picking a fixed subset of neighbors, and adopting the majority opinion of the subset with a pre-specified probability.

The basic observables of this model correspond to the *average magnetization* $\langle M \rangle$, the *exit probability* $E(m_0)$ that the entire population reaches the $+1$ consensus as a function of initial magnetization $m_0 \equiv (N_{+1} - N_{-1})/N$, where N_{+1} and N_{-1} are the initial number of $+1$ and -1 voters respectively, and the *consensus time* $T(m_0)$, the average time for each consensus as a function of m_0. Computationally, an agent i agents undergoes discrete time-dependent transformations of its state $s_i(t)$. To facilitate computation of observables described above, $s_i(t)$ follows the mapping

$$\begin{aligned} +1 &\mapsto 1 \\ -1 &\mapsto 0. \end{aligned} \tag{1}$$

Our interest in this work is to determine whether the observables of simple dyadic and higher-order voter models behave differently under different topological and implementation perturbations.

3.1 Dyadic Voter Model

In a dyadic voter model, agents interact pairwise depending on the adjacency matrix that captures the topology of the network containing them. We are then interested in two main quantities: the average predominance of a +1 opinion with N agents

$$\langle f(t)\rangle = \frac{1}{N}\sum_{i=1}^{N} s_i(t) \tag{2}$$

and the predominance of opinion +1 across the neighborhood of an agent i,

$$f_i(t) = \frac{1}{n_i}\sum_{j\in n_i} s_j(t). \tag{3}$$

Using (2), one can recover the average magnetization of a particular voter model by computing

$$\langle M(t)\rangle = 2\langle f(t)\rangle - 1, \tag{4}$$

which ensures $\langle M(t)\rangle \in [-1, +1]$.

For our majority voting strategy σ for time $t+1$, we define the state transition function

$$s_i(t+1) = \begin{cases} 1, & f_i(t) \geq 0.5 \\ 0, & f_i(t) < 0.5. \end{cases} \tag{5}$$

As observed from (3) and (5), this model definition pertains exclusively to dyadic interaction models, similar to how inter-atomic potentials are computed in classical molecular dynamics [30].

3.2 Higher-Order Voter Model

Recent literature has shown significant model differences depending on whether simulations portray dyadic and higher-order interactions [22]. Consequently, we explore in this work the effects of higher-order interactions on the observables described above. To do so, we rework (5) by assuming that partitioning the set of N agents

by k—albeit possibly imperfectly- leads to $\lceil N/k \rceil$ random groups during each time step whose opinion will be updated simultaneously, according to an updated majority rule. Each k-set of agents is therefore a hyperedge in a corresponding hypergraph. Contrary to the work of Papanikolaou et al. [22], our model is not adaptive, since it assumes that hyperedges remain the same size across simulation time (i.e., no coalescence of hyperedges with homogeneous opinions). Hence, (5) becomes

$$\overline{s_\alpha}(t+1) = \begin{cases} 1, s_\alpha(t) + f_\alpha(t) \geq 0.5 \\ 0, s_\alpha(t) + f_\alpha(t) < 0.5. \end{cases} \tag{6}$$

where α denotes the centroid of the ℓ-th k-group of voter agents and $\overline{s_\alpha}$ the simultaneity of the update for all agents connected to the hyperedge centered at α, $\ell \in \{1, \cdots, \lceil N/k \rceil\}$.

4 Case Study: Differences Under Various Perturbations

Given the simple nature of the models above, we focused on two types of perturbations: (a) changing the underlying topology that defines neighborhoods, and (b) changing the implementation of the simulation based on adjacency matrices and state vectors, and a fully-fledged ABM. In this section, we review the experimental context and the results obtained with our code base.

4.1 Software Implementation

During software development, we built a minimal code base that follows research software engineering principles and practices [9] as a way to ensure code transparency, reproducibility and extensibility. Our code was implemented using Python 3.9, and benefited from `networkx` to simulate various topologies of interest [13] and `Mesa` as the core ABM framework [18]. The code base is open source, available on GitHub, and our analysis is fully reproducible.[1]

Structurally, the `AbstractVoterModel` class defines the properties and methods any voter model must implement and/or call in the course of a single experimental run. The `NetworkEnsembleFactory` class generates ensembles of networks that share the same specification and are used to then produce ensemble computations by the `ModelDriver` class. Instead of storing information in memory or in a CSV file, we opted for database storage as a means to (a) prototype future data storage designs for SPEC-ABM [21], and (b) facilitate data analysis using `R` [23]. We used `SQLite` [6, 15] as our database engine due to the small size of our simulations and the convenience for users. Data storage for agents, ensembles and simulations

[1] See: https://github.com/SPEC-collab/CSSSA2022.

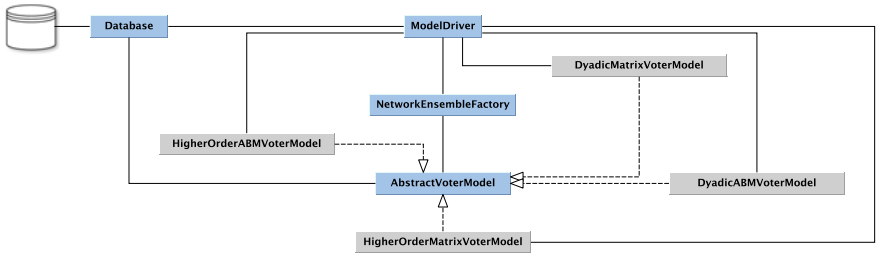


Fig. 1 Software architecture for voter model simulations. Entities in gray indicate specializations, while entities in blue constitute primary classes

is managed by the `Database` class. Figure 1 depicts the resulting software architecture.

The simulation proceeds as follows. Based on input parameters, the `ModelDriver` connects to a new or existing database, stores the parameters as a new simulation with a unique `uuid` as its identifier. Then, the `NetworkEnsembleFactory` produces an ensemble of size K with N nodes corresponding to agents in the voter network. Once the network ensemble is available, the `ModelDriver` iteratively instantiates a model implementing the `AbstractVoterModel` corresponding to the parameters controlling for interaction type and matrix/ABM implementation. During each simulation step, `AbstractVoterModel` commits fine-grained, individual agent information to the database (i.e., `Record` objects), and individual observables (i.e., `Summary` objects) are stored. As each ensemble point finalizes its execution, the `ModelDriver` commits observables. In order to avoid excessive computation, each ensemble point executes until it has converged. Convergence time is tested dynamically by identifying the time-step when the corresponding trace has not varied during the last five consecutive steps.

Our implementation makes use of the *random activation* scheduler for all classes of simulation as a way to maintain comparability of analysis. However, a major difference exists between matrix and ABM implementations. In matrix implementations, agent changes are buffered for each time step, and updated simultaneously once all agents update their state. In ABM implementations, no buffering is provided, and agents update their state in the sequence in which the random activation function dictates. Our rationale for this decision is to make explicit and challenge the assumption that, at least for simple cases, ABMs and matrix implementations should produce similar outcomes.

4.2 Experimental Design

Our central aim in this study was to elicit potential differences across various implementations of voter models. Given that our models produce ensemble computations,

Table 1 List of parameters for the voter model.

Parameter	Range
`simulation_type`	Matrix-based (`matrix`) or ABM-based (`abm`)
`network_type`	Regular 2D grid (`12dr`, periodic), hypercube (`hc`), power-law cluster network (`pl`), or Erdös-Rényi (`er`)
`interaction`	Dyadic (`dyn`) or higher order (`hord`)
N	Population size (`1024`)
`interactants`	Agents per hyperedge (`3`)
`max_steps`	Maximum number of steps per ensemble point (`500`)
`ensemble_size`	Number of ensemble points (`100`)
`initial_state`	Initial number of agents with +1 opinion (`0.25, 0.5, 0.75, 0.8, 0.9, 0.95`)

and that observable of network ensembles vary significantly depending on the type of network [1], we hypothesized that varying the network constituted a significant perturbation capable of maximizing the observed differences since a voter model corresponds to a diffusion process across a lattice, whether ordered or disordered [14]. Such differences manifest, for example, for the lattice Boltzmann equation when irregular lattices are introduced; the irregularity plays the role of a modulator of viscosity [20], and we suspect an analogous phenomenon for voter models manifests, based on prior work on social viscosity using ABM [25]. Similarly, since the value of m_0 determines the initial set point of the system, we hypothesized that exploring a relatively wide range of values would also maximize such differences. The value of N where differences manifest sufficiently, while yielding tractable computation times, was found to be 1024. Population sizes are expected to be of the form $N = 2^{2p}$ for $p > 1$ such that square lattices and hypercubes are correctly generated. Table 1 summarizes parameter choices and their variation.

4.3 Results

Computational experiments were executed in a MacBook Pro, 2.3 GHz 8-Core Intel Core i9 with 32 GB 2667 MHz DDR4 under base memory usage. We observed that Mesa-based implementation was 40x times slower, on average, than the matrix-based implementation using the same simulation structure.

Variety in $\langle M(t) \rangle$ response is maximized by $m_0 = 0.50$ and dyadic interactions. Computation of average magnetization and subsequent faceting reveal a wide diversity of outcomes. First, convergence is modulated by m_0 as expected, yet the magnitude of the response over $\langle M(t) \rangle$ shows unexpected variation between 0.50 and 0.80

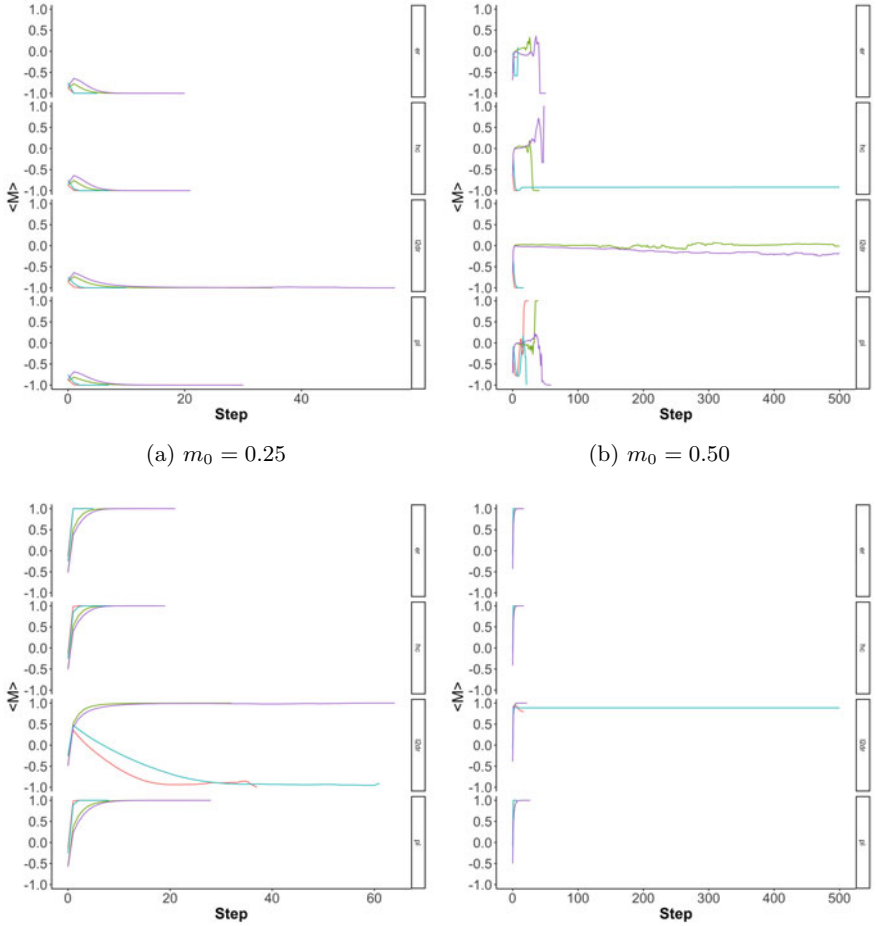
(Fig. 2). Noticeably, simulations using a 2D regular lattice consistently exhibit longer convergence times. For most network types, higher-order interactions (green, purple) and dyadic voter models group together. This is not the case, however, for power-law clustered networks, in which the simultaneity (teal, purple) or non-simultaneity (green, red) of updates partitions final state convergence for the case with the maximally mixed state.

Focusing on the structure of the 2D regular lattice (Fig. 3), convergence time for the matrix dyadic model increases with m_0; while the same appears to occur for the ABM dyadic simulation as well, up to $m_0 = 0.8$. Remarkably, $m_0 = 0.50$ maximizes the convergence time for higher-order models in the 2D regular lattice, and for the matrix dyadic model in the hypercube. Dyadic interactions maximize the diversity of average convergence curves (Fig. 4). Finally, we observe that for the relatively low value of $m_0 = 0.50$, average magnetization oscillates in power-law cluster networks (yellow curve in Fig. 4a, b) with opposite final average magnetization.

Exit probability depends drastically upon network and simulation types We now turn our attention to the exit probability $E(m_0)$ as described by the distribution of $M(m_0)$ once it reaches consensus. First, consider $m_0 = 0.50$ as a point of maximal initial mixing of voter states and hence maximum entropy (Fig. 5). For each one of the simulation types, changing the underlying topology yields significant changes in the final outcome. Dyadic and higher-order models partition data into two different categories of outputs. Regular grids (2D, hypercube) produce almost unimodal distributions for dyadic models, while the rest of the combinations yield multimodal distributions. Higher-order models produce bimodal, (almost) balanced distributions, while four cases yield trimodal ones (i.e., two of them for higher-order models in 2D regular grids, one for the ABM dyadic model in a power-law clustered network, one for the hypercube with a matrix dyadic model).

To better understand these differences across simulation types, we studied the effect of increasing values of m_0 across multiple network types (Fig. 6). Consensus to the +1 state occurs for all simulation types and most networks at $m_0 = 0.75$, except the 2D regular lattice. As expected, $m_0 = 0.50$ contains a collection of final state mixtures, with non-regular networks producing distributions biased toward the –1 final state. Higher-order models in the same initial state exhibit a uniformly distributed mixture of final states. Again, the deceivingly simple structure of the 2D regular lattice shows unexpected variation at or near thresholds depending on the implementation (Fig. 7).

Interaction type separates consensus times except for 2D regular lattices Simulation data indicate that lattice structures (i.e., 2D lattice, hypercube) yield shorter consensus times than power-law clustered and Erdös-Rényi networks (Fig. 8). Dyadic and higher-order models exhibit a clear separation, except for the 2D lattice in which consensus times become mixed between 0.50 and 0.90, with non-normal distributions for all simulation types. This fact, alongside the prior analysis, suggests that 2D regular lattices induce a significantly more complex response than expected, despite being one of the most frequent network type used across ABM research. For non-regular network types, m_0 modulates the average and standard deviation of consensus



(a) $m_0 = 0.25$ (b) $m_0 = 0.50$

(c) $m_0 = 0.75$ (d) $m_0 = 0.90$

Fig. 2 Average magnetization $\langle M(t) \rangle$ per initial state m_0. Color key: teal—matrix dyadic, purple—matrix higher-order, red—abm dyadic, green—abm higher-order. Vertical facets correspond to network types. Plots for $m_0 = 0.8$ and $m_0 = 0.95$ available in GitHub repository

times in a possibly symmetrical manner (e.g., see $m_0 = 0.25$ versus $m_0 = 0.75$) with maximum dispersion at $m_0 = 0.50$.

Regular lattices induce band structures for matrix dyadic voter models Figure 4a contained an intriguing feature that deemed a closer look. The variance for the average magnetization remained constant for the entire 500 simulation steps in the 2D regular lattice at $m_0 = 0.80$ and the hypercube at $m_0 = 0.50$, instead of varying due to some ensemble points achieving consensus earlier than others. The outcome revealed individual agents that could not achieve consensus, yet the magnetization value remained constant in matrix dyadic voter models (Fig. 9). While these trajec-

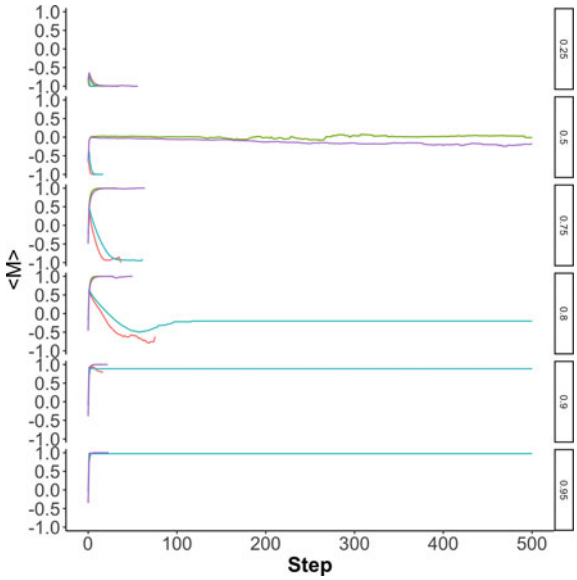
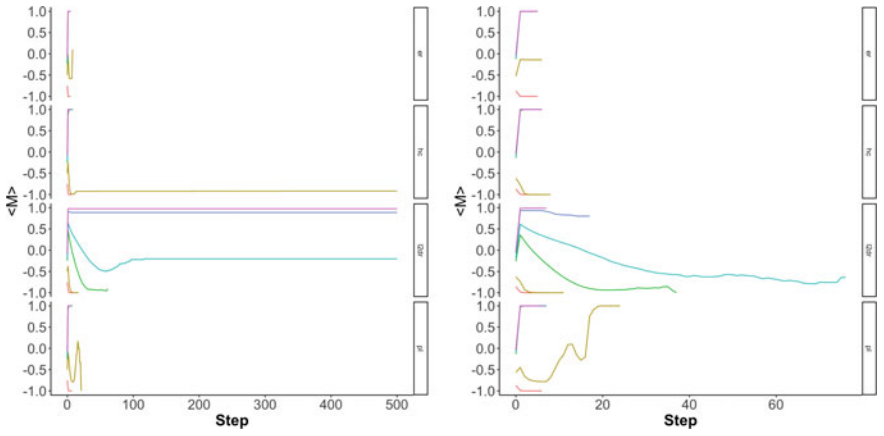


Fig. 3 Average magnetization $\langle M(t)\rangle$ for the 2D regular lattice. Color key: teal—matrix dyadic, purple—matrix higher-order, red—abm dyadic, green—abm higher-order. Vertical facets correspond to m_0



(a) Matrix dyadic (b) ABM dyadic

Fig. 4 Average magnetization $\langle M(t)\rangle$ for dyadic models. Color key: red—0.25, yellow—0.50, green—0.75, teal—0.8, blue—0.90, magenta—0.95. Vertical facets correspond to network types. Plots for higher-order interaction models available on GitHub

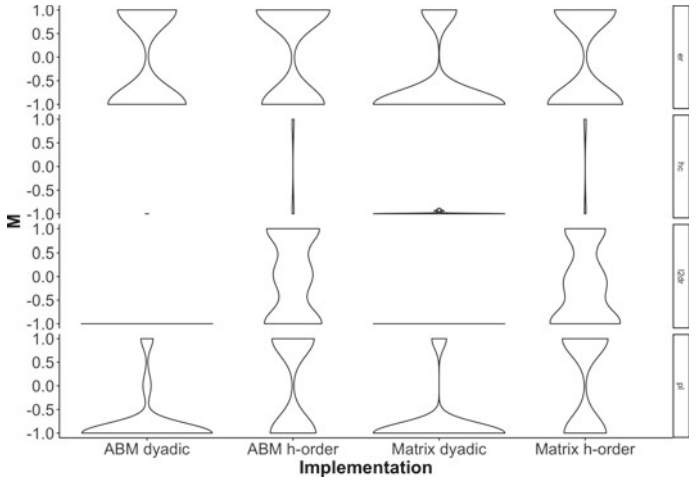
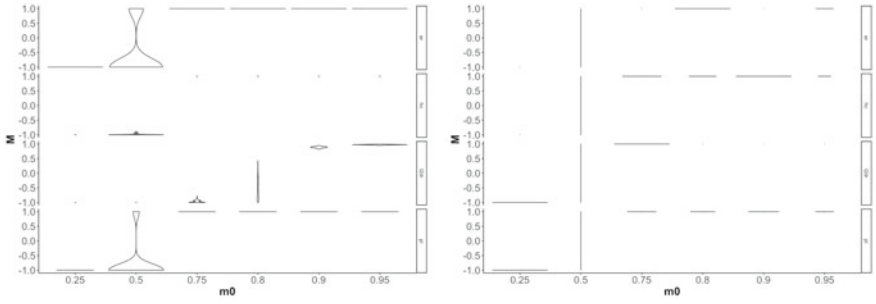
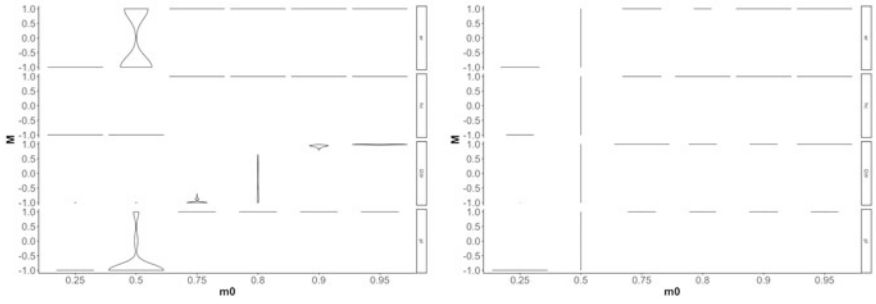


Fig. 5 Exit probability distribution $E(m_0)$ for $m_0 = 0.50$. Color key: teal—matrix dyadic, purple—matrix higher-order, red—abm dyadic, green—abm higher-order. Vertical facets correspond to m_0



(a) Matrix dyadic (b) Matrix higher-order



(c) ABM dyadic (d) ABM higher-order

Fig. 6 Exit probability distribution $E(m_0)$ per simulation type m_0. Vertical facets correspond to network types

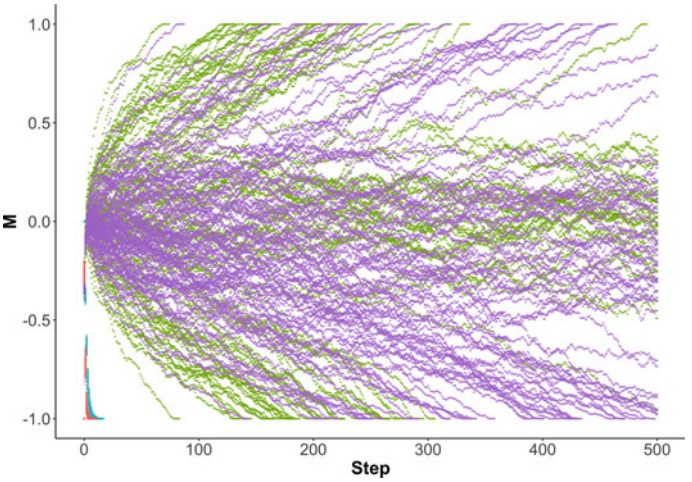
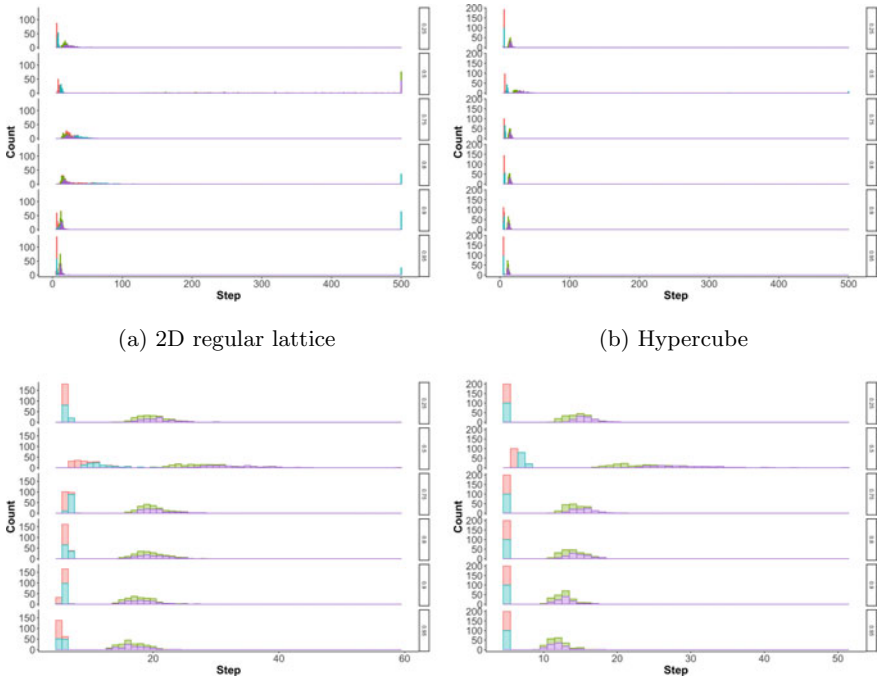


Fig. 7 Exit probability distribution $E(m_0)$ for the 2D regular lattice at $m_0 = 0.50$. Color key: teal—matrix dyadic, purple—matrix higher-order, red—abm dyadic, green—abm higher-order



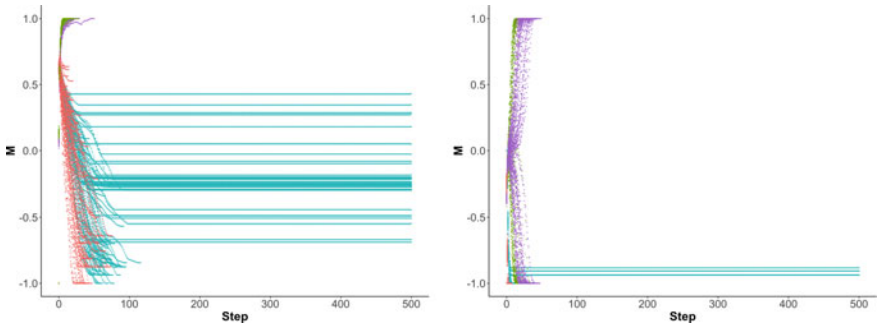
(a) 2D regular lattice

(b) Hypercube

(c) Power-law clustered

(d) Erdös-Rényi

Fig. 8 Average consensus time per simulation type. Color key: teal—matrix dyadic, purple—matrix higher-order, red—abm dyadic, green—abm higher-order. Vertical facets correspond to values of m_0



(a) `l2dr`, $m_0 = 0.80$ (b) `hc`, $m_0 = 0.50$

Fig. 9 Magnetization trajectories for the 2D regular lattice at $m_0 = 0.80$ and the hypercube at $m_0 = 0.50$. Color key: teal—matrix dyadic, purple—matrix higher-order, red—abm dyadic, green—abm higher-order

tories are far more numerous for the 2D regular lattice, they appear to form band gap structures around certain magnetization values.

4.4 Discussion

The experimental results obtained here reveal unexpected and unpredictable simulation differences originating from relatively small changes, most of which give the impression of being innocuous to the trajectory and final outcomes of ensembles and individual points in them. The richness of response to these changes of the 2D regular lattice was most surprising of all since (a) it tends to be the most frequent geometry for spatially-dependent simulations, (b) its constrained structure intuitively suggests that changes in how the simulation is implemented should at most change convergence time but not exit probabilities—2D voter models at the macroscale behave at the thermodynamic limit as a reaction-diffusion system [10, 17], hence changes altering the diffusion coefficient D-, and (c) its structure is strongly constrained, periodic and with a small number of degrees of freedom per lattice site. Consensus time was the only observable for which these differences were less clear across model implementations (Fig. 8a). The only expected feature was slower consensus times on average due to fewer degrees of freedom than the other network used. For the ABM modeling community, these results suggest the need to rigorously evaluate when 2D regular lattices are used under various types of interactions, implementations, schedulers and initial conditions. For this, identifying which parameters are suitable for sensitivity analyses becomes a strategic concern.

In most simulations, dyadic versus higher-order became easily distinguishable (Fig. 6). Introducing higher-order interactions to voter models produced different outcomes for the three observables of interest in our study. Power-law clustered net-

works, however, constitute an exception around the $m_0 = 0.50$ threshold. First, the dominant feature becomes the implementation (i.e., matrix vs. ABM) instead of the interaction for final average magnetization value (Fig. 2b). Even more striking, final average magnetization reverses entirely between matrix and ABM implementations of the dyadic voter model. We hypothesize that the peculiar structure of the power-law clustered network leads to limited percolation of +1 states when simultaneous matrix updates translating into a single diffusion front, in contrast to deeper percolation by individual agents in the ABM which becomes a fractal-like diffusion front [26]. In turn, the latter leads to stronger state correlations as local +1 hubs form progressively in ABM implementations [12]. Moreover, the non-universality [7] and local dependence of diffusion exponents [11] can help explain the pattern observed in Fig. 2b, in which and initially low average magnetization reaches a local maxima, then a local minima, and converges finally to the +1 state; its matrix counterpart exhibits a contrary motion in time. Our results provide an insight into discrete processes analogous or equal to reaction-diffusion in networks whose structure is closer to those present in real social systems.

We observe that the main effect of using ABM versus matrix implementations appears to be one of *time compression* (Fig. 8): even with mixed results found in the 2D regular lattice for $m_0 = 0.75$ and $m_0 = 0.80$, ABM simulations reached shorter consensus times than matrix implementations. On the one hand, simultaneous global updates give the impression of deciding the fate of the system sooner due to an all-or-non local update. Nevertheless, this is not the case. ABM models with incremental agent updates per simulation step suggest that topological dependencies in network structures induce emergent synchronization across the network [3] due to the interplay between structure (i.e., the computation of f_i with respect to each agent's local topology) and function (i.e., computation of $s_i(t + 1)$ in the context of other updates). We hypothesize that simultaneous global state updates, particularly for higher-order interaction models in which hyperedges are randomly selected and thus some agents may not participate in the update rule, may produce alternating opinion motifs that take longer to converge to either –1 or +1 due to varying coverage around certain key agents. Despite the fact that the present study only makes use of the random activation scheduler for the agents, the simultaneity of the update rule suffices to imprint differences across relevant observables in the various voter models.

Finally, the band structures that emerge in regular lattices (Fig. 9) present an interesting challenge. After looking at all ensembles, band structures appear for $m_0 \geq 0.8$ in the 2D regular lattice, and only for $m_0 = 0.50$ for the hypercube, and only occurs for specific values of M. Even when the value of magnetization is fixed, the models do not achieve consensus. We looked carefully at the implementation, and could not find any systematic error capable of reproducing the band structure without altering other results. We thus hypothesize that such band structure results from the appearance of spatio-temporal alternating motifs that, while preventing convergence, allow the simulation to remain at a stable trajectory. Since this particular model is implemented using pairwise interactions and the agent states are minimally simple (i.e., one bit), we believe that these motifs resemble entities such as gliders in Conway's Game of Life, which is know to exhibit self-organized criticality [4].

Contrary to Game of Life, however, voter models do exhibit various conservation laws [8, 16, 27, 29], belonging to a different class of universality in this sense. The difference in the range of values of m_0 in which the band structure occurs for 2D regular lattices versus hypercubes appears much simpler to explain, as the hypercube provides more topological degrees of freedom for agents to avoid becoming trapped in these motifs except when state mixing is at its maximum. We also hypothesize that our observations may be also explained as a result of the distribution of eigenvalues in the Markov transition matrix corresponding to each voter model implementation. Clearly, elucidating the origin and properties of this particular feature requires further experimentation and analysis.

5 Conclusion

In this study, we provide strong evidence indicating that voter model observables vary widely depending on whether state updates are simultaneous or staggered per model step (i.e., matrix vs. ABM), if interactions are pairwise or higher-order, or if the underlying topology changes even when the abstract specification of the voter model is the same. Contrary to common intuition, the largest variation occurs when 2D regular lattices describe the underlying topology of interactions: this observation suggests that voter-like models, and reaction-diffusion models more generally must undergo significant scrutiny. Our approach suggests that, particularly for complex models, experimental design leading to a robust sensitivity analysis becomes essential; this is particularly when the simulation aims to reconstruct social systems with some degree of fidelity to reality.

Our experiments pose several questions for future research. First and foremost, our voter models can be further perturbed by choosing different schedulers, in line with prior research performed by SPEC [28]; based on our current results, we expect new kinds of differences to appear when event ordering per each time step deviates from that provided by *random activation*. Increasing the number of interactants per hyperedge is required to understand the effect of larger random partitions of the network per time step in higher-order models. Achieving more resolution between $m_0 = 0.50$ and $m_0 = 0.75$ is needed to understand the sharp transition observed across exit probability distributions for networks other than the 2D regular lattice. H Van Dyke Parunak (personal commucation) has suggested compting similar results with triangular and hexagonal lattices, since these have interesting properties in Ising models when rescaled by an appropriate renormalization group. Exploring parameter spaces of current and new non-regular network ensembles must be performed in order to better understand percolation differences between matrix and ABM models. Other voter model variants become available by changing number of possible (discrete) opinions agents can have, as well as allowing continuous opinion values. Finally, our voter models can be extended to continuous time using the Gillespie algorithm [5].

Regarding the relation between our experiments and existing theory, we found connections to reaction-diffusion phenomena, fractal dynamics, correlation expo-

nents across regular and non-regular networks, percolation theory, network synchronization, cellular automata and self-organized criticality. The unexpected richness of the space spanned by our study on voter models suggests that devising ABM simulations for social phenomena of the sort explored here (and others) remains a challenging task even for relatively simple cases. The latter indicates that agent states and their interactions are insufficient to derive intuitions about the dynamics of an ABM, and that even with complete information about how a model operates it may not be possible to do so. As the complexity of an ABM increases, computational experiments sample decreasing portions of all possible model trajectories, making brute force approaches much less successful. Hence, our view –and current research direction- is that tighter integration between what is usually thought of as computational in ABM research and statistical physics is not only desirable in principle, but inevitable in practice.

We thank the reviewers and participants in CSS 2022 for their insightful comments and suggestions.

References

1. Albert, R., Barabási, A.: Statistical mechanics of complex networks. Rev. Mod. Phys. **74**(1), 47 (2002)
2. Aldous, D.: Interacting particle systems as stochastic social dynamics. Bernoulli **19**(4), 1122–1149 (2013)
3. Arenas, A., Díaz-Guilera, A., Kurths, J., Moreno, Y., Zhou, C.: Synchronization in complex networks. Phys. Rep. **469**(3), 93–153 (2008)
4. Bak, P., Chen, K., Creutz, M.: Self-organized criticality in the'game of life. Nature **342**(6251), 780–782 (1989)
5. Bernstein, D.: Simulating mesoscopic reaction-diffusion systems using the gillespie algorithm. Phys. Rev. E **71**(4), 041103 (2005)
6. Bhosale, S.T., Patil, T., Patil, P.: Sqlite: Light database system. Int. J. Comput. Sci. Mob. Comput. **44**(4), 882–885 (2015)
7. Bunde, A., Harder, H., Havlin, S.: Nonuniversality of diffusion exponents in percolation systems. Phys. Rev. B **34**(5), 3540 (1986)
8. Caccioli, F., Dall'Asta, L., Galla, T., Rogers, T.: Voter models with conserved dynamics. Phys. Rev. E **87**(5), 052114 (2013)
9. Cohen, J., Katz, D.S., Barker, M., Hong, N.C., Haines, R., Jay, C.: The four pillars of research software engineering. IEEE Softw. **38**(1), 97–105 (2020)
10. Theodore Cox, J., Durrett, R., Perkins, E.: Voter Model Perturbations and Reaction Diffusion Equations (2011). arXiv:1103.1676
11. Gallos. L.K., Argyrakis, P.: Reaction-diffusion processes on correlated and uncorrelated scale-free networks. Phys. Rev. E **72**(1), 017101 (2005)
12. Gallos, L.K., Song, C., Makse, H.A.: Scaling of degree correlations and its influence on diffusion in scale-free networks. Phys. Rev. Setters **100**(24), 248701 (2008)
13. Hagberg, A., Swart, P., Chult, D.S.: Exploring network structure, dynamics, and function using networkx. Technical report, Los Alamos National Lab (LANL), Los Alamos, NM (United States) (2008)
14. Haus, J.W., Kehr, K.W.: Diffusion in regular and disordered lattices. Phys. Rep. **150**(5–6), 263–406 (1987)
15. Hipp R.D.: SQLite (2022)

16. Horstmeyer, L., Kuehn, C.: Adaptive voter model on simplicial complexes. Phys. Rev. E **101**(2), 022305 (2020)
17. Howard, M., Godreche, C.: Persistence in the voter model: continuum reaction-diffusion approach. J. Phys. A: Math. Gen. **31**(11), L209 (1998)
18. Kazil, J., Masad, D., Crooks, A.: Utilizing python for agent-based modeling: the mesa framework. In: Thomson, R., Bisgin, H., Dancy, C., Hyder, A., Hussain, M. (eds.) Social, Cultural, and Behavioral Modeling, pp. 308–317. Springer International Publishing, Cham (2020)
19. Koller, D., Friedman, N.: Probabilistic Graphical Models: Principles and Techniques. MIT Press (2009)
20. Nannelli, F., Succi, S.: The lattice boltzmann equation on irregular lattices. J. Stat. Phys. **68**(3), 401–407 (1992)
21. Núñez-Corrales, S., Friesen, M., Srikanth, M., Venkatachalapathy, R., Graham, J.: In-Silico models with greater fidelity to social processes: towards abm platforms with realistic concurrency. In:Conference of the Computational Social Science Society of the Americas. Springer, Berlin (2020)
22. Papanikolaou, N., Vaccario, G., Hormann, E., Lambiotte, R., Schweitzer, F.: Consensus from group interactions: an adaptive voter model on hypergraphs. Phys. Rev. E **105**(5), 054307 (2022)
23. R Core Team.: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria (2021)
24. Redner, S.: Reality-inspired voter models: a mini-review. Comptes Rendus Physique **20**(4), 275–292 (2019)
25. Salamanca, J., Núñez-Corrales, S.: Social viscosity, fluidity, and turbulence in collective perceptions of color: An agent-based model of color scale convergence. In: Conference of the Computational Social Science Society of the Americas, pp. 191–212. Springer, Berlin (2020)
26. Sapoval, B., Rosso, M., Gouyet, J.-F.: The fractal nature of a diffusion front and the relation to percolation. Journal de Physique Lettres **46**(4), 149–156 (1985)
27. Ángeles Serrano, M., Klemm, K., Vazquez, F., Eguíluz, V.M., Miguel, M.S.: Conservation laws for voter-like models on random directed networks. J. Stat. Mech.: Theory Exp. (10):P10024 (2009)
28. Srikanth, M., Núñez-Corrales, S., Venkatachalapathy, R., Graham, J.: Scheduler dependencies in agent-based models: a case-study using a contagion model. In: Conference of the Computational Social Science Society of the Americas. Springer, Berlin (2021)
29. Suchecki, K., Eguiluz, V.M., Miguel, M.S.: Conservation laws for the voter model in complex networks. EPL (Europhys. Lett.) **69**(2), 228 (2004)
30. Voter, A.F.: Interatomic potentials for atomistic simulations. MRS Bull. **21**(2), 17–19 (1996)
31. Wilensky, U.: Netlogo. Center for Connected Learning and Computer-Based Modeling. Northwestern University, Evanston, IL (1999). http://ccl.northwestern.edu/netlogo/
32. Wilensky, U., Rand, W.: An Introduction to Agent-based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo. Mit Press (2015)

Modeling Macaque Fighting Dynamics with the Evolutionary Model Discovery Framework to Understand Its Application and Utility

Alex Isherwood, Melanie Jutras, Matthew Koehler, David Slater, William Thompson, and Maria Yelenick

Abstract In this study we use a simple case study of macaque fighting dynamics to develop and use the Evolutionary Model Discovery (EMD) framework [7, 8]. Rather than focus on the macaque populations EMD created, we will focus on our use of the EMD framework. Application of the EMD framework is not straight forward and involves making many decisions that may ultimately impact one's results. Here we highlight many of these and explain how we navigated this process.

1 Introduction

Since Axelrod [1] highlighted the utility of simulation within the social sciences, its use has continued to grow. Social systems are, after all, complex; therefore, an efficient way to explore their potential future state is through simulation [3]. Furthermore, the most natural way to capture the important components of a social system is via an agent-based model (ABM) or simulation [2, 5]. This has been characterized as social science from the ground up [6]. However, this process has usually taken the form of a single implementations of "hand crafted" formulations tuned and tweaked until they resemble the social phenomena of interest. This, of course, begs the question: would other model formations also generate the phenomena of interest?

A new field is coalescing to address this criticism: inverse generative social science (iGSS, www.igss-workshop.org). iGSS attempts to create families of models that all

A. Isherwood · M. Jutras · M. Koehler (✉) · D. Slater · W. Thompson · M. Yelenick
The MITRE Corporation, 7515 Colshire Dr, McLean, VA 87507, USA
e-mail: mkoehler@mitre.org

W. Thompson
The Vermont Complex Systems Center at the University of Vermont, Burlington, VT 05405, USA

Z. Yang and S. Núñez-Corrales (eds.), *Proceedings of the 2022 Conference of The Computational Social Science Society of the Americas*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-031-37553-8_9

generate the social phenomena of interest directly from data, from a combination of data and theory, or an articulation of plausible behaviors derived from theory.[1]

In this work we examine one of the iGSS techniques: Evolutionary Model Discovery. This technique was first described by Gunaratne [7] and again by Gunaratne and Garibay [8]. It uses genetic algorithms to build up collections of agent behavior via a genetic program that optimize a fitness function or create output data that closely resembles a reference dataset. Moreover, by utilizing a genetic algorithm, this system creates a population of candidate agent behavior rules that may then be analyzed to highlight commonalities, uniqueness, and other interesting features. This technique will be discussed in more details can be found in our supplementary materials. In order to better understand the maturity of the approach, the impact of stochasticity, as well as other challenges and pitfalls we construct an experiment for which we define the ground truth(s) explicitly and examine how the technique performs relative to this defined ground truth.

We begin with a discussion of related work in this nascent field which draws from a variety of academic research areas. Next, we outline our approach and the methodology based on the pioneering work of the Evolutionary Model Discovery [7] technique. The agent-based model and our design of experiments are described in Sect. 5. In the remainder of the paper we present experiment results, as well as discussion and topics for future work.

2 Related Work

Inverse Generative Social Science (iGSS) is a nascent academic field, the most recent work in this space can be found at: www.igss-workshop.org. iGSS is an emerging academic discipline aimed at addressing one of the major criticisms of simulation-based analyses, and specifically of agent-based models of social phenomena. More specifically, most ABM-based studies detail a single set of agent behaviors that replicate some social phenomena of interest, which begs the question, "how many other sets of agent behaviors would also replicate the phenomena?" iGSS tries to answer that question by automatically discovering many sets of behaviors that replicate a given referent phenomena. More information on iGSS may be found at www.igss-workshop.org.

There is a potential danger here, however. The blind combination of agent rules may be able to recreate a given phenomenon, but they may be meaningless theoretically. Especially in the case of simulating human systems, the defined potential behavior rules should be grounded in theory, be human readable, and the resulting combinations should be evaluated by how well the recreate the phenomena in question and by how theoretically plausible they are. This is discussed in more detail by Rand [10].

[1]

2.1 Rule Induction

Rule induction leverages machine learning techniques such as Class Association Rules (CAR) to build a connection between agent characteristics and behaviors/decisions. For example, as discussed by Rand [10], one application related agent demographic characteristics to decisions about how to commute. This technique differs from EMD as the universe of behavior options is known a priori. What is not known is how to assign agents to decisions, that is where machine learning algorithms such as CAR come in.

2.2 Computational Abduction

Computational Abduction is an implementation of the abductive loop to inform the development and iterative refinement of agent-based models, especially as applied to human socio-behavioral phenomena. The process as described by Ren et al. [11] involves first analyzing experimental data to create a theory of the underlying generating mechanism. The theory is then tested and refined via simulation. The process is repeated until the researcher is satisfied they have articulated the best explanation.

2.3 Inductive Game Theory

Inductive Game Theory (IGT) is a technique developed by DeDeo, Krakauer and Flack [4] that introduced an analytic technique to infer agent strategies from microlevel interaction data. Specifically, the authors collected data on 1096 fights taking place over a 158 h period within a pigtailed macaque primate population. The data was, essentially, which primates joined each fight over time. Using this micro-level interaction data, the authors sought to infer the decision rules used by the primates regarding whether or not to join an ongoing fight. To begin this process the authors assumed that an individual macaque's decisions about whether or not to join a fight was a function of whom they saw previously fighting. They then analyzed the temporal correlations between macaques fighting at time t and the correlations with those fighting at t–1. These correlations were then used to generate a new distribution of fight sizes and the one with the best fit to the original fight distribution was assumed to be correct.

3 Motivation for the Current Study

Inspired by the original EMD concepts, MITRE is building an open source toolkit (https://github.com/mitre/strategy-mining) in Java that supports multiple agent-based model frameworks (currently supporting NetLogo and MASON). We aim to provide a standardized way to explore the impact of perturbations to the system across plausible rulesets. This will include the development of standard measures and metrics.

Another motivation for this work was to lower the barriers to entry for using EMD. Some users coming from a social science background may be interested in the tool but have no prior Python or Java experience—only a model written in NetLogo. With our new toolkit, all standard evolutionary model settings are set to commonly-used default values and can be edited in a parameters text file. Advanced users who need more complex functionality can write custom Java functions to meet their needs.

Goals for the toolkit are:

- Easy parallelization both on a single machine and across nodes in a high-performance cluster
- No need to create a scheduler, spin up jobs, or work across programming languages
- Hello World examples to guide the user in using the tool
- Default implementations and templates for customizing required EMD functions
- Factors, setup commands, and fitness metrics.

In other words, a user with an existing MASON or NetLogo model should be able to use the tool without needing to know any other programming language than the one their model is written in. The setup should be minimal, and the output should be easy to interpret and use with a researcher's tool of choice.

Finally, we wished to explore EMD as a *process*. In other words, how does one use it to answer specific questions. How does one analyze collections of rules? How is the optimization process that underlies EMD impacted by the stochasticity so often inherent in an agent-based model? A thorough analysis of the EMD process has not yet been described in the literature but is critical if it is to be used to answer significant questions. It is this last piece that we focus on for the remainder of this paper.

4 The Model

To test this tool and process, we created a simple agent-based model. This model was created as part of an initial evaluation of a different iGSS technique, Inductive Game Theory (IGT) [4]. IGT was used to induce the rules macaques used to decide whether or not to join a fight. Inspired by this work, we created a simulation of macaques fighting. The simulation was relatively simple but could still support many different decision rules (see Table 1, below).

The simulation consists of a population of 50 macaques each of whom has a randomly assigned social rank, size, and social network. If no macaques are fighting,

Table 1 Macaque decision rules

Simulation rule	Decision Impact
Minority fighting	Returns true if less than 50% of the macaque population is currently fighting
Majority fighting	Returns true if more than 50% of the macaque population is currently fighting
Majority links fighting	Returns true if more than 50% of the macaques in the particular macaque's social network are fighting
Minority links fighting	Returns true if less than 50% of the macaques in the particular macaque's social network are fighting
Size in threshold	Returns true if the macaques currently fighting are within a user specified threshold of the particular macaque's size
Rank in threshold	Returns true if the macaques currently fighting are within a user specified threshold of the particular macaque's social rank

then two are chosen at random to fight. While the number of macaques fighting is greater than 0, all macaques are given the option to join the fight (if they are not fighting) or leave the fight (if they are fighting). This continues until the fight size stabilizes, at which point all macaques are "deactivated" and the process starts over.

When deciding to join or leave the fight, the macaques use some combination of logical criteria (shown in Table 1). These rules may be combined with either AND or OR (note this is not XOR). If the combination of decision rules returns true, then the macaque will fight, otherwise it will not.

5 Design of Experiments

The nature of genetic programming requires creating and evolving new rule trees. For this reason, and in order to explore a vast search space, our experiments were designed to be easily parallelized not only on a single machine, but also across nodes in a cluster. The high-performance computing system we utilized for the experiments described herein uses Slurm to schedule jobs across a Linux (CentOS) cluster computer.

As discussed above, EMD uses genetic programming to build up new agent behavior rules. This requires a way to measure fitness. This is typically done by comparing the results of the evolved system to a referent. Usually, this referent is a real-world phenomena or another simulation. In order to explore the EMD process thoroughly, before the experiment, we ran the genetic program and stored 75 randomly generated rules which were later used as our design point referents, specifically the distribution of fight sizes they created. For each of those 75 design points, we performed 30 replications of a genetic program experiment which had 30 generations of 30 individual simulations. In total, over two million (2,025,000) individual rule trees were created and evaluated. For this experiment, we minimized the Kruskal–Wallis test statistic as the fitness measure between the "true" underlying

referent data and the generated simulation data. The experiment took approximately 8 h to run.

6 Results

6.1 Comparing Simulation Output

Each of the 75 design points showed different performance in terms of the median fitness and the variation in fitness. Figure 1 shows a boxplot of the fitness distribution for each of the 75 unique design points (each boxplot contains all rules evaluated in all replications). Some design points were so easy to fit that their entire middle 50% fitness range is not visible due to its proximity to zero on the graph. For other design points, the genetic program tried many rules with bad fitnesses.

Similarly, Fig. 2 shows a boxplot with the percent of replications with mean fitness less than 1.32 for each of the design points (smaller is better). The data was aggregated by replication so that each replication had 900 data points: one for each individual simulation in each generation. The threshold fitness of 1.32 was chosen because it is the median fitness of all rules across all design points in generation 30, the final generation. This threshold is used to classify the performance of a rule as "good" or "bad" for all design points. Once again, some design points were consistently able to find rules with good fitnesses. The large variation in the location and size of these boxplots confirms that for some design points it is easier to discover well performing rules.

The genetic program is a heuristic optimization technique. As such, each generation should show an improvement in terms of the fitness distribution. Some design points find new "good" rules in all generations, as shown in Fig. 3. To be labeled as a new rule, it must have not been present in prior generations within that replication. The same threshold for "good" rules is used here (1.32). Once again, some design points find a small set of "good" rules at the end of their 30 generations and others

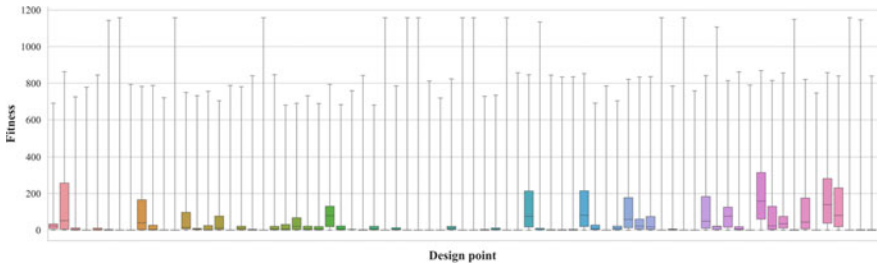


Fig. 1 Fitness distribution for each design point. This distribution includes individuals from all 30 generations in all 30 replications of the experiment. A lower fitness indicates a more similar distribution in ape fight sizes

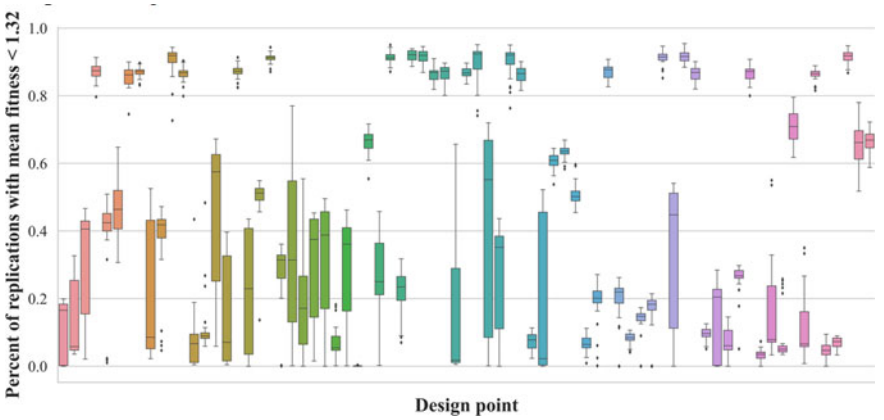


Fig. 2 Percentage of replications with mean fitness less than 1.32 for each design point

find a much larger set. This graph indicates that running the experiment for more than 30 generations may lead to the discovery of even more new "good" rules. Figure 4 shows that the range of fitnesses at each generation is constant, which indicates that each generation is still trying "bad" rules, even as it improves its middle 50% of fitnesses and continues to discover new "good" rules.

The distribution of fitnesses will vary based on the fitness metric used and the experiment setup. For this experiment, the fitness distribution is nearly log-normal, as shown in Fig. 5. The range of fitnesses is approximately [0, 1200], with most of

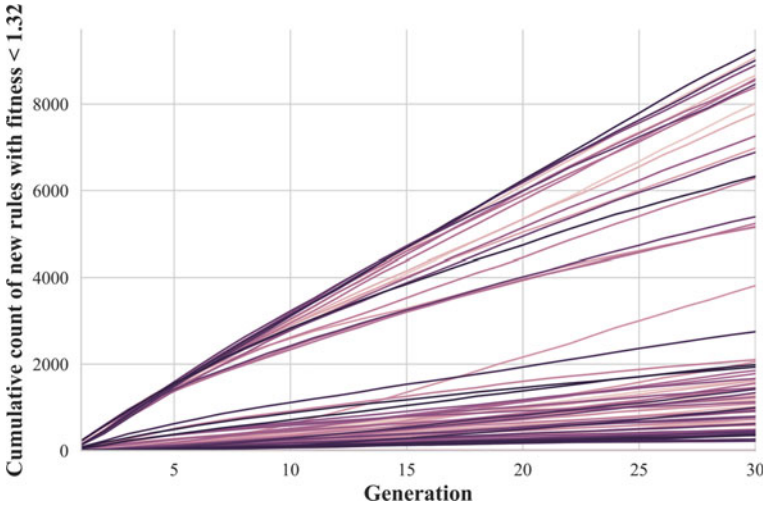


Fig. 3 Cumulative count of unique rules with fitness less than 1.32. Each colored line represents one of the 75 design points

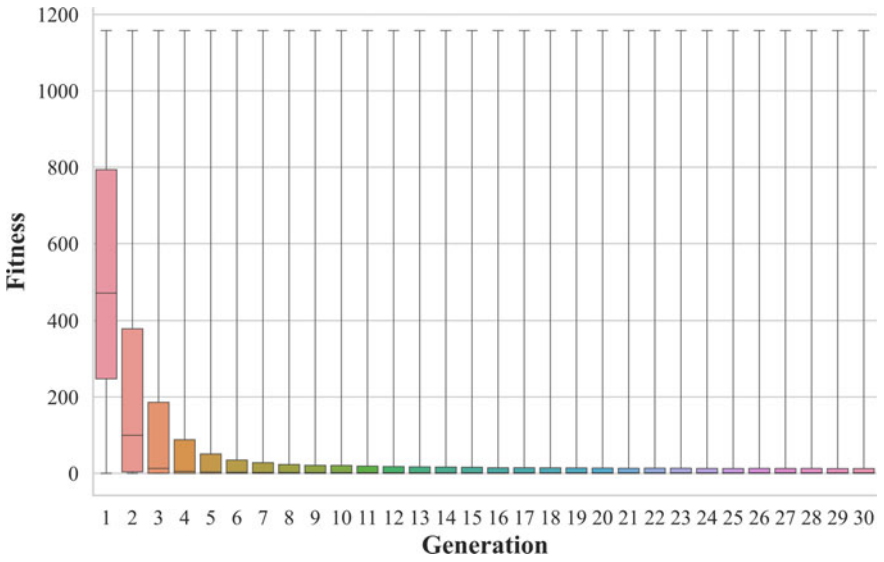


Fig. 4 The range and middle 50% of fitnesses for each generation

the fitnesses falling in the [0, 100] range. The sharp spikes in this plot are likely attributed to a small set of non-stochastic (or less stochastic) rules.

During the analysis of our experiment data, and after realizing the degree that stochasticity affects the fitness calculation, we ran an additional experiment to generate data with each design point being measured against itself. For each rule, we ran the model once to generate data, then ran the model nearly 200 times and calculated the fitness of the data against the original "ground truth" data generated

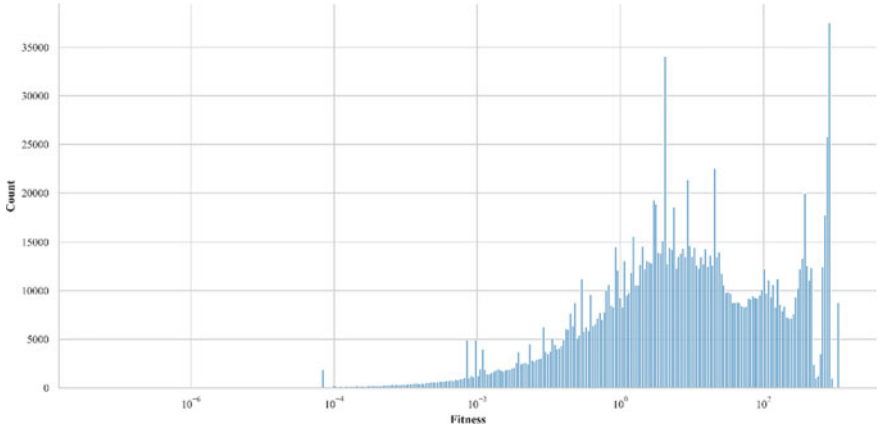


Fig. 5 The distribution of all fitnesses, shown with a logarithmic x-scale

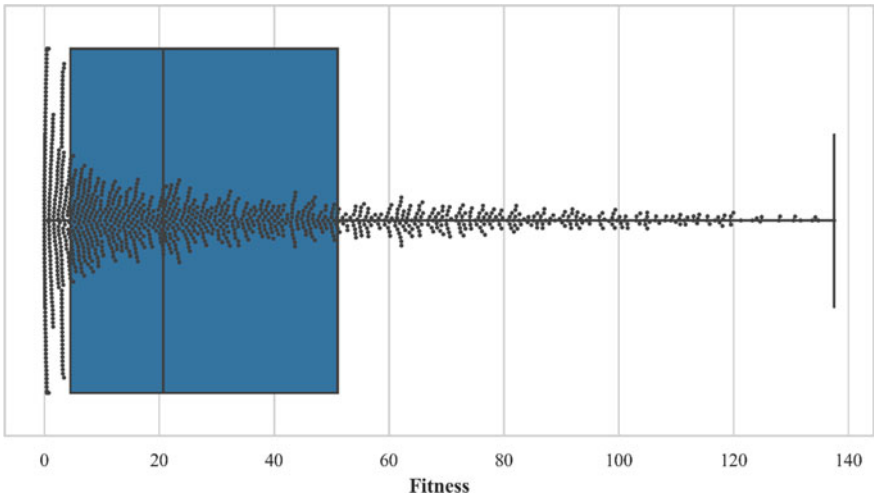


Fig. 6 The distribution of fitnesses when a rule was run 1225 times against itself

by the same rule. We expected that the fitnesses should be very close to or exactly zero. This data was generated for each design point. Figure 8 shows the distribution of fitnesses from the results of this experiment on just one design point. For this design point, only 12% of the fitnesses fell below the 1.32 threshold used in Figs. 2 and 3, even though all of them were generated from the same rule.

We combined this data with the original dataset and separated it to differentiate rules that were actually identical or different. Figure 6 shows the risk that even the "true" rule can have a bad fitness. For example, a fitness of 5 corresponds to an approximately 70% chance that the rule is identical. A fitness of 250 corresponds to a 50% likelihood that the rule is identical, and a fitness near 500 has a near-zero likelihood of coming from an identical rule.

Figure 6 also demonstrates that the Kruskal–Wallis test statistic used as the fitness should not be interpreted as a "true" test statistic. A test statistic greater than 5 is statistically significant at $\alpha = 0.05$ with one degree of freedom, but in our experiment, the likelihood of the distributions coming from identical rules crosses 0.05 at a fitness greater than 400. The metric chosen does not have to be interpreted in a statistical sense; it merely needs to be able to quantify a performance comparison, i.e., to say that one rule performed better than another, in this case measure the difference between two distributions (Fig. 7).

6.2 Comparing Generated Rules

Comparing simulation output to a referent is relatively straight forward and well defined. Trying to compare the sets of generated rules, on the other hand, is not at

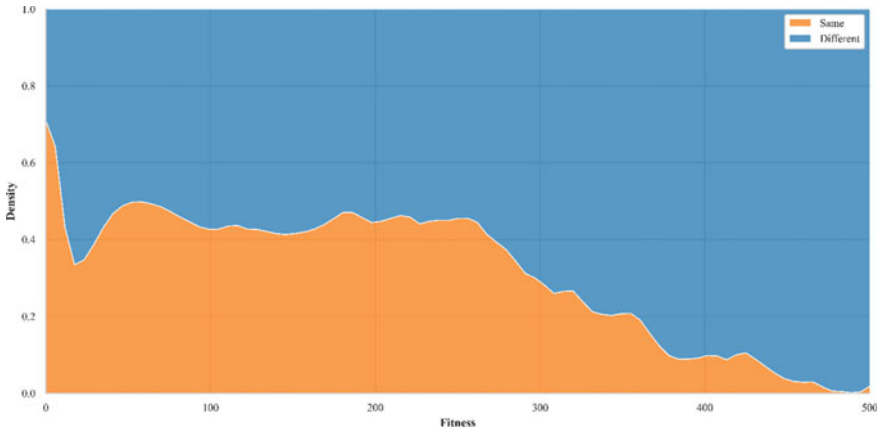


Fig. 7 Normalized stacked area plot showing the portion of the data at each fitness value coming from identical (orange) or different (blue) rules

all clear. This task is important to understand as the output of iGSS are the rules, essentially the agents, rather than the simulation output data. In the present case this task is somewhat simplified because all the rules are combinations of booleans; therefore, all rules may be expressed as a set of 0's and 1's combined by AND and OR. As initially defined, EMD uses a random forest regression to calculate the relative contribution of each atomic behavior rule to overall performance. While this is an effective means, it potentially misses combinations of behaviors that have synergistic effects on fitness. We are particularly interested in producing an analysis akin to that defined by Wagner [9] where metabolic genotype and phenotype are mapped into a fitness landscape and mutation-based movement across the landscape characterizes the "robustness" of the organism to mutation. Here, of course, we are not dealing with gene expression or the like. Rather, we want to explore changes to sets of rules and their impact on simulation fitness relative to a referent. This requires a well-defined measure of distance among rules.

As discussed above, in order to compare the simplest form of a distance between rules, we examined a distance matrix based on fitness. While we chose to use the Kruskal statistic, it should be noted that we used it simply as a method for comparing distributions and not strictly as a statistic. In order to compare not only simulation performance but also agent behavior rules, we attempted to find a way to measure differences in rules that corresponded with differences in performance. As the rules in this simulation were exclusively boolean, we examined Hamming distance and tree edit distance. There was little correlation between fitness distance (as measured by the KS statistic) and rule space distance (measured by either tree edit distance or Hamming distance). This is likely due to the lack of correlation between the expression and behavior of a rule within the simulation.

We found a similar result with clustering. We clustered the rules and then examined the relationship between cluster membership and performance. Once again, we found

little relationship between cluster membership and performance. However, the test case discussed here might not be generalizable. Here all the rules are boolean, which is likely not going to be the case with many other ABMs. However, we chose to do this as we felt it would make the analysis of the rules themselves somewhat easier. Even with this stylized, simplified test case we were unable to find an appropriate way to analyze the rule our EMD system created beyond a relatively basic frequency of subcomponents analysis and the factor regression described by Gunaratne and Garibay [8].

7 Conclusions

EMD is a very promising, efficient methodology for discovering combinations of agent rules that create a particular output. We were able to reproduce the EMD functionality as originally specified by Gunaratne and Garibay [8] in a fully Java framework; thus, creating an easier to use and maintain implementation. We then undertook a study to explore the overall EMD process. Our major findings include:

- All components of agent behavior must be expressible and expressed a priori
- The EMD system can find combinations of agent behaviors that produce simulation dynamics "close" to a referent
- The typically inherent stochasticity of agent models may create a problem with the signal to noise ratio complicating the GP heuristic search process
- Rules are not consistently tested (some are run once other are run thousands of times)
- The EMD framework only test agent rules, it does not explore rules and parameter values, this could be an issue if rule performance is impacted by the parameter space.

EMD has proven to be a very effective method for discovering families of agent behaviors that create a simulation that generate dynamics that are "close" to a referent. It is, however, not without its nuances that should be fully understood by researchers using it. As shown in Fig. 6, even when comparing a simulation to itself a wide variety of results may be obtained. This complicates the analysis of results and highlights the need to run multiple replicates of a simulation. In particular, if EMD is evaluating fitness based on similarity to non-EMD-generated output data, e.g., the "real world," then it may be trying to fit to an outlier, again, complicating the use of the EMD system.

We intend to continue to develop our EMD implementation and explore how best to utilize and understand the EMD process. Future work should include recommendations of the best ways to answer experimental questions, best ways to measure rule complexity, best ways to measure rule similarity (distance metrics), and how to finetune experimental parameters.

For citations of references, we prefer the use of square brackets and consecutive numbers. Citations using labels or the author/year convention are also acceptable.

The following bibliography provides a sample reference list with entries for journal articles [1], an LNCS chapter [2], a book [3], proceedings without editors [4], as well as a URL [5].

References

1. Axelrod, R.: Simulation in Social Sciences (2007)
2. Axtell, R.L.: Why Agents? On the Varied Motivations for Agent Computing in the Social Sciences (2000)
3. Buss, S., Papadimitriou, C., Tsitsiklis, N.: On the predictability of coupled automata: an allegory about chaos. In: Proceedings [1990] 31st Annual Symposium on Foundations of Computer Science, pp. 788–793 (1990)
4. DeDeo, S., Krakauer, D., Flack, J.: Inductive game theory and the dynamics of animal conflict. PLOS Comput. Biol. **6**(5) (2010)
5. Epstein, J.M.: Generative Social Science (2006)
6. Epstein, J., Axtell, R.: Growing Artificial Societies: Social Science from the Bottom Up (1996)
7. Gunaratne, C.: Evolutionary Model Discovery: Automating Causal Inference for Generative Models of Human Social Behavior (2019)
8. Gunaratne, C., Garibay, I.: Evolutionary model discovery of causal factors behind the socio-agricultural behavior of the Ancestral Pueblo. PLOS ONE **15**(12) (2020)
9. Hamilton, C., Strader, L., Pratt, J., Maiese, D., Hendershot, T., Kwok, R., et al.: The PhenX Toolkit: Get the Most From Your Measures. Am. J. Epidemiol. **174**(3), 253–260 (2011)
10. Rand, W.: Theory-Interpretable, Data-Driven Agent-Based Modeling, pp. 337–357
11. Ren, Y., Cedeno-Mieles, V., Hu, Z., Deng, X., Adiga, A., Barrett, C., et al.: Generative modeling of human behavior and social interactions using abductive analysis. In: Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 413–420 (2018)

Flocking with Only Two Parameters

Dashiell Bhattacharyya and William G. Kennedy

Abstract The boids flocking model, an agent based model simulating bird flocking behavior, relies on four parameters: the rates of alignment, separation, cohesion, and the vision radius of the birds. Only certain combinations of these parameters result in the birds flocking together. Reliance upon fewer parameters would result in a more efficient and faster modeling process along with clarifying the concepts behind the model's behavior. It was hypothesized that separation and cohesion could be combined into a single parameter and vision radius removed without affecting the resulting behavior. To test this hypothesis, these simplified models were created, run, and the result of every combination of parameters distilled to a set of numeric metrics. It was shown that a simplified model combining cohesion and separation to their ratio, termed friendliness, as well as fixing vision radius, was able to result in a functionally equivalent model, demonstrating that the two parameter model is a successful substitute for the original four parameter one.

1 Introduction

The boids flocking model is a classic agent-based model that simulates biological flocking behavior in a 2D grid. The birds, fish, or other agents act in relation to their neighbors, defined as the set of agents whose distance to a specific agent is below a certain threshold, the vision radius. The vision radius is also known as neighborhood distance and neighborhood radius. We will be using birds as representative agents. The birds traditionally move according to three rules.

- **Rule 1: Separation**: Birds turn away from their neighbors if they are too close. The rate at which these birds turn is dictated by the separation parameter, also referred to as avoidance, and scaled by the proximity of nearby birds.

D. Bhattacharyya · W. G. Kennedy (✉)
George Mason University, 4400 University Drive, Fairfax, VA 22030, USA
e-mail: wkennedy@gmu.edu

Z. Yang and S. Núñez-Corrales (eds.), *Proceedings of the 2022 Conference of The Computational Social Science Society of the Americas*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-031-37553-8_10

- **Rule 2: Cohesion**: Birds turn towards their neighbors. The rate at which these birds turn is dictated by the cohesion parameter. The combination of cohesion and separation results in the birds attempting to maintain a constant distance from each other.

- **Rule 3: Alignment**: Birds turn towards the same direction their neighbors are moving. The rate at which these birds turn is dictated by the alignment parameter, also referred to as consistency. The resulting effect of this rule is that birds in the same flock move in the same direction.

Neighborhood radius, separation, cohesion, and alignment are the four parameters needed for the classic boids model to function. Variations in the parameters modify the flocking behavior of the birds, and certain combinations of parameters remove the ability of the birds to form flocks altogether.

2 Background

Ever since Craig Reynolds [1] created the boids model it has been widely used to model not only bird flocking behavior but also employed in the field of robotics and for swarm intelligence [2]. Its uses have continued into artificial life simulations [3], particle physics [4], and data visualization [5]. Christopher Hartman and Bedrich Benes [6] provided a mathematical explanation for each of the three rules of the model. Separation was defined as the negative sum of the vectors to all nearby birds. Cohesion was defined as the vector towards the center of mass of all nearby birds. The center of mass was calculated by averaging the birds' positions. Finally, alignment was calculated as the average direction vector of all nearby birds. These three vectors were then weighted, summed, and normalized to the magnitude of a set constant speed. This implementation allows for all the functionality of Craig Reynolds's original paper. In Hartman and Benes's paper, the model was also extended to contain an additional rule; however, we will be working only with the original 3-rule model. It is clear from the model's widespread array of uses that attempts to simplify or optimize it would be beneficial. The model in past has been optimized for GPUs [7], and parameters of boids have also been optimized using genetic algorithms [8]. Despite extensive research and use, no attempt to simplify the model has been made to our knowledge.

3 Experiment 1

The goal of the first experiment was to determine an optimal number of steps for the rest of the experiments. All birds start in random positions, and a certain amount of time must have elapsed before their positions are determined by the parameters.

3.1 Methodology

For the purpose of the first experiment, 300 birds were used in a 500 unit by 500 unit grid, each with neighborhood radius set constant at 10 units. Separation, alignment, and cohesion parameters were all set to 1, as these are their default values and result in flocking behavior.[1] The environment was a 2D grid with wrapping edges, and all distance calculations accounted for this.

A way to define a model in terms of quantified metrics was also needed. Two metrics were created: the average distance between each bird and its neighbors, and the average number of neighbors per bird with a radius of 10. It was hypothesized that a graph of these metrics over time as a model was run would plateau once enough steps had elapsed for birds to be in a position dependent on their parameters. If so, the number of steps for the metrics to remain constant could be used in the following experiments to get the reliable data.

To find the reasonable minimum number of steps, the model with default parameters was run for 2000 steps, and the two metrics calculated at each step. The data collected from this was then output to a data file.

3.2 Results

Both the distance between neighbors and the number of neighbors plateau by step 1000, so that is a practical number of steps required to determine a complete assessment of the effect of the parameters on the state (see Figs. 1 and 2).

All further experiments therefore were calculated for 1000 steps.

3.3 Analysis

The graphs Figs. 1 and 2 plateau by approximately 1000 steps, and therefore the model does not need to be run for longer for all behaviors to have emerged.

These conclusions can be observed visually. Between 0 and 1000 steps, the flocking of the agents changes over time, but from 1000 to 1500 no more changes can be observed. From 1000 on, the approximate number of flocks and visual organization of the birds remains constant (see Fig. 3). These conclusions were also observed to

[1] Flocking behavior is defined by the following: First, most birds must have multiple neighbors with which they generally stay. Second, the birds must be moving in the same direction as the others in their flock. Finally, birds in a flock must have space between them, such that they are not consistently colliding. If a model was observed not to flock by this definition, it does not achieve the fundamental purpose of the algorithm, organic flocking, and therefore is not relevant in determining if two versions of the flockers model are functionally equivalent. For all experiments, it was tested in a modifed version of MASON's Flockers with UI and visually analyzed under the aforementioned three conditions.

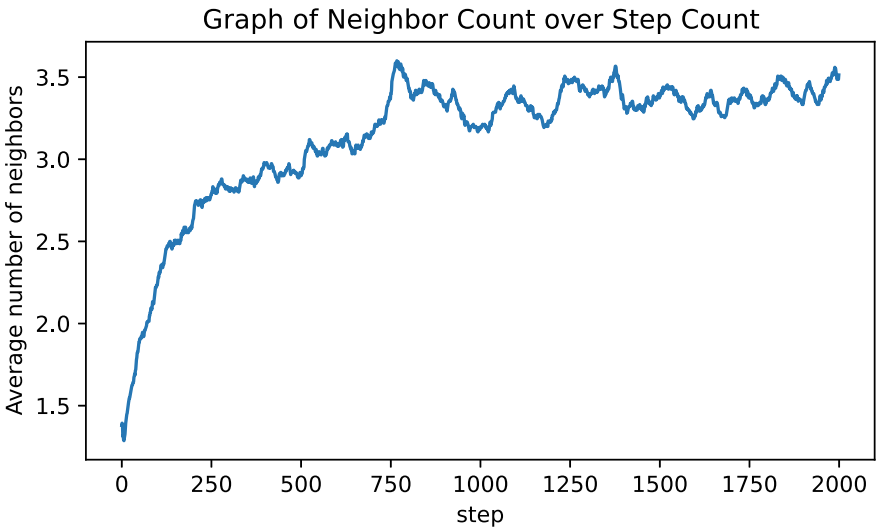


Fig. 1 Graph of the average number of neighbors over step count, run on a model of all weights at 1 and neighborhood radius at 10

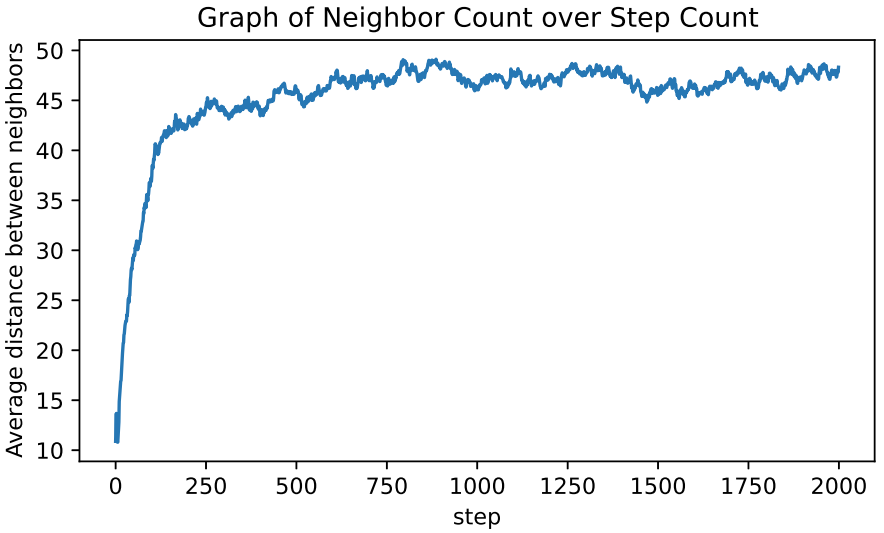


Fig. 2 Graph of the average distance between neighbors over step count, run on a model of all weights at 1 and neighborhood radius at 10

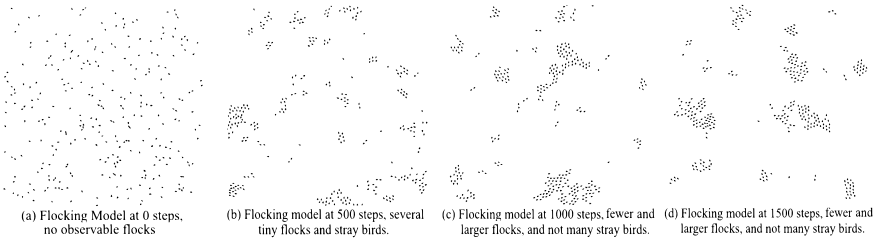


Fig. 3 Boids model visualization at 0, 500, 1000, and 1500 steps, done using MASON's Flockers with UI

apply to other combinations of parameters. It is unknown if 1000 is the optimal step count for other grid sizes or numbers of birds, but as all all further experiments are run in a 500 unit by 500 unit grid containing 300 birds, this is not a concern.

For the purpose of the future experiments, 300 birds were used in a 500 unit by 500 unit grid. Initial bird positions and rotations were random; however, across multiple trials, initial positions scarcely affected resultant metrics. Birds who flew out of the bounds of the grid wrapped around to the other side, and all distance calculations accounted for this. Experiments were run for 1000 steps as the standard for comparisons.

4 Experiment 2

The goal of the second experiment was to determine whether or not combining cohesion and separation into the single new parameter named friendliness prevented certain flocking behaviors from being possible.

We define the friendliness parameter as a combination of separation and cohesion. The individual bird's movement calculation is the same, but the rate at which the bird coheres is scaled by the new friendliness parameter, and the rate at which the bird separates is scaled by the inverse of the friendliness parameter. Therefore, the friendliness parameter acts as the ratio between coherence and separation.

4.1 Methodology

Two versions of the boids model were created. One, the original boids model, where alignment separation and cohesion were scaled by their respective values, and another, where alignment was scaled by its weight, cohesion was scaled by the friendliness parameter, and separation was inversely scaled by friendliness. Neighborhood radius was held constant throughout the whole experiment at 10.

The original boids model was run on every value of cohesion from [0, 6) with an increment of 0.5, along with the additional value of 0.25 (i.e. 0.25, 0.5, 1, 1.5 ... 5.5). With each of these values for cohesion, the model was run on every value of separation with the same range of values. Finally, it was run on every alignment value [0, 5) with an increment of 1 (i.e. 1, 2, 3, 4). These ranges were selected based on visual analysis of the range of parameters which resulted in flocking behavior. With these 3 parameters, the model was run for the number of steps determined in Experiment 1, and then evaluated on the two metrics. Then the cohesion weight, separation weight, alignment weight, average number of neighbors, and average distance between neighbors was output to a data file for analysis.

The modified 2 parameter model was run from every friendliness value from [0.25, 8) with an increment of 0.25 (ie 0.25, 0.5, 0.75 ... 7.75), and for each friendliness parameter, the model was run on every alignment value [0, 5) with an increment of 1. These numbers were likewise selected based on the range of values which resulted in flocking behavior. The model was run for the same number of steps, then evaluated on the same metrics, and the friendliness, consistency, average number of neighbors, and average distance between neighbors was output to a data file.

4.2 Results

After 1000 steps, the second experiment had models of various parameters forming three regions (see Fig. 4). The first region included both 4-parameter and 3-parameter models, very few neighbors per bird and a relatively small average distance per neighbor. Models in this region had relatively high separation and low cohesion (see Figs. 5 and 6), or low friendliness (see Fig. 9), and birds did not flock together due to insufficient cohesion and too much separation.

The second region on Fig. 4 had both types of models, relatively high neighbors per bird, and a wide range of number of neighbors. Upon testing several models in this region, it was determined that most models in the second region had flocking behavior.[2] The models in this region had a relatively wide range of parameters; however, cohesion was always greater than 1, likely because sufficient cohesion is necessary in order to have flocking by our definition (see Fig. 6). While birds with alignment values of zero were present in second region, they did not always flock (see Figs. 7 and 8), due to the birds in a flock not moving in the same direction.

Finally, the third region on Fig. 4 only had the original models with a high average number of neighbors, but extremely low average distance per neighbor. The birds

[2] Models in this region were tested under the assumption that models with similar metrics behaved similarly. Because of this reasoning, the extremes of each region were prioritized, and the centers tested in an even distribution. The assumption that similar parameters result in similar results was also made.

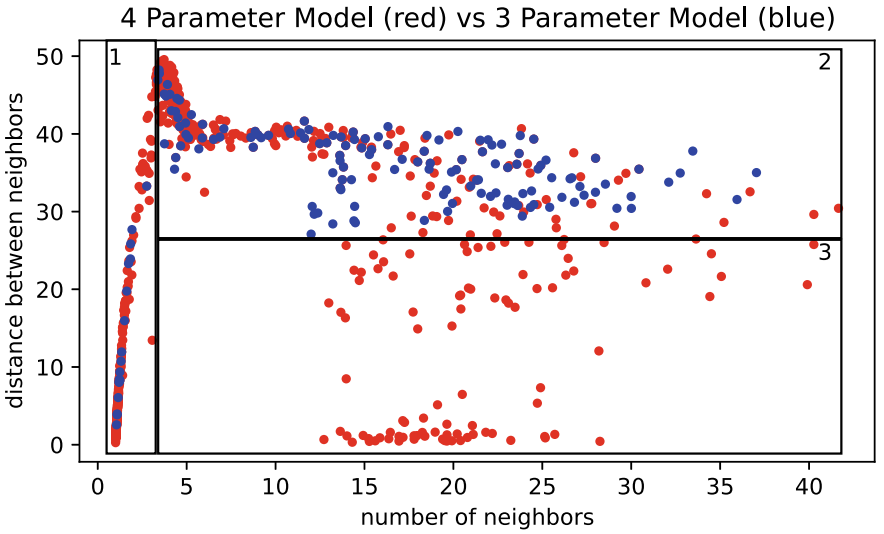


Fig. 4 Both the 4 parameter and 3 parameter models graphed in terms of both metrics, the mean distance between neighbors and the average number of neighbors. The graph was split into three regions, 1, 2, and 3, each expressing similar behavior

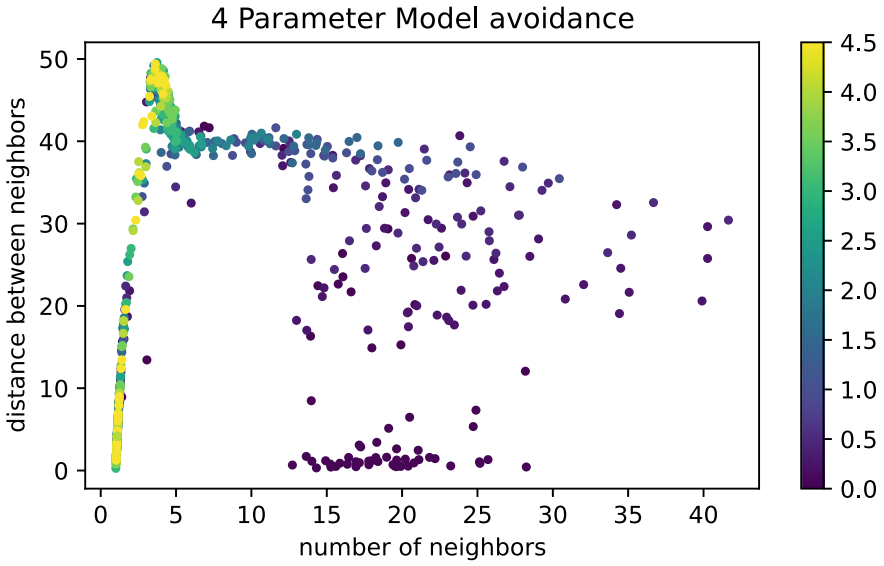


Fig. 5 The scatterplot of the original 4 parameter model colored by avoidance parameter over the range 0–5 units

here clustered together very compactly, and flocking was not achieved. The models in this region had little separation weight (see Fig. 5), and the birds clustered too tightly.

4.3 Analysis

In addition to these results, it can be observed that the models with extremely high separation often had a very low average number of neighbors per bird, although the average distance between those neighbors varied depending on cohesion (see Fig. 4). Region 1 tends to have low cohesion, which with high separation, results in the behavior of independent non-flocking birds. Despite the upper left corner of region 2 having high separation, the high cohesion seems to compensate for this effect. Region 3 had relatively high cohesion compared to separation (see Figs. 5 and 6). The effect of alignment on the movement of the birds is less concrete; however, every flocking combination with extremely low consistency had high cohesion to preserve flocking behavior. If alignment was too low, flocking behavior was never achieved (see Fig. 7).

For the 3 parameter model, without sufficient alignment, the birds were not able to achieve a higher number of neighbors due to the flocks often splitting, and most of the models without sufficient alignment did not flock (see Fig. 8). The effect of the ratio between cohesion and separation on this graph is much clearer since models with insufficient friendliness were not able to flock without high enough alignment to keep birds together. Higher friendliness allowed for closer neighbors and more neighbors per bird (see Fig. 9). Overall, the new friendliness parameter has a much clearer effect on the movement of the boids than either separation or cohesion does.

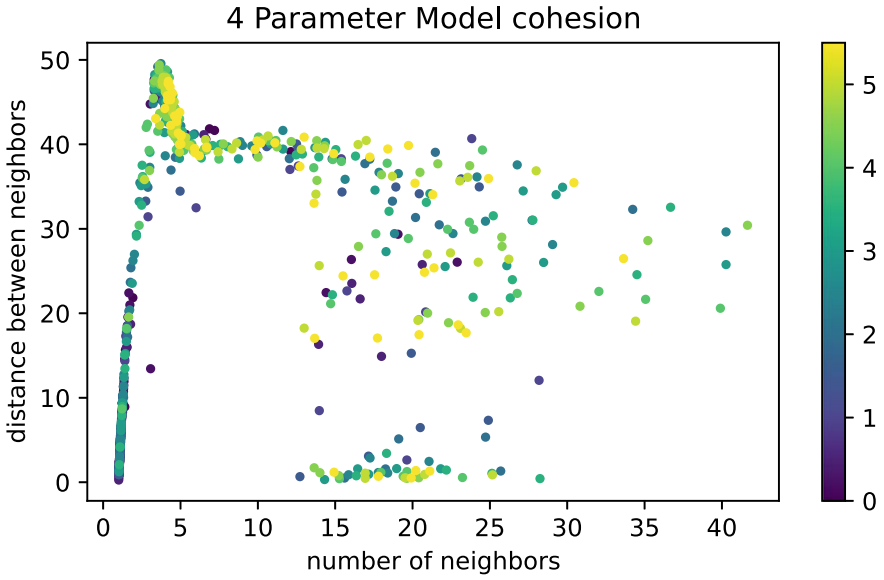


Fig. 6 The scatterplot of the original 4 parameter model colored by cohesion parameter over the range 0–6 units

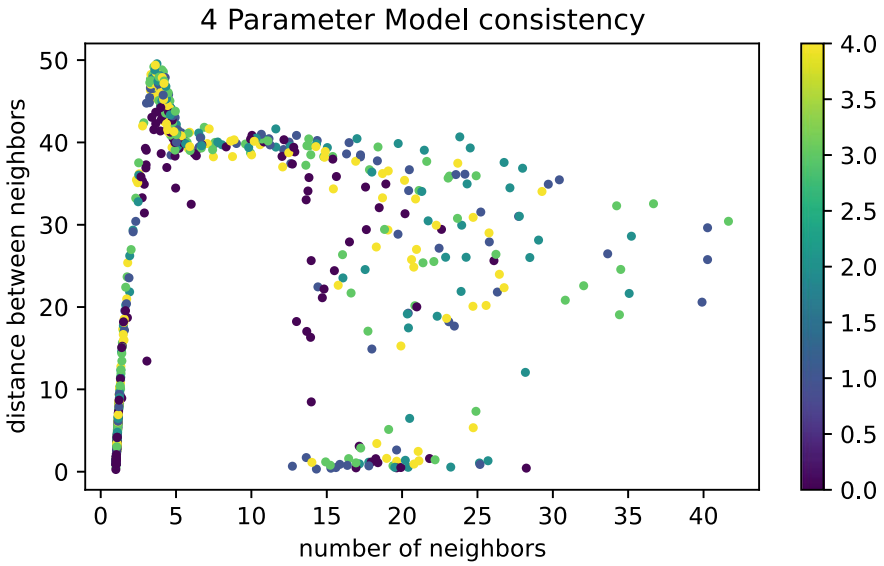


Fig. 7 The scatterplot of the original 4 parameter model colored by consistency parameter over the range 0–5 units

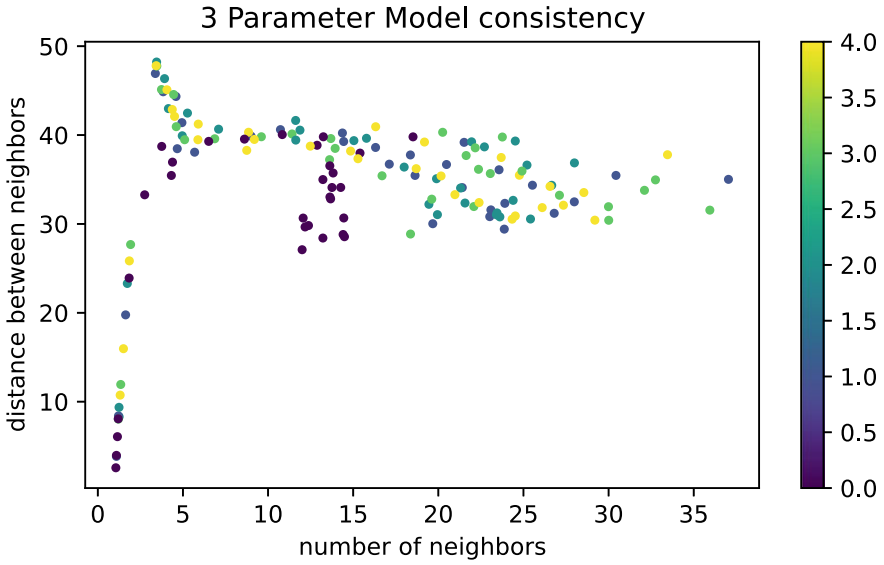


Fig. 8 The scatterplot of the 3 parameter model colored by consistency parameter over the range 0–5 units

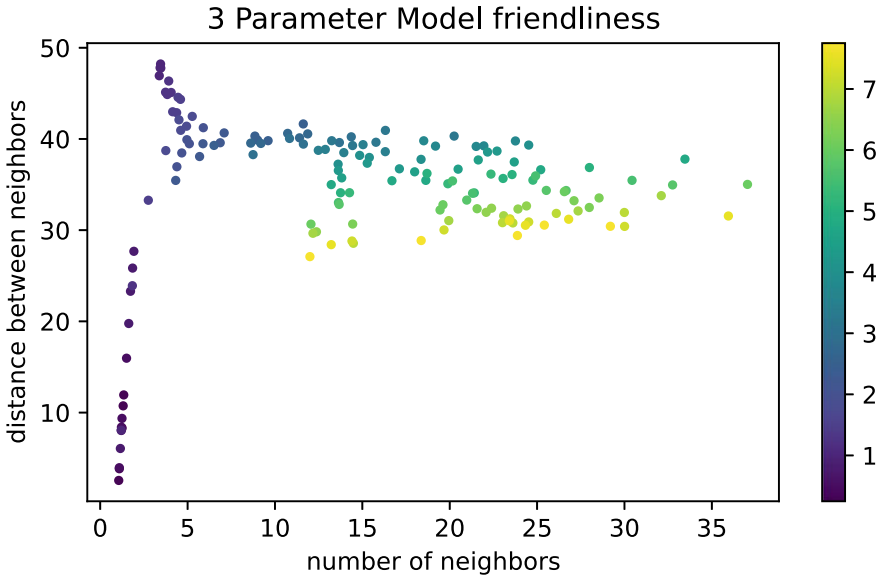


Fig. 9 The scatterplot of the 3 parameter model colored by friendliness parameter over the range 0–8 units

Because only region 2 had flocking behavior and included models of both types, both models are equally capable of producing varying types of flocking behavior. Therefore, Experiment 2 demonstrates that combining the values of cohesion and separation into one parameter, friendliness, successfully deparameterizes the model.

5 Experiment 3

The goal of the third experiment was to determine if changes in the neighborhood radius of the birds could be substituted by changes in the other parameters. If so, then the neighborhood radius could be held constant, allowing for the removal of an additional parameter.

5.1 Methodology

For this experiment, a modified version of the friendliness-based model that was used in Experiment 2 was used. However, this version no longer kept neighborhood radius constant. This way, a version with and without keeping the neighborhood

radius constant could be compared, to see if the parameter neighborhood radius was necessary.

For the purpose of the third experiment, we used the same environment: 300 birds in a 500 unit by 500 unit grid. Birds who flew out of the bounds of the grid wrapped around to the other side, and all distance calculations accounted for the screen wrapping.

This model was run on every combination of friendliness [0.5, 8) with an increment of 0.5, alignment [0, 5) with an increment of 1, and neighborhood radius [8, 13) with an increment of 1, for the same 1000 steps as before. As in Experiment 2, these ranges selected based on the range of values resulting in flocking. Then, friendliness, alignment, neighborhood radius, average number of neighbors, and average distance between neighbors were output to a data file.

These data were compared to the data on the friendliness model from Experiment 2 without a changing neighborhood radius, to see if changes in neighborhood radius create more flocking combinations of both metrics than the base model.

5.2 Results

In Experiment 3, the version of the 2 parameter model with a changing neighbordhood radius almost exactly followed the model with a constant neighborhood radius (see Fig. 10).

5.3 Analysis

Increases in neighborhood distance could be compensated by decreases in friendliness and alignment, since models with varying neighborhood distances could have the same average numbers of neighbors and average distances by adjusting the other parameters to match (see Figs. 11, 12 and 13).

In addition to these results, it was observed that the only area where the 3 parameter model is solely present in Fig. 10 is with a radius value of 8, too small for flocking to occur unless an extraordinarily high friendliness is present. As these models here do not result in flocking by our definition, they can be ignored, and the parameter can still be removed (see Fig. 11). Models with higher neighborhood radius will remain in the same region with a lower friendliness. A higher neighborhood radius increased the strength of cohesion since it acts on more birds, so a lower friendliness was required to exhibit the same results (see Fig. 12). The trend observed in Fig. 11 repeats itself in Fig. 13: because a higher neighborhood radius causes alignment to be more powerful, acting on more birds, a high neighborhood radius is almost always accompanied by a lower alignment in order to preserve the same behavior (see Fig. 13).

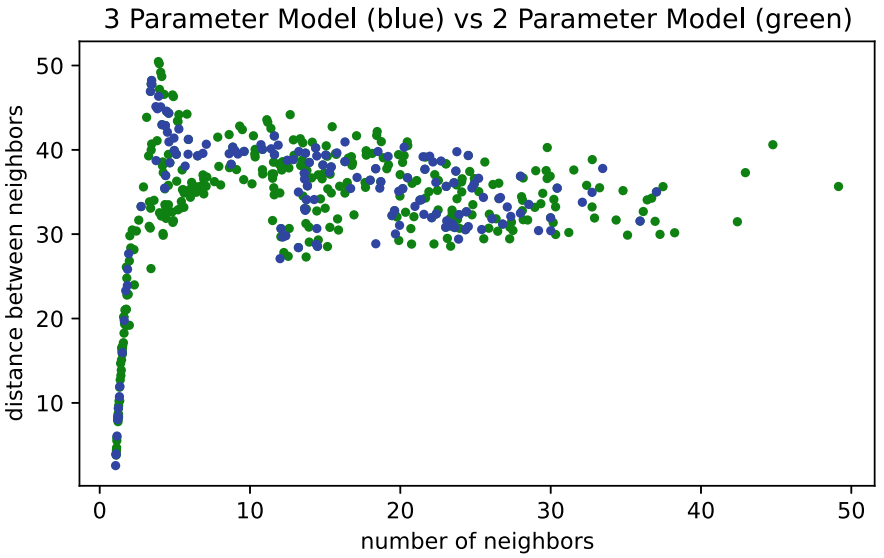


Fig. 10 A scatterplot of the average distances between neighbors and numbers of neighbors across the 3 parameter model done in Experiment 1 with changing neighborhood radius, and the 2 parameter model with constant neighborhood radius, colored green and blue respectively

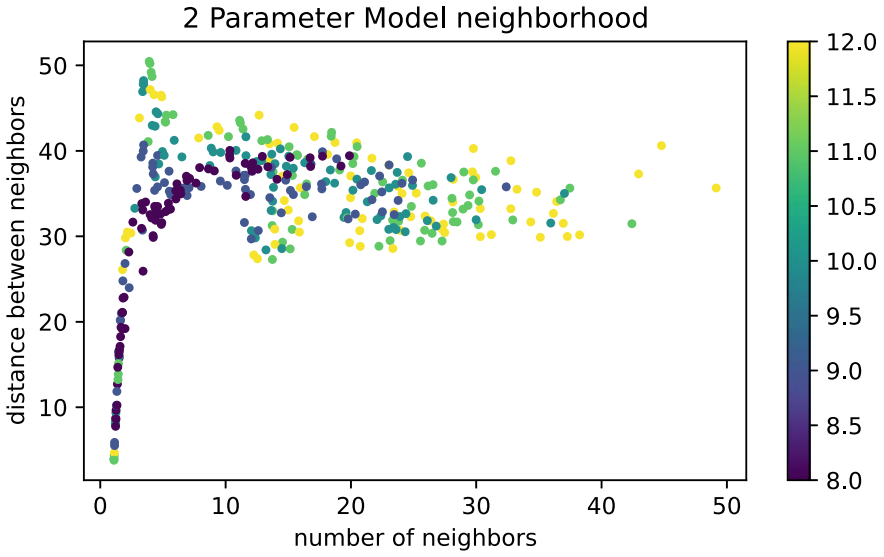


Fig. 11 A scatterplot of the average distances between neighbors and the number of neighbors of the 3 parameter model with varying neighborhood radius, colored by neighborhood radius on the range 8–12

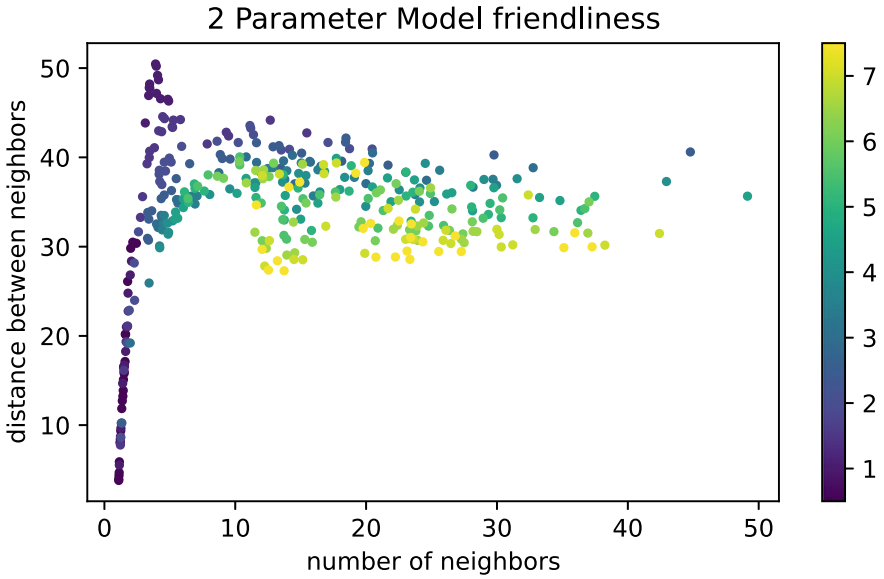


Fig. 12 A scatterplot of the average distances between neighbors and the number of neighbors of the 3 parameter model, colored by friendliness on the range 0–8

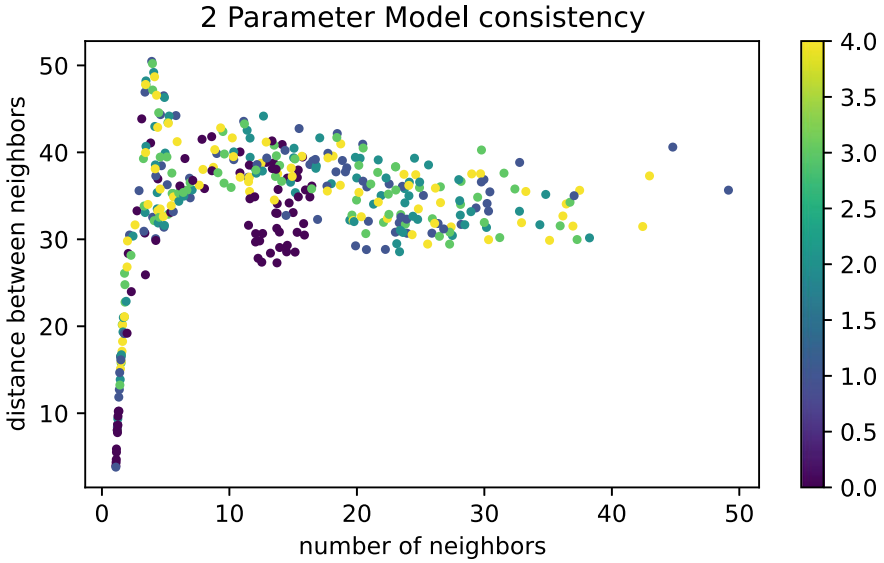


Fig. 13 A scatterplot of the average distances between neighbors and the number of neighbors of the 3 parameter model, colored by alignment on the range 0–5

Because changes in the neighborhood radius parameter can be substituted by changes in the other parameters, and because all flocking combinations of the number of neighbors and distances between neighbors can be achieved by both a model with constant radius and one with varying radius, the parameter neighborhood radius can be held fixed to 10.

6 Conclusions

Because the 2 parameter model and 3 parameter model with changing neighborhood radius were shown to be functionally equivalent for flocking combinations, and the 3 parameter model and 4 parameter model with varying neighborhood radius were determined to be functionally equivalent for flocking combinations, it can be derived that the boids flocking model, with the 4 parameters separation, consistency, alignment, and neighborhood radius, can be reduced to a 2 parameter model model with the parameters friendliness, alignment, and a constant vision radius.

For the implementation of the simplified boids model, the cohesion and separation parameters can be replaced by **friendliness**. Where cohesion is used, friendliness should be placed, and where separation is used, 1 should be. Neighborhood radius can be removed as a parameter and set to a constant 10.

To convert between the original 4 parameters to a 2 parameter version with equivalent behavior, substitute increases in neighborhood radius with increases in friendliness and alignment, and use the following formula.

$$\text{friendliness} = \frac{\text{cohesion}}{\text{separation}}$$

With these tools it is possible to recreate almost all flocking combinations of the boids model in a simplified 2 parameter version of the model. The model is therefore comparable to the original with half the parameters, and any instance of modeling behavior with the boids model can be done more efficiently with fewer parameters to tune.

Further work would involve a more comprehensive sweep of the 4, 3, and 2 parameter models, to verify the equivalence of the simplified model at the extremes of parameters, and to fine tune the parameter conversion calculation between models. Additionally, it would be informative to test the simplified model on existing applications of the boids model to see if the functional equivalence continues into more practical applications. It would also be useful to test multiple implementations of the boids model too see if results hold.

Further mathematical analysis of the model could also be done to determine significance to flocking behavior itself, derive a numerical relationship between the models, and determine the necessity and effect of the remaining parameters.

Acknowledgements The first author, a student, is grateful to the Aspiring Scientists Summer Internship Program, a program at George Mason University which gives research opportunities to students, for creating the opportunity for this research. The authors are also grateful to The Center for Social Complexity at George Mason University.

References

1. Reynolds, C.W.: Comput. Graph. **21**(4), 10 (1987)
2. Multiple Aerial Vehicle Formation Using Swarm Intelligence.: In: Guidance, Navigation, and Control and Co-Located Conferences (2012). https://doi.org/10.2514/6.2003-5729
3. Choi, T.J., Ahn, C.W.: Artificial life based on boids model and evolutionary chaotic neural networks for creating artworks. Swarm Evol. Comput. **47**, 80–88 (2019). https://doi.org/10.1016/j.swevo.2017.09.003
4. Quillen, A.C., Smucker, J.P., Peshkov, A.: Boids in a loop: Self-propelled particles within a flexible boundary. Phys. Rev. E **101**(5), 052618 (2020). https://doi.org/10.1103/PhysRevE.101.052618
5. Moere, A.V.: Time-varying data visualization using information flocking boids. In: IEEE Symposium on Information Visualization, pp. 97–104 (2014). https://doi.org/10.1109/INFVIS.2004.65
6. Hartman, C., Benes, B.: Autonomous boids. Comput. Animat. Virtual Worlds **17**(3–4), 199–206 (2006). https://doi.org/10.1002/cav.123
7. Silva, A.R.D., Lages, W.S., Chaimowicz, L.: Boids that see: using self-occlusion for simulating large groups on GPUs. **7**(4) (2010). https://doi.org/10.1145/1658866.1658870
8. Kobayashi, K., Huang, X., Nakao, Z.: Genetic algorithms for optimization of boids model. In: Gabrys, B., Howlett, R.J., Jain, L.C. (eds.) Knowledge-Based Intelligent Information and Engineering Systems, pp. 55–62. Springer, Berlin (2006). https://doi.org/10.1007/11893004_7
9. Wilensky, U.: NetLogo Flocking model. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL (1998). http://ccl.northwestern.edu/netlogo/models/Flocking
10. Wilensky, U.: NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL (1999). http://ccl.northwestern.edu/netlogo/

Understanding Genocide Through Emotion Detection in Historic Documents

Elizabeth M. von Briesen, Michael Garvin, and Samira Shaikh

Abstract Artificial simulations such as agent-based models provide an environment in which to explore the dynamics of human behavior, and can contribute to the work of genocide researchers seeking to understand and prevent this atrocity. Validation is essential to the development of such models, as it verifies that the simulation can reproduce expected results. This work explores methods for quantifying emotion in historic presidential and elite speech data for later use in model validation. Here, data were collected for cases in which genocide was expected and either did or did not occur. After digitization of a subset of the speech archive, early emotion analysis using a modern lexicon on translated text shows variations in emotions that correspond as expected to events along the case timeline, thus providing preliminary measures for historic case comparison in general social science research and appropriate artificial simulations. Future improvements in analytical technique and digitization processes will allow for continued work and contributions to the domain of genocide studies. A repository of digitized files, along with the full set of original texts, are also made available.

Keywords Natural language processing · Genocide · Computational social science · Agent-based modeling

1 Introduction

Working to prevent future genocides, historians, social scientists, and others have probed into the darkest of humanity's corners to determine its causes. Agent-based models can assist in this domain by providing an exploratory environment in which to understand the evolution of genocide; however, validation is essential in order to show

E. M. von Briesen (✉)
Elon University, Elon, NC, USA
e-mail: evonbriesen@elon.edu

M. Garvin · S. Shaikh
UNC Charlotte, Charlotte, NC, USA

Z. Yang and S. Núñez-Corrales (eds.), *Proceedings of the 2022 Conference of The Computational Social Science Society of the Americas*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-031-37553-8_11

that the model can accurately reproduce known dynamics. This research extends the work of von Briesen et al. to develop and validate an agent-based model of genocide. Their model explores the dynamics of genocide from the perspective of individual bystander motivation to intervene and protect persecuted out-group members, and includes a variable that allows for global restraining factors that influences micro-level bystander decision-making. Higher levels of global restraint yield a society in which the average bystander is more likely to work toward the protection of out-group members. Given that micro-level data of human emotion and experience has been sparse until recent years, historic event-validation is especially difficult in this type of model [4].

Genocide researchers rely on historic records to understand the past and influence the future for the better. This may include collection and analysis of elite speech contained in official government records, newspapers, and so on. Returning to the problem of validation of von Briesen et al's model [4], one method of obtaining relevant data is to mine elite speech for emotion, and use that to bridge global, societal-level factors and individual motivation and behavior. Expert qualitative analysis of these types of data is optimal; however, this methodology may not scale sufficiently with larger data sets. With this in mind, we seek to answer the following question: Can emotion mining of historic documents across scenarios provide genocide researchers with useful, quantitative measures to supplement their work, and can such measures be useful for validation of artificial simulations of genocide?

Scott Straus hypothesizes that there is much to be gained through understanding why genocide does not occur when it was expected (a negative case), particularly with respect to how these societies differ from those in which genocide was expected and *did* occur (a positive case) [15]. As part of his work, Straus made available a database of speeches given by presidents and other elite actors from five Sub-Saharan African nations [17].[1] In these speeches, he examines long-standing "themes" present in societies in the years prior to crisis points [16]. The data digitized and analyzed here are a subset of documents from two cases included in the database: Rwanda (positive), and Côte d'Ivoire (negative).

This research combines historic context and sentiment analysis to yield quantitative measures of relevant emotions in select historic texts. We compare and contrast differences in emotion between and within historic cases. We find that despite the limitations of using a modern lexicon on translated language, we are able to measure notable differences in emotion before and after significant events in each country. These results then provide an early starting point for the process of event-validation in von Briesen et al's model of genocide [4].

[1] Straus' original website is currently offline. All speeches and digital transcriptions are available at https://github.com/muniravb/AfricanPresidentialSpeeches. Original speech scans were kindly provided by Dr. Straus.

Fig. 1 Presidential speech document examples. Left: Rwanda, 1-Jul-1969, Right: Côte d'Ivoire, 2-Jan-1979

2 Data Description

Figure 1 shows two samples from Straus' original database. On the left is a document that can be digitized with high accuracy by OCR software, while the poor quality of the document on the right causes a much higher error rate, complicating the digitization process.

Table 1 details the date ranges and significant events for the selected subset of Ivorian and Rwandan speeches. The events are notable along each country's timeline because they mark a point after which ethnic tensions dramatically increased. This subset of documents are selected from the database to provide a similar baseline for comparison. Both cases present elite language from decades-long, post-colonial presidencies. The main inconsistency between these data is the inclusion of speeches following the death of Houphouët-Boigney in Côte d'Ivoire, with no similar data from Rwanda after Habyarimana's death. This is due to the constraints of Straus' original data set.

Table 2 details the total number of digitized documents obtained for both cases and descriptive statistics for word counts.[2] Documents in the database range from governmental publications to newspaper articles. The majority of speeches are in French, with some provided as English translations.

[2] Note that there were a few instances in which the original scan contained multiple speeches. These are split and digitized as separate documents.

Table 1 Description of selected speeches from Straus' African Presidential Speech Database [17]

Country	Date range	Content	Significant event
Rwanda	5-July-1973 to 26-Feb-1994	Presidential and other elite actors' speeches during the presidency of Juvénal Habyarimana.	October 1990—Rwandan Patriotic Front begin invasions of Rwanda from their base in Uganda [11]
Côte d'Ivoire	6-Aug-1961 to 30-Oct-2005	Presidential and other elite actors' speeches during the presidency of Félix Houphouët-Boigney through 12 years after his death	7-Dec-1993—death of President Houphouët-Boigney [12]

Table 2 Descriptive statistics for presidential speeches digitized and analyzed

	Rwanda	Côte d'Ivoire
# Documents	63	60
Min word count	224	122
Max word count	13798	5281
Mean word count	3106	924
Standard dev.	3471	1373

3 Hypothesis

In his research, Straus offers qualitative insights into the speeches he collected. In Rwanda and other positive cases, Straus finds a common presence of "...ideological dominance among the political and military elite of a hierarchical, nationalist founding narrative" [16, p. 275]. Using Habyarimana's 7-December-1990 speech as a example, he states that, "...we see how the founding narrative shapes how Habyarimana interpreted and spoke about the nature of the military. The fight was no ordinary one. It was between those who would protect the revolution and those who would destroy it" [16, p. 294]. Contrast this with Straus' finding that in Côte d'Ivoire, Houphouët-Boigny advocated for "...unity, dialogue, economic growth, stability, peace, and multiethnicity..." as core Ivorian values [16, p. 151]. He hypothesizes that these difference were important restraining factors in negative cases of genocide, as seen in Côte d'Ivoire in the early 2000s.

Straus' database presents a unique opportunity for the field of natural language processing to contribute to social science domains. These data contain elite language from countries that are culturally similar, given all are located in Sub-Saharan Africa, yet significantly different with respect to their ideological narratives and views of

identity groups at the points in time reflected in the database. As such, computational analyses have the potential to quantify measures of sentiment and emotion as they change over time, both within and between cases, supplementing the work of researchers like Straus.

At this early stage, we hypothesize that quantification and analysis of the change in emotion over time in the selected speeches will highlight critical dates and trends. Crocker and Canevello's Ecosystem and Egosystem Theory of Motivational Orientation provides a starting framework for determining how affective states correspond to the motivations of a given individual. They find that emotions such as competitiveness, conflict, confusion, and fear are associated with what they term the "egosystem" motivational state. Here, a person's goals are tied to "self-image" as an extension of the "evolved motivation" for self-preservation. In contrast, "ecosystem" motivation is characterized by emotions such as cooperativeness, peace, love, and clarity. Here, the individual's goals are "compassionate," and extend from the human need for species-preservation [5].

Connecting Crocker and Canevello's theory to Straus' findings in genocide studies, we hypothesize that negative emotions like **anger** and **fear** can be measured relative to the positive emotion of **joy** to quantify the overall motivational orientation of a given speech. Lower values of anger and fear with respect to joy should reflect speech that corresponds to compassionate goals, and a more peaceful, less polarized society. In contrast, higher values should correspond to self-interested goals and a more polarized society prone to conflict. Note that this hypothesis is grounded in the assumption that there is more to be gained by examining composite rather than individual measures of emotion, as this measure tells a larger story of the overall sentiment and motivations of the speaker.

4 Methods

For each speech, we manually removed any text that was *not* an elite actor's speech, such as additional news articles on the same page. When the document quality was sufficient to allow for digital transcription, we used ABBYY FineReader to extract text, manually correcting transcription errors [1]. In some cases, low resolution scans, background lines, smudging, and blurred text in original documents led to extensive digitization errors. We manually transcribed these documents to ensure we curated a complete and accurate set of transcriptions.

As noted above, documents in this data set are in either French or English. Araujo et al. find that translated text can yield better results if it allows for the use of more robust analytical tools [2], and Windsor et al. showed that machine translation can be a stable and reasonably accurate means of reconciling the problem of misalignment between source language and analytic tools [19]. With this in mind, we chose to use the Google API to translate all French language to English [7], and then used the NRC lexicon across all documents to quantify the fraction of words per paragraph expressing anger, fear, and joy [10]. We took the mean of these paragraph-level

fractions over each individual document to calculate a document-level score for each emotion. This process yielded quantification of a discrete set of emotions, by document, through the selected time-frames of each case.

Returning to our hypothesis regarding the usefulness of a composite emotional score, we then calculated document-level composite mean scores as follows:

$$DocScore = (anger_D + fear_D) - joy_D \tag{1}$$

where $anger_D$, $fear_D$, and joy_D, represent the mean score of that emotion for the document.

5 Results and Analysis

Document-level composite scores for Rwanda and Côte d'Ivoire are shown below in Figs. 2 and 3 respectively. Note that in both graphs, negative values indicate that the overall magnitude of joy in a document exceeded the sum of anger and fear. Both figures are annotated to show the approximate location of the "Significant Event" (detailed in Table 1) along that country's timeline. In both cases, anger and fear are clearly the dominant emotions relative to joy after an event that caused a significant increase in ethnic tensions within that nation.

Figure 4 shows the mean composite emotion score for documents before and after the specified event. In both cases, we see a substantial rise in the levels of anger and fear, and decline in joy, after the event. Of note here is that all emotions levels measured higher in the Ivorian data than in the Rwandan data, except for joy in its post-event range. This was unexpected, and requires more work to determine qualitative variations that may explain this result. Additionally, the absence of country-specific terms for hate speech in the lexicon may have artificially lowered the measures of anger and fear, particularly in the Rwandan data.

It is clear that despite many limitations, our methodology does produce a quantification of emotion in these texts that corresponds to expected variations given historic events. The composite measure of document-level emotion from (1) yields expected and informative patterns along the timeline, and the magnitude of mean scores shown in Fig. 4 align with expected changes given significant events.

6 Related Work

Here we address related work as it applies to the limitations of our current methodology. First is the issue of using a modern lexicon on historic language. Recent work in this area has included text annotation and lexicon development with the help of subject-matter experts, leading to performance evaluation for a variety of sentiment and emotion analysis techniques [8, 13]. Variations of these methods would allow

us to determine how closely our measures of emotion fit those of a subject-matter expert, and if customized lexicons and other advanced approaches improve performance. One important goal is to ensure that computational analyses using a lexicon or other methods align with domain expert analyses.

Next, it is essential to account for the use of hate speech in future work. Current research focused on hate speech detection and classification provides insight into methodologies that could be useful [6, 9]. However, given the significance of specific derogatory terms such as "inyenzi" in Rwanda (translated as "cockroaches and used to dehumanize Tutsis") [18, pp. 114–115], a simpler first step would be to customize the NRC lexicon to include terms of this nature for a specific country and time frame.

Finally, difficulties with digital transcription have delayed further progress in converting the remaining speeches in the database to text. Smith and Cordell present an extensive research agenda to address the wide variety of issues that arise when using OCR on historic texts [14]. We look forward to continued development in this area, as greater ability to train software to recognize character and document patterns would reduce the overall time requirement for digitization.

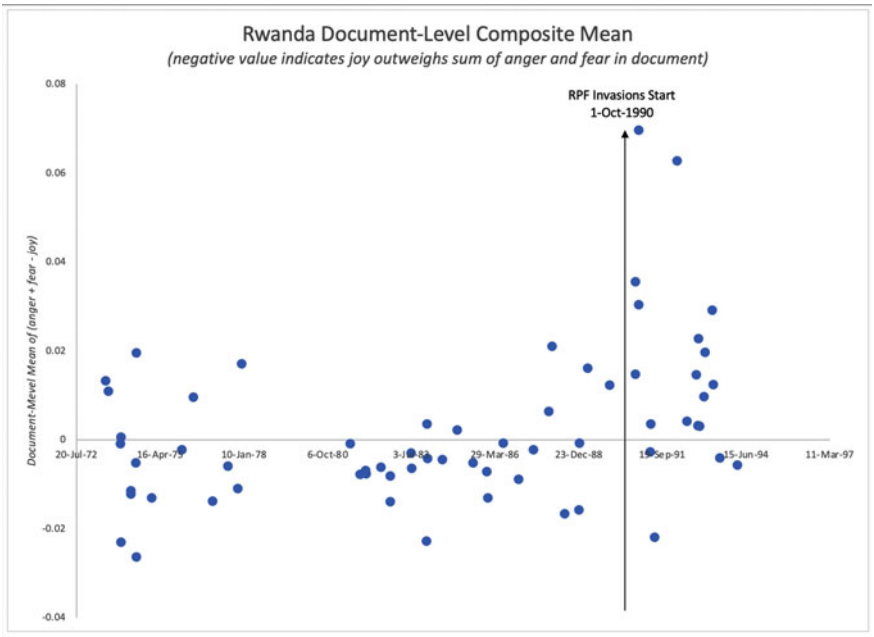


Fig. 2 Rwanda document-level composite emotion scores by date

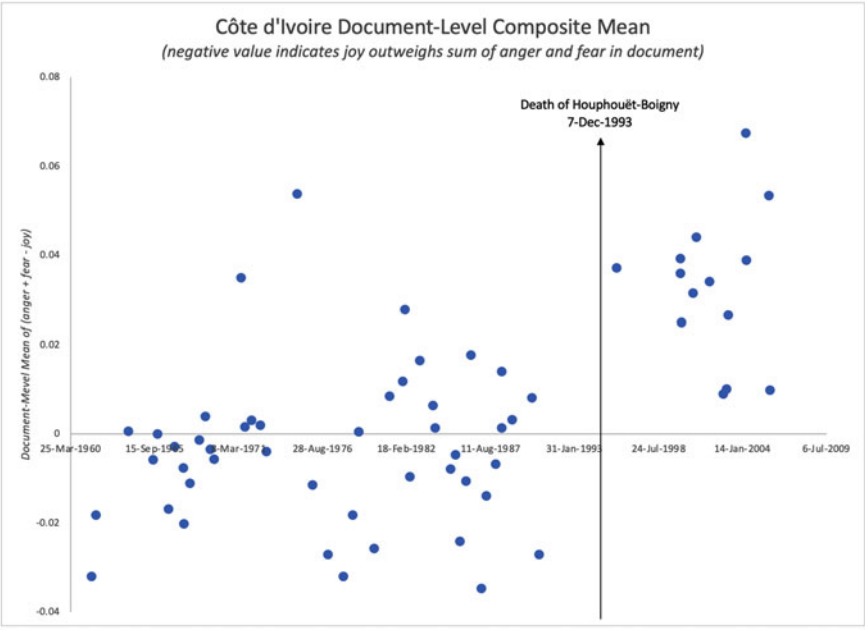


Fig. 3 Côte d'Ivoire document-level composite emotion scores by date

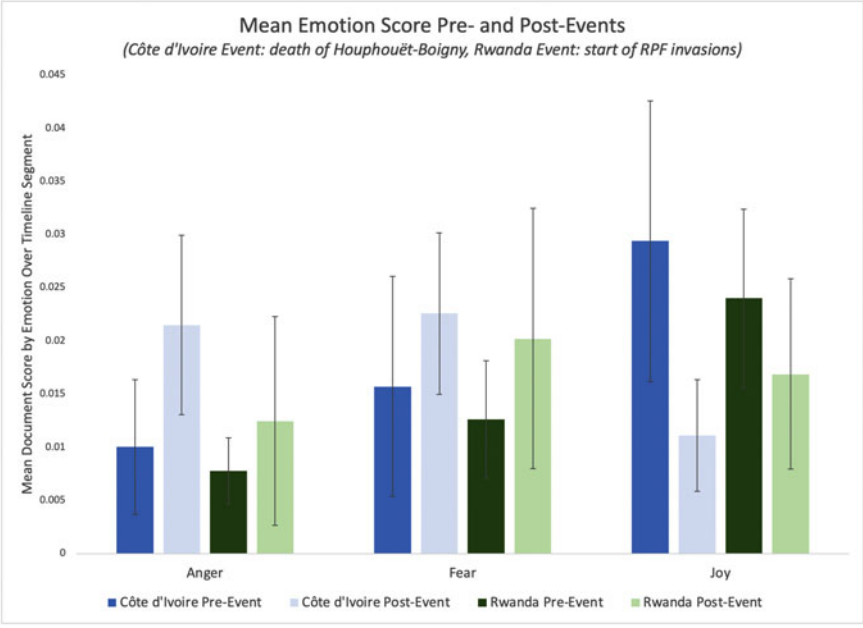


Fig. 4 Mean emotion scores segmented by timeline events

7 Conclusions and Future Work

The research presented here contributes a new approach to analyzing and understanding political text from the field of genocide studies. We have drawn from the work of Scott Straus, digitizing and analyzing the emotional content of a subset of his database of presidential speeches [3, 16, 17]. These first steps confirm that despite the limitations of using a modern lexicon on translated text, measures of emotional content correspond to expected shifts along the historical timeline. However, this approach did not yield results that were useful for quantifying emotional differences *between* the Ivorian and Rwandan cases, which would align Straus' qualitative findings [16].

Confirming Straus' qualitative finding of differing sentiment between scenarios is essential, as the validity of computational results only holds when they align with the findings of subject-matter experts. Achieving this will also be more useful in validating von Briesen et al's model, as this could allow for event-specific calibration and comparison of outcomes. In order to achieve this goal, future work should explore including additional emotions in the composite score (see 1), annotation and evaluation of digitized texts by subject-matter experts to allow for a more careful assessment of analytical accuracy, improved lexicons to measure hate speech in different cases, adoption of more advanced OCR technologies as they becomes available, and comparison of results when applying this methodology to modern, elite speech.

References

1. ABBYY: FineReader OCR Pro for Mac (2013). https://www.abbyy.com/en-us/finereader/pro-for-mac/
2. Araújo, M., Pereira, A., Benevenuto, F.: A comparative study of machine translation for multilingual sentence-level sentiment analysis. Inf. Sci. **512**, 1078–1102 (2020). https://doi.org/10.1016/J.INS.2019.10.031
3. von Briesen, E.: African Presidential Speeches. https://github.com/muniravb/AfricanPresidentialSpeeches
4. von Briesen, E., Canevello, A., Shaikh, S., Cox, J., Hadzikadic, M.: Modeling genocide: an agent-based model of bystander motivation and societal restraints. In: Proceedings of the 2019 International Conference of The Computational Social Science Society of the Americas. Springer, Berlin (2019)
5. Crocker, J., Canevello, A.: From egosystem to ecosystem: motivations of the self in a social world. In: Elliot, A. (ed.) Advances in Motivation Science, 1st edn., vol. 5, chap. 2, pp. 41–86. Elsevier Inc., New York (2018). https://doi.org/10.1016/bs.adms.2018.01.003, http://www.sciencedirect.com/science/article/pii/S2215091918300038
6. Gitari, N.D., Zuping, Z., Damien, H., Long, J.: A lexicon-based approach for hate speech detection. Int. J. Multimed. Ubiquitous Eng. **10**(4), 215–230 (2015). https://doi.org/10.14257/ijmue.2015.10.4.21
7. Google: googletrans 2.4.0 (2020). https://pypi.org/project/googletrans/
8. Hellrich, J., Buechel, S., Hahn, U.: Modeling word emotion in historical language: quantity beats supposed stability in seed word selection. In: Proceedings of the 3rd Joint SIGHUM

Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature, pp. 1–11. Association for Computational Linguistics (ACL), Minneapolis, USA (2019). https://doi.org/10.18653/V1/W19-2501, https://aclanthology.org/W19-2501
9. Martins, R., Gomes, M., Almeida, J.J., Novais, P., Henriques, P.: Hate speech classification in social media using emotional analysis. In: Proceedings—2018 Brazilian Conference on Intelligent Systems, BRACIS 2018, pp. 61–66. Institute of Electrical and Electronics Engineers Inc. (2018). https://doi.org/10.1109/BRACIS.2018.00019
10. Mohamad, S.: NRC Emotion Lexicon (2010). http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm
11. Nguyen, K.: TIMELINE-Rwandan genocide sparked by burning ethnic tensions (2014). https://news.trust.org/item/20140402113037-u315s/
12. Noble, K.B.: Felix Houphouet-Boigny, Ivory Coast's Leader Since Freedom in 1960, Is Dead—The New York Times (1993). https://www.nytimes.com/1993/12/08/obituaries/felix-houphouet-boigny-ivory-coast-s-leader-since-freedom-in-1960-is-dead.html
13. Schmidt, T., Burghardt, M.: An evaluation of lexicon-based sentiment analysis techniques for the plays of gotthold ephraim lessing. In: Proceedings of the Second Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature, pp. 139–149. Association for Computational Linguistics, Santa Fe, NM (2018). https://aclanthology.org/W18-4516
14. Smith, D., Cordell, R.: A Research Agenda for Historical and Multilingual Optical Character Recognition—DRS (2018). https://repository.library.northeastern.edu/files/neu:f1881m035
15. Straus, S.: Retreating from the Brink: Theorizing Mass Violence and the Dynamics of Restraint. Perspect. Polit. **10**(02), 343–362 (2012). https://doi.org/10.1017/S1537592712000709
16. Straus, S.: Making and Unmaking Nations: War, Leadership, and Genocide in Modern Africa, 1st edn. Cornell University Press (2015)
17. Straus, S.: African Presidential Speeches Database (2020). https://faculty.polisci.wisc.edu/sstraus/african-presidential-speeches-database/
18. Twagilimana, A.: Historical Dictionary of Rwanda—Aimable Twagilimana—Google Books, 2nd edn. Rwoman & Littlefield, London (2016)
19. Windsor, L.C., Cupit, J.G., Windsor, A.J.: Automated content analysis across six languages. PLOS ONE **14**(11), e0224425 (2019). https://doi.org/10.1371/JOURNAL.PONE.0224425, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0224425

An Agent-Based Model to Explore Belief and Behavioural Change in a Classroom

Keegan Fernandes, Daniel Davison, and David Wang

Abstract A student classroom is a complex environment of social dynamics through which instructors attempt to navigate and find ways to improve their students' behaviour. These complex dynamics make it difficult to predict how a student will react to an instructor's interventions. We build an agent-based model to examine how an individual's beliefs might spread through a classroom, and how they may affect behaviour, to be able to identify how an instructor can approach a classroom with the hopes of encouraging better note taking behaviour. The model is built on existing psychological models and consists of three parts: a behavioural change component in the form of the reasoned action approach, a social influence component, and an outcome feedback component. The agent's beliefs determine their behaviour and are influenced by their peers' beliefs. The results of this paper show that the model exhibits expected psychological phenomena such as polarisation and clustering. Additionally, we explore various possible instructor influence techniques such as an "all student" technique, an "influencer" technique, a "rotating" technique, and a more "targeted" technique.

1 Introduction

The student classroom provides an interesting backdrop for social dynamics. There are many effects at play such as student pressures, instructor feedback, and group structures, all of which give rise to complex social phenomena. Within this elaborate environment instructors are constantly trying to find ways to improve student performance. One technique often used is to encourage good studying habits such as note taking, scheduling, and self-regulated learning [12]. Studies show that proper note taking is very beneficial for recall and allowing people to better process information [10]. The aim of this paper is to use fundamental psychological and sociological effects to develop an agent-based model that can simulate the behaviours of stu-

K. Fernandes (✉) · D. Davison · D. Wang
University of Waterloo, Waterloo, ON N2L 3G1, Canada
e-mail: ka3ferna@uwaterloo.ca

Z. Yang and S. Núñez-Corrales (eds.), *Proceedings of the 2022 Conference of The Computational Social Science Society of the Americas*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-031-37553-8_12

dents in a classroom setting. More specifically, a scenario where an instructor tries to encourage their students to take up better note taking habits in order to improve academic success is used. This model could aid instructors in understanding how to effectively implement interventions to encourage better study habits and has the potential of positively impacting students by improving their understanding of course material.

While instructors have many strategies to try and encourage the use of good studying behaviours, this paper focuses on four main techniques. The first is the "all student" technique, where the instructor encourages the class as a whole by supplying information about a behaviour and its benefits, providing a generalised intervention [8]. Second, the instructor takes advantage of the class hierarchy to spend their time encouraging the most influential students in a class with the hopes that their behaviour will propagate to the rest of the class through social pressures [16]; this is called the "influencer" approach. Third, the "rotating" technique involves the instructor rotating daily through different individual students to try and encourage them to perform the behaviour, providing more specific interventions but at a cost of spending less time with each student [11]. Lastly, there is a "targeted" approach, where the instructors identifies the student most likely to change beliefs about the behaviour and encourages them to do so (see Fig. 1). The goal of this paper is to be able to identify the effects of these different strategies in a class and perform simulations to get initial insights into what might happen.

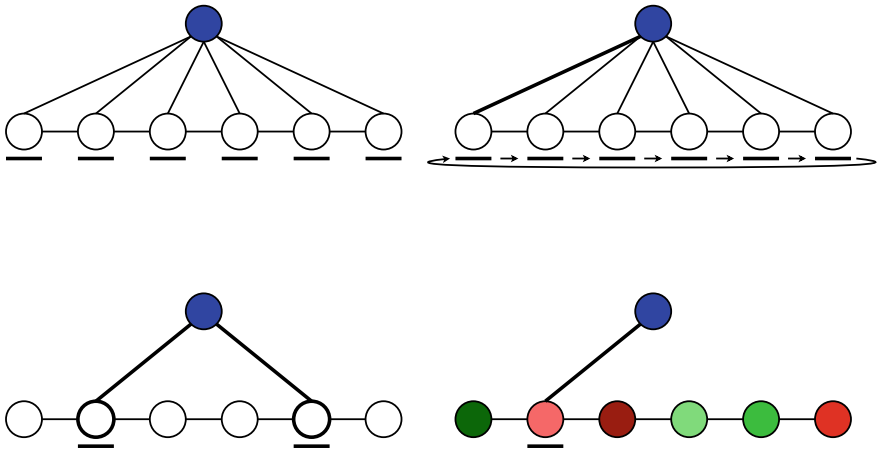


Fig. 1 Top Left: Instructor weakly influencing many students (all student). Top Right: Instructor interacting with individual students one at a time in rotation, identified by moving underlines (rotating). Bottom Left: Instructor strongly influencing few students, who in turn influence more students (influencer). Bottom Right: Instructor strongly influencing student on the verge of changing beliefs (dark red being negative and dark green being positive) (targeted)

2 Model Components

To simulate how an individual's behaviour and beliefs evolve in a group setting, the model incorporates three main components:

- A behavioural change model, in the form of the Reasoned Action Approach (RAA), to explain how and why human behaviours are changed.
- A social influence effect, in the form of conformity and confirmation bias, to explain how social interactions affect people's beliefs.
- Outcome feedback, to encapsulate how the outcomes of a person's actions affects their beliefs.

2.1 Reasoned Action Approach

The understanding of human behavioural change has been explored by many researchers and has resulted in models such as the reasoned action approach (RAA) [2, 8]. RAA states that, given a clearly identified behaviour, a person will choose to perform this behaviour by implicitly considering underlying information and their beliefs. These beliefs fall into three categories: behavioural beliefs, normative beliefs, and control beliefs. Each of these beliefs respectively determine a person's attitude towards performing the behaviour, a perceived social norm towards the behaviour, and a perceived behavioural control with regard to the behaviour. The theory then states that a person's attitudes, perceived norms, and perceived behavioural control help determine whether they perform a behaviour. A schematic of the reasoned action approach can be found in Fig. 2.

The frequency that an agent i performs a behaviour, $B_i \in [-1, 1]$, is determined by a weighted average of their attitude, $A_i \in [-1, 1]$, perceived norm, $N_i \in [-1, 1]$, and perceived behavioural control, $C_i \in [-1, 1]$. When $B_i = -1$ the agent does not perform the behaviour, when $B_i = 1$ the agent performs the behaviour frequently (daily), and when $B_i = 0$ the agent performs the behaviour moderately (weekly). Equations can be found in Appendix A. In the classroom scenario the behaviour of interest is note taking. The attitude then would be whether a person thinks note taking is good or bad, the perceived norm is the perception that others think note taking is good or bad, and the perceived control is whether the person believes they are capable of note taking.

RAA further states that a person's attitudes, perceived norms, and perceived behavioural control can be determined by a weighted sum of the individual's beliefs using an expectancy-value model [3, 5, 7]. It states that a person's behavioural beliefs, $b_{i,l}$, normative beliefs, $n^I_{i,l}$ and $n^D_{i,l}$, and control beliefs, $c_{i,l}$, are weighted by their respective importance and whether they are beneficial or not, $e_{i,l}$,$m^I_{i,l}$ and $m^D_{i,l}$, and $p_{i,l}$. Equations can be found in Appendix A. In the classroom setting the behavioural beliefs being considered are that note taking improves memory retention

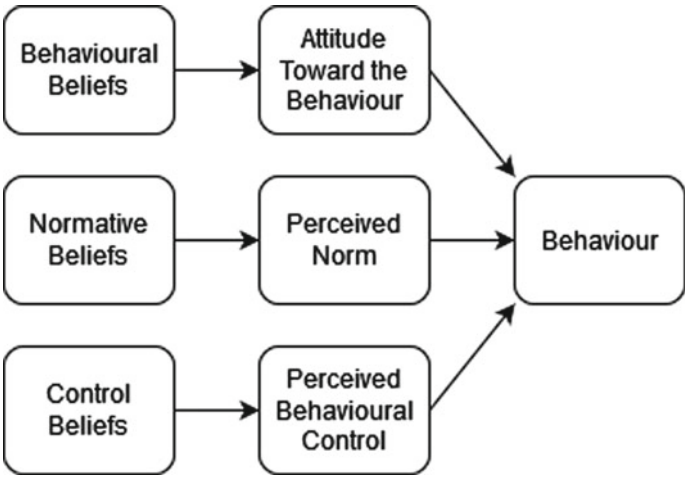


Fig. 2 Schematic of the reasoned action approach model [8]

and that note taking will take time away from other assignments. The normative belief being considered is that their peers believe that note taking is beneficial. Finally, the control belief being used is the belief that the person can access the resources necessary to allow them to spend time note taking. Each of these beliefs determines their respective attitudes, perceived norms, or perceived control.

2.2 Social Influence

Social influence is an aspect of human interaction whereby individuals modify their opinions, beliefs, attitudes, and behaviour towards that of those they are close to and interact with [16]. The reasons behind these influences vary from being persuaded [13] to a social pressure to conform with social norms [1, 6, 19]. The means of social influence is studied in the field of social psychology, where one of the key ideas is that of conformity [4]. Additionally, ideas such as confirmation bias try to explain why some people tend to have polarised or different opinions [15].

Conformity Conformity is the phenomenon of people changing their behaviour or beliefs to match those around them [4]. It is a social effect that arises from a norm being established by a group of individuals. People tend to conform to this norm rather than their own desires because it is easier to do so, and they are sometimes persuaded to do so via social pressures [1].

This phenomenon can be represented by an assimilative model which assumes that all connected individuals can influence each other [9]. This model states that an individual's beliefs move towards, at some convergent rate μ, the weighted average of the beliefs of everyone in the system. According to [9], this form of social influence

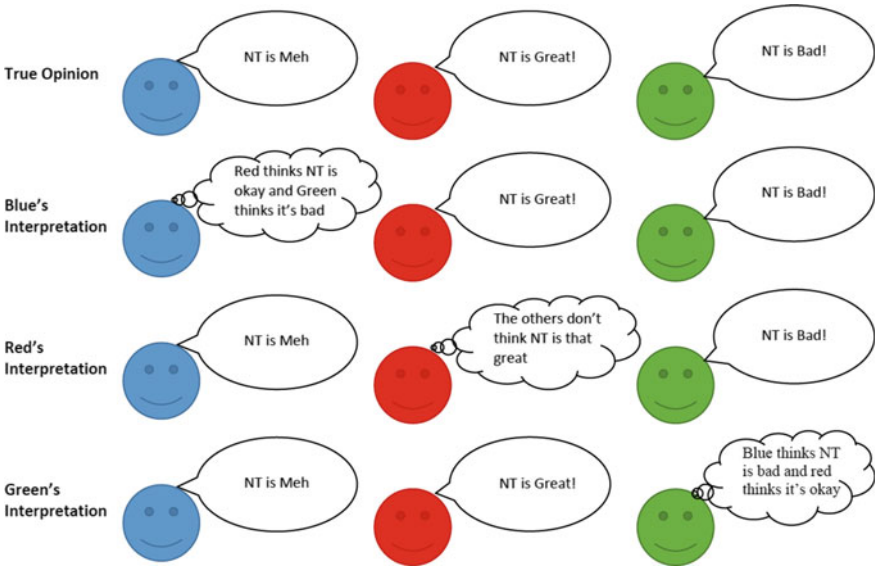


Fig. 3 Demonstration of interpretation bias

leads to the eventual consensus of opinion in all individuals in the network, consistent with the idea of conformity.

Confirmation Bias Social influence theory also includes the idea that not all individuals conform to one another. There are situations in which individuals have differing opinions and do not necessarily come to a consensus. In the extreme case, individuals get polarised and have diametrically opposing views. Confirmation bias is a possible mechanism for this phenomenon. Confirmation bias is the tendency for people to find evidence that validates their existing beliefs [15]. The form of confirmation bias implemented in this paper is a biased interpretation. This bias leads individuals, who are presented the same information, to interpret the information differently based on their existing beliefs, favouring their view [17]. Figure 3 provides an example where three individuals, Blue, Red, and Green each believe that note taking (NT) is okay for you, good for you, and bad for you respectively. Additionally, Blue has a large interpretation bias, Red has a small bias, and Green has a medium bias. The figure demonstrates how the interpretation biases can lead to differing opinions.

This bias is included in the model by simply adding an offset when an agent interprets any information presented to them. This offset is added in the direction that supports the agent's underlying belief; for example, it is positive if the agent's belief is positive and vice versa.

2.3 Outcome Feedback

In addition to social influence, the reasoned action approach mentions that certain attributes such as behavioural beliefs, normative beliefs, and control beliefs are susceptible to feedback based on the performed behaviour [8]. These beliefs are influenced by an individual's own behaviour, attitudes, and the outcomes of the behaviour performed. For example, if agent i performs the desired behaviour, they can compare the outcome, R_i, of the behaviour to their behavioural belief (did note taking result in improved retention?) and will attempt to match their belief to the outcome. This results in an agent's beliefs changing to match the outcome of the behaviour.

2.4 Building the Model

The overall agent model is built by incorporating the RAA model, the conformity and confirmation bias effects, and outcome feedback together. A simplified diagram for the case of two agents is shown in Fig. 4.

We can combine all the models presented in this section to get a generalised model for n agents. For simplicity, the resulting equations for a two agent system are listed in Appendix A. The system presented is a discrete-time system where k is a step of a day in the classroom where all students have interacted with those in their

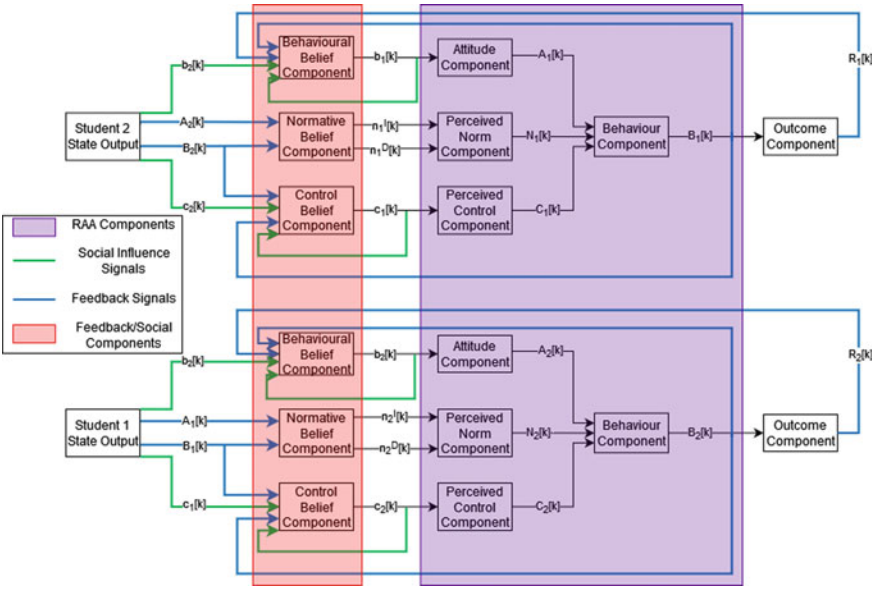


Fig. 4 Combined model block diagram for 2 agents

social circle. For the outcome component, we assume that note taking does result in increased memory retention as well as time taken from other assignments. This means that the outcome is equal to the amount the behaviour is performed.

The behavioural component (see Fig. 4) determines how a behaviour will be performed based on a weighted average of the attitudes, perceived norms, and perceived behavioural control, as discussed in Sect. 2.1.

The attitude component, perceived norm component, and perceived control component (see Fig. 4) each explain how one's behavioural beliefs, normative beliefs, and control beliefs affect attitude, perceived norm, and perceived behavioural control respectively. These beliefs are summed and weighted by their value, motivation to comply, and power respectively.

The next components of behavioural belief, normative belief, and control belief (see Fig. 4) update their respective beliefs with feedback and social influence effects. In each of these cases there are two phases involved with updating the core beliefs, one pertaining to the feedback step, and the other to the effects of social influence. Both incorporate signed biases in their input signals to account for confirmation bias.

The behavioural belief component is first updated with an outcome feedback effect, as described in Sect. 2.3. The effects of conformity are then accounted for, moving the agent's beliefs towards a weighted average of the connected behavioural beliefs, as described in Sect. 2.2. The normative belief component is similarly handled in two steps. First, attitudes and behaviours are fed back to the agents. Next, conformity moves the agents normative beliefs towards the weighted average of all of those that interact with each other. Lastly, the control belief component is also handled in the same way. Feedback effects, this time involving behaviour, are first accounted for. After this, conformity moves the agent's control beliefs towards the weighted mean.

In the reasoned action approach, the means of changing peoples' behaviour are usually interventions designed to inform an individual about their respective beliefs [8]. In the model designed in this section, such interventions are initiated by the instructor and would appear in the form of additional sources of influence in the belief components. Therefore, our control inputs would appear as an additional control agent (i.e., the instructor) in our multi agent system. This control agent would be an additional source of influence on the other agents. However, the control agent itself is unaffected by other agents (i.e., students) in the system. In the case of the classroom example, an instructor would be able to influence the class with their own beliefs with respect to the behaviour. By informing students of the benefits of note taking, they may be able to sway the students' beliefs about the behaviour.

3 Simulation and Results

In this section, the model is first simulated to validate that it reflects existing psychological expectations. The model is then investigated using different forms of instructor influence. The behaviour being enforced is note taking, and consists of

two behavioural beliefs (note taking helps with memory retention and that it takes time away from other tasks), one normative belief (their peers are performing and believe in the behaviour), and one control belief (they can access resources to achieve the behaviour). There are 30 students in the system and all students are able to communicate with each other (it is a fully connected system). The instructor is considered another agent in the system who's dynamics are unaffected by others and who's states are fully controllable.

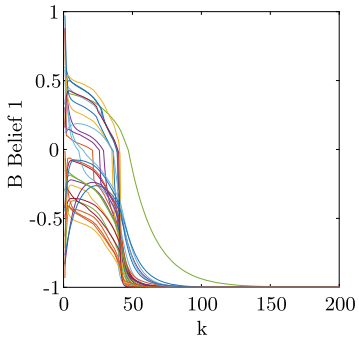
3.1 Validation Results

In this section no instructor is included. The simulation was run for 200 time steps and repeated 100,000 times with parameters and initial conditions for each agent randomly chosen from a uniform distribution. From observing these results, it is clear that there are two main types of outcomes with this model, polarisation of beliefs (Fig. 5), or clustering of beliefs (Fig. 6).

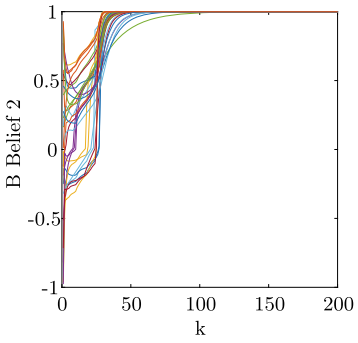
Figure 5 is a sample of one of the many simulations run. From this figure we can see that the students begin with a spread of initial beliefs and behaviours. However, in this case, the parameters of these students were such that the students with negative beliefs eventually overwhelm those with positive beliefs and lead to a negative outcome. We see that, initially, the students begin to converge in beliefs, moving towards the average of the class. However, as students begin to change beliefs from being optimistic about the behaviour to being pessimistic, the effects of confirmation bias causes significant changes in beliefs as the students begin to interpret information in a way that suits their new found direction. This confirmation bias leads to students interpreting that their peers' beliefs are more negative than they actually are, leading to a vicious feedback loop that causes beliefs to spiral further and further towards the negative.

This spiraling behaviour leading to strongly polarised beliefs is reminiscent of group polarisation [14]. Group polarisation is the tendency for a group of people to have more extreme behaviours and beliefs than that of the individuals in the group, leading to the individuals to eventually have a stronger belief than what they stated with. Group polarisation is sometimes explained by informational influences [18], where the sharing of new information helps an individual become more convinced of their initial beliefs. The emergence of this effect is understandable due to the confirmation bias involved in the model and lends credence to its validity.

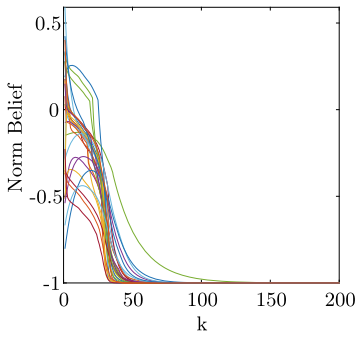
Another sample of the simulations, Fig. 6, displays the other dominant outcome, that of clustered opinions. From the figure we can see that with a mixture of initial conditions the students eventually get to a point where they have clusters of varying beliefs forming. The main driving force of this is the confirmation bias in each student. While there is a convergence effect that causes students to try and reach a consensus, some students have a large confirmation bias causing them to misinterpret their peers' beliefs. This leads to students having a perceived consensus of beliefs where in reality they have different beliefs.



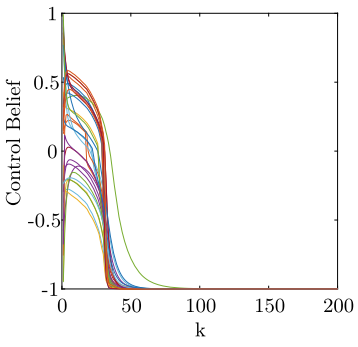
(a) Change in First Behavioural Belief



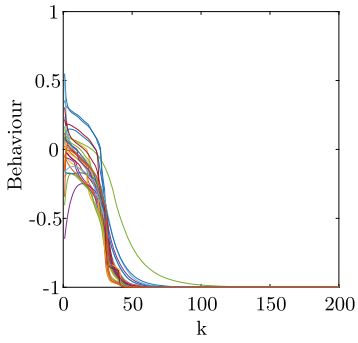
(b) Change in Second Behavioural Belief



(c) Change in Normative Belief

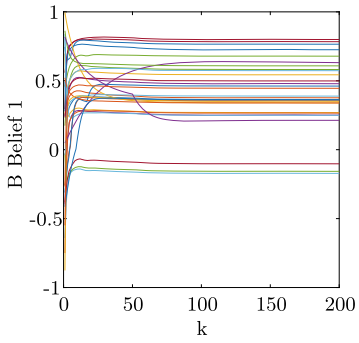


(d) Change in Control Belief

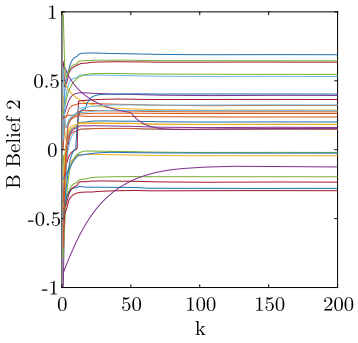


(e) Change in Behaviour

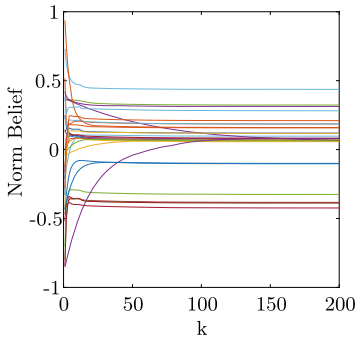
Fig. 5 Polarisation behaviour of a 30 person class



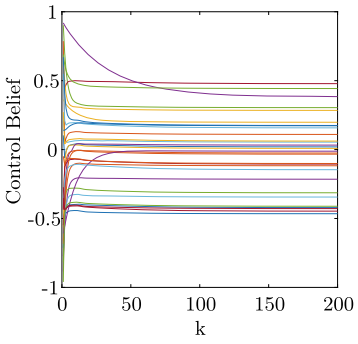
(a) Change in First Behavioural Belief



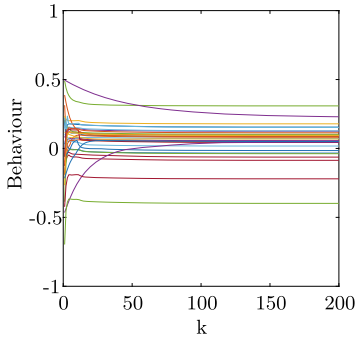
(b) Change in Second Behavioural Belief



(c) Change in Normative Belief



(d) Change in Control Belief



(e) Change in Behaviour

Fig. 6 Clustering behaviour of a 30 person class

This clustering of beliefs is reminiscent of what is seen in a real classroom with groups of students forming varying sets of beliefs about a subject and is reflective of what would be expected from similarity based social models [9], further validating the model.

3.2 Instructor Influence

The end purpose of this model is to find techniques to enable instructors to best influence students to take up good studying habits. With that in mind, we simulate the effects of four types of influence approaches: influencing all students, focusing on influential students, focusing on one different student at a time, and focusing on students that are close to changing opinions. Since, according to the reasoned action approach [8], behaviour is determined by an individual's underlying beliefs and the suggested intervention for changing behaviour is to target these beliefs, the influence methods stated in this section will focus on changing an individual's beliefs. Specifically, we will focus on the first behavioural belief (Note taking helps with memory retention) and, if successful, this same technique can be used on the other stated beliefs as well.

Each technique was simulated for 200 time steps and repeated 100,000 times with random initial conditions and parameter values. This enables us to better understand which approach is more effective in general. We are interested in increasing the beliefs of each student in the class, so to measure the performance of each technique we define a metric equal to the total change of the belief value compared to the no input case (see (19)). For example, if each of the 30 students increases their belief by 0.2 after the technique has been applied, the metric of the technique would measure $(0.2 \times 30) = 6$. Using this metric, the best possible value achievable would be if all students started with their belief at -1 and ended with belief at +1, leading to a metric of $(1 - (-1)) \times 30 = 60$. However, using a uniform distribution initial condition, we can expect that on average the students will start with a belief of 0. This leads to an expected best case metric of $(1 - 0) \times 30 = 30$. A summary of the results is provided in Figs. 7 and 8.

All Student For the first influence technique, we simulate a case where the instructor is trying to convince the entire class simultaneously of the benefits of note taking towards memory retention. By taking this generalised approach the instructor would have less of an ability to convince any one student in the class. That is, the instructor has less influence on each student. To account for this, the model reduces the amount of influence the instructor has proportionally to the number of students they interact with.

Figure 7 shows that the all student technique, based on 100,000 runs, has a distribution where $\sim$30% of the time the metric was greater than 5. This means that the strategy of the instructor trying to persuade the whole class does, in general, improve the beliefs of the class. However, it does not have a strong improvement, with each student's belief only increasing marginally on average. There are also some instances

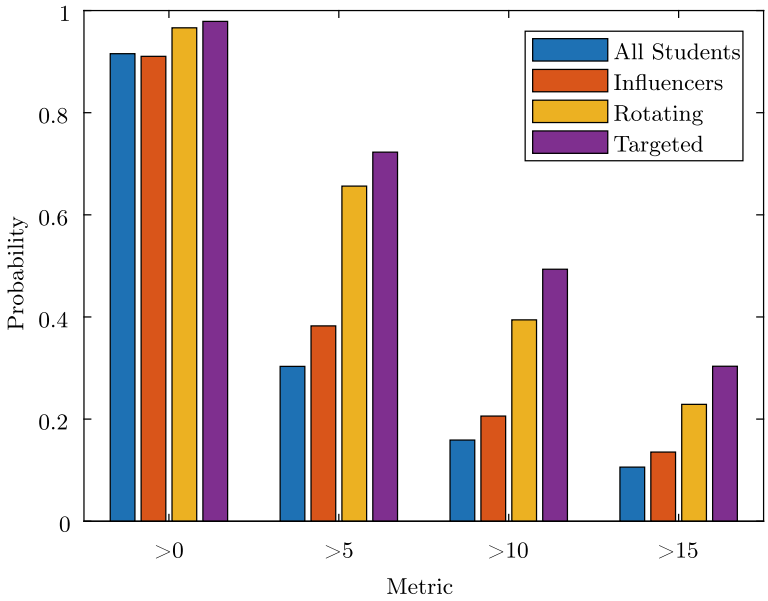
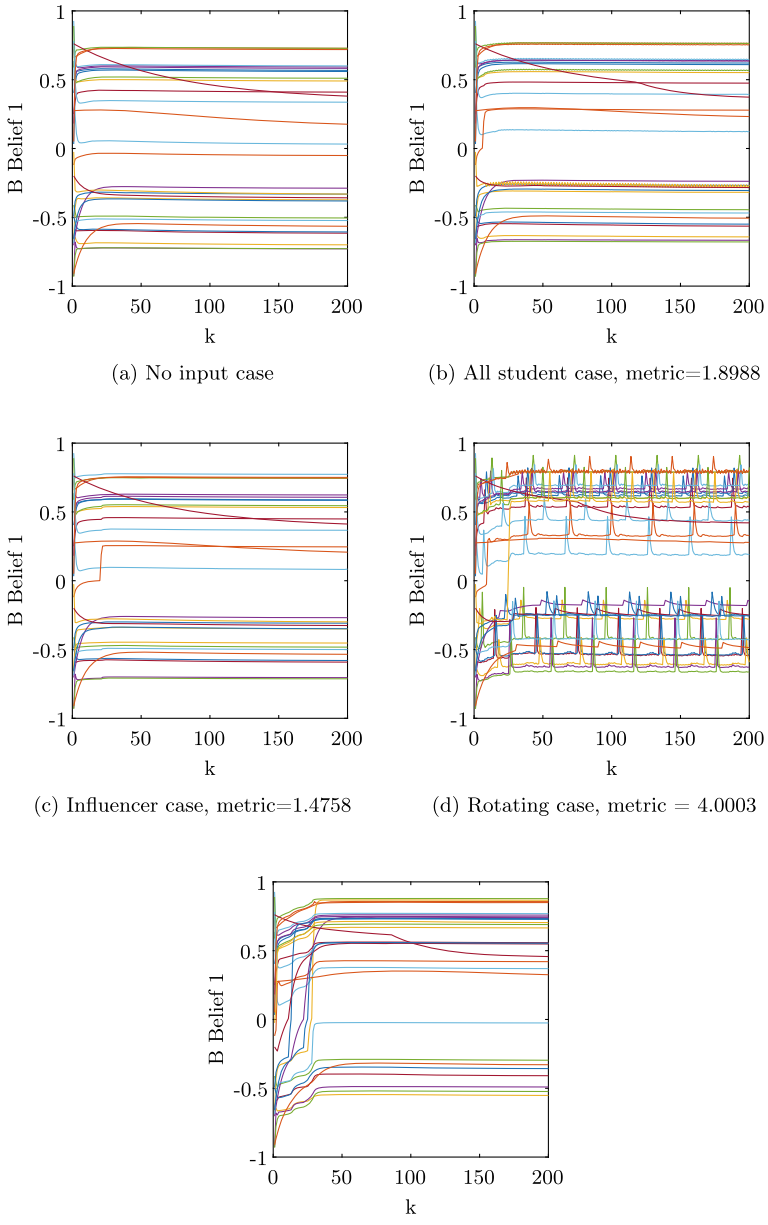


Fig. 7 Probabilities of instructor techniques going above a metric threshold based on 100,000 random simulations

where this technique, and the other listed techniques as well, can have a negative metric. This translates to some scenarios where the influence technique causes the class to end up with a worse belief compared to if no influence was being exerted. This arises from situations where the student begins to perform the behaviour while still having a large negative bias towards the belief being beneficial. When this happens the student interprets the outcome of the behaviour as being overall negative, and can cause the student to believe that the behaviour is not beneficial. This shows that is important to implement a technique strategically, as sometimes, the class can end up worse than if no intervention was applied.

Figure 8b shows a typical sample of the 100,000 simulations run. From this figure, we can see that the technique was able to improve the belief of the students slightly when compared to the no input case. The largest contribution to this improvement comes from a student who flips from having a negative belief to having a positive belief at $k = 5$. When this happens, the now positive confirmation bias of the student kicks in, and causes the student to quickly increase in their belief. This is a key observation.

Influencers The second technique examined takes advantage of the existing influential students in the class to allow the instructor to more easily spread beliefs. The class simulations are designed with two students that are, on average, twice as influential as any other student in the class. More influential students can sway the beliefs of their peers more easily than less influential students. In this technique, the instructor focuses their energy on only these two students, with the hope that, by



(a) No input case

(b) All student case, metric=1.8988

(c) Influencer case, metric=1.4758

(d) Rotating case, metric = 4.0003

(e) Targeted case, metric = 11.1962

Fig. 8 Sample of effects of different instructor influence techniques

more effectively convincing the influential students that the behaviour is beneficial, these students can then convince their peers of its usefulness. By only approaching two students, the instructor's relative influence is kept high, since they are not being spread thin by 30 students.

Figure 7 shows that the influencer technique has a distribution with ∼38% of cases had a metric greater than 5. This is higher than the all student technique, but not by a lot. However, this technique also has a better chance of achieving higher metrics. From the sample example shown in Fig. 8c, we see a case where the influencer technique performs slightly worse than the all student technique, which can happen but isn't typical. Similar to the all student case, the students improve in their beliefs, with the largest contribution happening from a student flipping form negative to positive belief.

Rotating The third technique is one where the instructor spends time with only one student each day, repeatedly cycling through all of the students. This allows the instructor to convince the student to the best of their ability for the one day they spend with them. However, this poses the risk of the student reverting back to their original state in the time between visits.

Figure 7 shows that this rotating technique has a much higher probability of achieving a greater than 5 metric at ∼66%. This result is dramatically higher than the previous two methods examined and seems like a promising technique. Not only does it have a higher probability of achieving a metric greater than 5, but it also has a significantly better chance of achieving metrics greater than 10 and 15 as well. Figure 8d demonstrates how this technique operates. At first glance this graph is very noisy, but upon closer inspection, it is clear that there are large, periodic increases in a student's belief whenever the instructor spends time with them. The instructor is able to impart a large amount of influence on a student, causing their belief to spike. As the instructor moves over to a different student this spike diminishes over time, until the instructor comes by once again. These large spikes in belief can occasionally cause students to flip from a negative belief to a positive one leading to the effects previously demonstrated. With the rotating technique, the instructor's large influence is, in general, able to pull more students over from having a negative belief to a positive belief.

Targeted The final technique being examined is that of identifying and targeting key students. Inspired by the ideas demonstrated in the previous simulations a technique is proposed to try and maximise the number of students that flip from having negative beliefs to positive beliefs. To achieve these flips, this technique identifies students that have the negative beliefs but are close to flipping (i.e. the least negative beliefs). The instructor should, in theory, easily be able to flip these student's beliefs.

Figure 7 shows that the targeted approach has a large probability of achieving a greater than 5 metric at ∼72%. This is a large improvement over the other methods and even has a larger greater than 10 and 15 probability when compared to the rotating technique. This technique of targeting those students that are close to flipping beliefs but haven't yet seems to be very effective. Figure 8e provides a sample of how this approach operates. We can see that students quickly switch sides one by one, until the instructor is stuck with a reluctant student. This leads to a large number of flipped

beliefs, as evident by the few remaining students with negative beliefs. This technique could further be improved by allowing the instructor to move over to another student if one seems to be unaffected by the interventions being applied.

4 Conclusion

In this paper we proposed a hypothetical model for simulating the behaviours of students in a classroom as well as proposed techniques to better encourage students to take up better note taking habits. The model was able to recreate common psychological phenomena such as polarisation and clustering, lending credence to its validity. By using this model we can see how students might be influenced to perform certain behaviours as well as how beliefs spread within a classroom, and how to take advantage of these phenomena. We experimented with different influence techniques to gain insight into how an instructor could try and convince a class to take up better note taking habits. This exploration reveals the potential benefits of using such a model to evaluate teaching techniques. It could also allow instructors to test unintuitive strategies such as the targeted case presented in Sect. 3.2.

This paper has limitations, namely, the model currently does not use real world data for its parameter values. This model can be further refined by only allowing parameter values that fall in a distribution of more likely real world possibilities. To obtain these values we plan on running a survey in the style proposed in the RAA [8].

Further studies for this research include exploring more efficient influence techniques. One approach would be to try and find optimal influence techniques that maximise the metric of a given class while minimising instructor effort and resources. Additionally, a strategy of optimising group connections could be tested. It would be of interest to see how incorporating class groups would affect the spread of beliefs. By making students form groups the amount of influence their group mates have should increase relative to their peers; this could be used to encourage students with negative beliefs to more easily conform to their more positive peers.

A. Agent Model Equations

Equations for a two-agent system are listed here. The outcome component of the system is as follows,

$$R_{i,l}[k] = B_i[k] \tag{1}$$

where $R_{i,l}[k]$ is the outcome component for agent i associated with belief l, B represents the agent's behaviour, and k represents the day.

Next, the behavioural component determines how a behaviour will be performed,

$$B_i[k] = \lambda_i A_i[k] + \beta_i N_i[k] + \gamma_i C_i[k],\ \lambda_i + \beta_i + \gamma_i = 1\ \lambda_i, \beta_i, \gamma_i > 0 \quad (2)$$

where A represents attitude, N represents perceived norms, and C represents perceived behavioural control. $\lambda_i, \beta_i, \gamma_i$ are gains that weight how these components affect behaviour.

The attitude component, perceived norm component, and perceived control component are dictated by the following equations,

$$A_i[k] = \frac{\sum_l b_{i,l}[k] e_{i,l}}{\sum l} \quad (3)$$

$$N_i[k] = \frac{1}{2\sum l}\left(\sum_l n^I_{i,l}[k] m^I_{i,l} + \sum_l n^D_{i,l}[k] m^D_{i,l}\right) \quad (4)$$

$$C_i[k] = \frac{\sum_l c_{i,l}[k] p_{i,l}}{\sum l} \quad (5)$$

where l is the number of beliefs associated with each perception.

Each of the next components of behavioural belief, normative belief, and control belief (see Fig. 4) involves two phases when updating the core beliefs, one pertaining to the feedback step, the next to the effects of social influence with additional signed biases. The behavioural belief component is as follows,

$$sat(x) = \min(\max(x, -1), 1) \quad (6)$$

$$R_{i,l}[k]_{bias} = sat(R_{i,l}[k] + sgn(b_{i,l}[k]) h^R), \quad (7)$$

$$\hat{b}_{i,l}[k+1] = b_{i,l}[k] + \mu_i^{AR}(\alpha^{AR} R_{i,l}[k]_{bias} - b_{i,l}[k]), \quad (8)$$

$$\hat{b}_{j,l}[k+1]_{bias} = sat(\hat{b}_{j,l}[k+1] + sgn(b_{i,l}[k+1]) h^b), \quad (9)$$

$$\begin{aligned} b_{1,l}[k+1] &= \hat{b}_{1,l}[k+1] + \mu_1^A(\hat{b}_{2,l}[k+1]_{bias} - \hat{b}_{1,l}[k+1]) \\ b_{2,l}[k+1] &= \hat{b}_{2,l}[k+1] + \mu_2^A(\hat{b}_{1,l}[k+1]_{bias} - \hat{b}_{2,l}[k+1]) \end{aligned} \quad (10)$$

where $\hat{b}[k]$ is the update step related to feedback effects, while $b[k+1]$ is the step that takes into account the social influence effects, μ is the rate of convergence. Additionally, h is the added offset value to account for confirmation bias and $sat(x)$ is used to ensure that values stay bounded. The normative belief component is described by,

$$B^N_{i,l}[k]_{bias} = sat(B_i[k] + sgn(n^D_{i,l}[k]) h^{BN}), \quad (11)$$

$$\begin{aligned} n^D_{1,l}[k+1] &= n^D_{1,l}[k] + \mu_1^{NB}(\alpha^{NB} B^N_{2,l}[k]_{bias} - n^D_{1,l}[k]) \\ n^D_{2,l}[k+1] &= n^D_2, l[k] + \mu_2^{NB}(\alpha^{NB} B^N_{1,l}[k]_{bias} - n^D_{2,l}[k]) \end{aligned} \quad (12)$$

$$A_{i,l}[k]_{bias} = sat(A_{i,l}[k] + sgn(n^I_{i,l}[k])h^A), \tag{13}$$

$$\begin{aligned} n^I_{1,l}[k+1] &= n^I_{1,l}[k] + \mu^{NA}_1(\alpha^{NA} A_{2,l}[k]_{bias} - n^I_{1,l}[k]) \\ n^I_{2,l}[k+1] &= n^I_{2,l}[k] + \mu^{NA}_2(\alpha^{NA} A_{1,l}[k]_{bias} - n^I_{2,l}[k]) \end{aligned} \tag{14}$$

Finally, the control belief component is described similarly to the behavioural belief model, as such,

$$B^C_{i,l}[k]_{bias} = sat(B_i[k] + sgn(c_{i,l}[k])h^{BC}), \tag{15}$$

$$\begin{aligned} \hat{c}_{1,l}[k+1] &= c_{1,l}[k] + \mu^{CB}_1(\alpha^{CB} B^C_{1,l}[k]_{bias} - c_{1,l}[k]) \\ \hat{c}_{2,l}[k+1] &= c_{2,l}[k] + \mu^{CB}_2(\alpha^{CB} B^C_{2,l}[k]_{bias} - c_{2,l}[k]) \end{aligned} \tag{16}$$

$$\hat{c}_{j,l}[k+1]_{bias} = sat(\hat{c}_{j,l}[k+1] + sgn(\hat{c}_{i,l}[k+1])h^c), \tag{17}$$

$$\begin{aligned} c_{1,l}[k+1] &= \hat{c}_{1,l}[k+1] + \mu^C_1(\hat{c}_{2,l}[k+1]_{bias} - \hat{c}_{1,l}[k+1]) \\ c_{2,l}[k+1] &= \hat{c}_{2,l}[k+1] + \mu^C_2(\hat{c}_{1,l}[k+1]_{bias} - \hat{c}_{2,l}[k+1]) \end{aligned} \tag{18}$$

The metric equation for evaluating the performance of technique r is as follow,

$$Metric = \sum_i (b_i - b_{i,noinput}) \tag{19}$$

B. Summary ODD

The overall *purpose* of our model is to predict the behavioural dynamics of students in a classroom setting. Specifically, we are addressing the following questions: How does an instructors influence affect the change in student beliefs and behaviour. to consider our model realistic enough for its purpose, we use *patterns* of belief change. The model's ability to reproduce existing psychological phenomena will be used to validate the model's realism.

The model includes the following *entity*: students. The *state variables* characterising these entities are listed in Table 1. Students have a set of beliefs that that determine their likelihood of or effectiveness at performing a behaviour. The temporal resolution and extent are 1 day for each time step, with simulations run for a school year of approximately 200 days.

The most important *processes* of the model , which are repeated every *time step*, are the evolution of student beliefs relative to other students based on conformity effects as well as the incorporation of feedback from the outcomes of performing a given behaviour.

The most important *design concepts* of the model are the way students can incorporate outcome feedback into their belief updates, as well as how confirmation bias

affects their beliefs by encouraging like minded thought.utcome feedback into their belief updates, as well as how confirmation bias affects their beliefs by encouraging like minded thought.

Table 1 Important parameters and variables from the combined model

Symbol	Description
$B[k] \in [-1, 1]$	State representation of the behaviour of all individuals in the group, $B[k] = [B_1 B_2 \dots B_n]^T$ for n individuals in the group
$A[k] \in [-1, 1]$	State representation of the attitude of all individuals in the group, $A[k] = [A_1 A_2 \dots A_n]^T$
$N[k] \in [-1, 1]$	State representation of the perceived norm of all individuals in the group, $N[k] = [N_1 N_2 \dots N_n]^T$
$C[k] \in [-1, 1]$	State representation of the perceived behavioural control of all individuals in the group, $C[k] = [C_1 C_2 \dots C_n]^T$
$R[k] \in [-1, 1]$	State representation of the behavioural result of all individuals in the group, $R[k] = [R_1 R_2 \dots R_n]^T$, where R_i is itself a state containing all results relevant to individual i
$b[k] \in [-1, 1]$	State representation of the behavioural beliefs of all individuals in the group, $b[k] = [b_1 b_2 \dots b_n]^T$
$n^I[k] \in [-1, 1]$	State representation of the injunctive normative beliefs of all individuals in the group, $n^I[k] = [n_1^I n_2^I \dots n_n^I]^T$
$n^D[k] \in [-1, 1]$	State representation of the descriptive normative beliefs of all individuals in the group, $b[k] = [n_1^D n_2^D \dots n_n^D]^T$
$e_{i,l} \in [-1, 1]$	The associated value of a belief l held by an individual
$m_{i,l}^I, m_{i,l}^D \in [0, 1]$	The associated motivation to comply of a normative belief l held by an individual
$p_{i,l} \in [-1, 1]$	The associated power of a control belief l held by an individual
$\mu \in [0, 1]$	The rate of convergence for an associated socially influenced belief
$\omega_{i,j} \in [0, 1]$	The impact of an individual j on another individual i
$h \in [0, 1)$	The level of confirmation bias for an associated agent

References

1. Asch, S.E.: Studies of independence and conformity: I. A minority of one against a unanimous majority. Psychol. Monogr.: Gen. Appl. **70**(9), 1 (1956)
2. Bandura, A.: Social cognitive theory of mass communication. Media Psychol. **3**(3), 265–299 (2001)
3. Carlson, E.R.: Attitude change through modification of attitude structure. J. Abnorm. Soc. Psychol. **52**(2), 256 (1956)
4. Cialdini, R.B., Goldstein, N.J.: Social influence: compliance and conformity. Annu. Rev. Psychol. **55**, 591–621 (2004)
5. Feather, N.T.: Expectations and Actions: Expectancy-Value Models in Psychology. Routledge (1982)
6. Festinger, L., Schachter, S., Back, K.: Social Pressures in Informal Groups; A Study of Human Factors in Housing. Harper (1950)
7. Fishbein, M.: A consideration of beliefs and their role in attitude measurement. In: Readings in Attitude Theory and Measurement, pp. 257–266 (1967)
8. Fishbein, M., Ajzen, I.: Predicting and Changing Behavior: The Reasoned Action Approach. Taylor & Francis (2011)
9. Flache, A., Mäs, M., Feliciani, T., Chattoe-Brown, E., Deffuant, G., Huet, S., Lorenz, J.: Models of social influence: Towards the next frontiers. J. Artif. Soc. Soc. Simul. **20**(4) (2017)
10. Friedman, M.C.: Notes on Note-Taking: Review of Research and Insights for Students and Instructors, pp. 1–34. Harvard Initiative for Learning and Teaching (2014)
11. Knauder, H., Koschmieder, C.: Individualized student support in primary school teaching: A review of influencing factors using the theory of planned behavior (TPB). Teach. Teach. Educ. **77**, 66–76 (2019)
12. Lobdell, M.: Study Less, Study Smart: A Guide to Effective Study Techniques and Enhanced Learning. Createspace Independent Publishing Platform (2015)
13. Myers, D.G.: Polarizing effects of social interaction. In: Group Decision Making, pp. 125–161 (1982)
14. Myers, D.G., Lamm, H.: The group polarization phenomenon. Psychol. Bull. **83**(4), 602 (1976)
15. Nickerson, R.S.: Confirmation bias: a ubiquitous phenomenon in many guises. Rev. Gen. Psychol.Rev. Gen. Psychol. **2**(2), 175–220 (1998)
16. Rashotte, L.: Social influence. In: The Blackwell Encyclopedia of Sociology (2007)
17. Stanovich, K.E., West, R.F., Toplak, M.E.: Myside bias, rational thinking, and intelligence. Curr. Dir. Psychol. Sci. **22**(4), 259–264 (2013)
18. Vinokur, A., Burstein, E.: Effects of partially shared persuasive arguments on group-induced shifts: a group-problem-solving approach. J. Pers. Soc. Psychol. **29**(3), 305 (1974)
19. Wood, W.: Attitude change: Persuasion and social influence. Annu. Rev. Psychol. **51** (2000)

The Role of Communication and Network Technologies in the Dynamics of Social Movements

Krystyna Marcinek, Rushil Zutshi, Omair Khan, Justin Grana, Marek Posard, Todd Helmus, and Aaron Frank

Abstract We investigate the multi-faceted role of information technologies in the spread and dynamics of social movements. Specifically, we ask two main questions: (1) how do communication *and* network technologies impact the number and connectivity of movement participants, and (2) how does more efficient and more accurate surveillance technology impact an authority's ability to learn about the movement. Importantly, our simulation model includes both homophily and social influence, two established tenants of social movements and social relationships more broadly. Our results show that communication technology that increases spontaneous interaction helps to ignite social movements, while improvements in networking technology are more effective at accelerating the growth of social movements in their intermediate stages. However, when aents are allowed to join the movement, outreach is more

Reproduced from "The Role of Communication and Network Technologies in the Dynamics of Social Movements" by Krystyna Marcinek, Rushil Zutshi, Omair Khan, Justin Grana, Marek N. Posard, Todd C. Helmus, Aaron B. Frank with permission from the RAND Corporation.
This work was completed when still affiliated with the RAND Corporation.

K. Marcinek · J. Grana · M. Posard · T. Helmus · A. Frank
RAND Corporation, 1200 South Hayes Street, Arlington, VA 22202, USA
e-mail: marcinek@prgs.edu

J. Grana
e-mail: justin.grana@microsoft.com

M. Posard
e-mail: mposard@rand.org

T. Helmus
e-mail: helmus@rand.org

A. Frank
e-mail: afrank@rand.org

K. Marcinek · R. Zutshi (✉) · O. Khan · T. Helmus
RAND Corporation, 1776 Main Street, Santa Monica, CA 90401, USA
e-mail: rzutshi@prgs.edu

O. Khan
e-mail: okhan@prgs.edu

Z. Yang and S. Núñez-Corrales (eds.), *Proceedings of the 2022 Conference of The Computational Social Science Society of the Americas*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-031-37553-8_13

effective at accelerating the growth of the number of participants. Our results also show that authority can gain highly accurate beliefs simply by observing network links (instead of individual actors) in all but the smallest social movements.

The ability to communicate with peers, both proximate and distant, is a central component of social movements. Historically, American slaves invented codes in the form of songs [1] and quilts [2] to facilitate communication and escape along the underground railroad. The advent of coffee shops in 17 and 18th century England is often acknowledged as an expansion of communication channels and a key contributor to the Age of Enlightenment [3]. Online communication platforms allowed Hong Kong protesters "to form small groups to initiate and coordinate their actions [4]." And in early 2021, political protesters used social media and other internet forums to share information, which ultimately resulted in violent protests at the U.S. Capitol.

As protesters develop and leverage advances in communication technologies, the empowered authority also leverages advances in data collection, storage, and processing technologies to better monitor and subdue the movement. For example, in the 2011 Syrian Civil Uprising, the Syrian government engaged in "monitoring and controlling a user's dynamic web-based activities," and implemented systems that were "capable of capturing webcam activity, logging keystrokes, [and] stealing passwords [5]." This suggests that both protesters and authority figures are leveraging cutting-edge technologies to achieve their objectives, but it is still unclear how this technological co-evolution impacts the overall dynamics of social movements.

This work develops and analyzes an agent-based model (ABM) of network formation to investigate such dynamics. Specifically, we first investigate how an increased ability to form communication channels impacts the overall growth (both nodes and links) of social movements. Then, we investigate how improved surveillance technologies—motivated by technological advances in data collection, storage, and processing power—impact an authority's ability to monitor and potentially disrupt a social movement. Ultimately, we characterize the interaction between the participants and the authority in the context of their technological endowments.

In order to remain as general as possible, the model examines how changes in communication technologies impact the number of participants and the connectivity among them without any reference to the ultimate goal of the social movement. Alternatively put, we abstract away from how—once connected—participants mobilize and coordinate resources in pursuit of an objective. Instead, we focus on the preliminary step of how the participants coalesce in the first place. The main benefit of this approach is that the model is agnostic to the participants' goals and thus provides general insights into *any* social movement. Of course, the main drawback is that the mechanisms through which a group achieves its goal are relevant for understanding (and predicting) any specific social movement. However, by establishing a framework that explains how social movements grow and connect, the model can serve as a building block onto larger models that consider both communication and mobilization. We also recognize that the surveilling body might not always be a government or law enforcement agency but may include private corporations or special interest

groups. However, we use the term "authority" loosely to mean an agent with the desire and means to monitor a social movement

We follow the same logic for the authority. That is, we abstract away from any particular goal of the authority and focus on the authority's need to decide which of the agents are participants. The value of this approach rests on the assumption that, except in the most extreme cases, regardless of the authority's goal, acquiring information about who is a participant is a necessary pre-requisite for action. For example, if an authority wanted to disrupt a social movement by severing communication links between participant factions, it would need to know which of the citizens are participants that serve as links between disparate groups. Again, while this means that our model does not apply to *particular* social movements, it serves as a framework for building models that include the authority's disruptive action. We also recognize that the surveilling body might not always be a government or law enforcement agency but may include private corporations or special interest groups. However, we use the term "authority" loosely to mean an agent with the desire and means to monitor a social movement.

In order to ensure that the results are rooted in established principles, we include both homophily and social influence in the model. Homophily is the notion that similar individuals are more likely to have contact than two individuals that are different [6]. Such phenomena are well documented in domains such as adolescent friendship formation [7], entrepreneurial relationships [8], and, perhaps unsurprisingly, political opinion networks [9]. Social influence is the notion that an individual's opinions and beliefs are shaped by their peers [10]. Like homophily, social influence has been observed in a variety of contexts, including exercise habits [11] and product purchase decisions [12]. Both homophily and social influence have been observed in social movements [13–15].

Our model yields several interesting results. It illustrates how a change in communication technologies that increase participants' ability to contact agents outside of their immediate network (a term we call *outreach*) is most effective in igniting a social movement, both in terms of increasing the number of its supporters, as well as the accelerating connectivity among them. On the other hand, a change in networking technologies that increases participants' ability to contact agents within close network proximity (i.e., friends of friends) is most effective at accelerating the growth of the social movement through its intermediate stages. The final number of social movement participants in our model is determined by the underlying characteristics of the population—social influence, homophily, and the initial seed of the movement. However, for some sets of population characteristics, the final number of participants is very difficult to predict—it is possible that no new agent becomes a participant (that is, the social movement does not grow beyond the initial seed), as well as that all of them do, turning the entire population into supporters of the social movement.

On the surveillance side, we find that the authority figure is able to form accurate beliefs about the participation status of agents even when the node observations may be noisy so long as it can reliably observe the links between agents. Expectedly, we also find that the accuracy of these beliefs improves when agents are sampled more frequently. When we then compare improvements in the accuracy of these beliefs

across these three parameters over different kinds of social movements (based on size and connectivity), we find that the accuracy of the authority's beliefs is most improved when the node observation process is less noisy, but this is less the case as the social movement becomes larger and more connected. Most of the same results hold even with social influence, but the dynamic nature of the social movement size results in less monotonic relationships between the accuracy of the authority's beliefs and the surveillance parameters.

Communication, Social Movements, and Agent-Based Models

Our work builds on—and incorporates—insights from several traditionally disparate disciplines. As mentioned above, our model incorporates homophily and social influence, two well-established tenants in the sociology of network formation. We quantitatively define[1] degrees of homophily and social influence and contribute to the literature by evaluating how changes in such degrees impact the dynamics of a social movement.

Our work is also related to the vast literature on behavior spread [16] and dynamic networks [17]. Using both analytical (examples can be found in [18, 19]) and computational (examples can be found in [20, 21]) methods, the goal is to derive properties of networks that form according to some underlying stochastic process. Our work is related to this line of research in that we also propose an underlying stochastic process for the evolution of the network. However, our work is different in several ways. First, we are interested in the state of the network as a function of time and nodes. This contrasts with the dynamic network literature that typically (though not always) focuses on the limits as time and the number of nodes gets large. Furthermore, we incorporate an authority figure that does not participate in the network but *monitors* the network. This is an element that is absent from the study of dynamic networks.

Our work is most related to other ABMs of communication and social movements. This includes models of political insurgency [22], revolution and technology [23], civil violence [24–26], radicalization [27] and more [28]. We also build on game theoretic models of mass action [29]. Our model adds to this literature in several ways.

First, our model trades intention for generality. That is, we focus specifically on how well participants can form a network and how well an authority can monitor the network without any regard for either party's objectives. This means that our model is not constrained to a specific social movement but can be used to explain network formation more generally. Of course, the trade-off is that our work stops short of expressing how changes in communication technologies affect how well the

[1] We do not claim that our quantitative definition of homophily and social influence is the unique definition, we only claim that it is reasonable in the context of social movements.

participants achieve their goals, but our model is meant to be modular and can be included in models that specifically incorporate participants' goals.

Secondly, we include technological change from both the participants' and the authority's perspectives. Many recent ABMs examine the impact of more accessible social media on the evolution of a social movement [30]. However, these models typically ignore that as participants are exposed to more technologically sophisticated tools, the authority's technological endowments are also increasing. So, for example, while participants are more able to connect, the authority's increased processing power and algorithmic advances may also increase its ability to monitor a social movement and eventually disrupt it. Our model explicitly takes these dynamics into account.

Third, our model is not only concerned with "steady-state" distributions but instead asks "how fast" do the participant's network and the authority's knowledge evolve? This contrasts with many ABMs of social movements that focus on the distribution of end-states [31]. That is not to say that our model is better or more relevant than those focusing on end-states. Instead, it only illuminates a different aspect of complex social movements.

Model Overview

In this section, we present a narrative overview of the model. The model overview suffices for understanding the key components of the model and the results. The full technical specification—presented in the Appendix—provides the necessary detail to replicate the model. The model overview followed by the the full technical specification adhere to the ODD Protocol.

Purpose

The model seeks to understand how changes in communication technologies (the increased ability for outreach and networking) impact the growth of the number of participants and the connectivity between them. The first insight tells us how different communication technologies can influence whether a social movement is a slow boil, explosive, or some combination of the two. Connectivity is important under the assumption that regardless of the participants' goals, having more communication channels facilitates the mobilization of resources towards the participants' ultimate goal. The model also focuses on how the authority collects information about the social movement. The key question we address is how changes in surveillance and data processing technologies impact how an authority determines which citizens are participants.

Entities

Our model is a dynamic model of network formation with both *citizens* and an *authority*. A fixed number of citizens are deemed social movement *participants* while the rest of the citizens are non-participants. Each citizen, either participant or non-participant, has an additional attribute called its *identity*. The identity is independent of whether the citizen is a social movement participant and represents characteristics of the citizen other than its proclivity to join the movement. For example, it might capture demographic information, tastes, interests, or education levels. At the start of the model, all citizens are disconnected, and there are no communication channels between any two citizens.

Process overview

Participant Link Formation and Evolution

In each time step, an agent is chosen at random to potentially form links with other agents under two mechanisms, which we label *outreach* and *networking*. Outreach is the process where the agent can survey a fixed number of random agents. If both agents are participants or both agents are similar enough in identity, they form a link. This process represents a simple notion of homophily and can be interpreted as serendipitous networking such as meeting a person at a public park or cafe. As a digital example, an agent might 'stumble' upon another agent's blog while on the internet and subsequently connect to that person through a social media platform.

Networking refers to the process where the agent can survey a fixed number of agents who have connections with its current connections. For example, if agent Alex is connected to agent Bob and agent Bob is connected to agent Chris, then Alex can survey Chris because it is connected to Chris through Bob. Again, homophily dictates that agents under this process form a link if they are either both participants or if their identities are sufficiently similar. A digital example of the networking mechanism would be two agents that connect via a closed online group they were both invited to by a common friend.

The parameters that govern how many agents a given agent can survey in one time step can be interpreted as the inverse of the communication costs to forming links. For example, if an agent has a fixed amount of time it can dedicate to communicating and meeting new people, raising the number of queries an agent can make to other agents is equivalent to lowering the communication cost. A key feature of our model is that there are separate parameters that govern outreach and networking, and thus it is possible to isolate the impacts of advances in networking and advances in outreach technology. In real-world communication systems, advances in social media platforms allow providers to suggest connections based on friends of friends and thus reduce the cost of networking. On the other hand, websites centered around filling job vacancies facilitate communication between parties that are not otherwise

linked and thus reduce outreach costs. Our model is able to disentangle the impacts of these similar but subtly different notions of advances in communication technologies. The results ultimately show that the relative impact of increasing networking versus increasing outreach depends on how much the movement has already progressed and the initial prevalence of the social movement.

Finally, we allow agents that were not originally participants to become participants if they are connected to enough participants. There is a parameter that specifies the agent's propensity to become a participant, and that parameter captures the degree of social influence in the model. While the parameter is the same for every agent, we run several experiments—including a base case with no social influence—to explore the impacts of various degrees of social influence.

Authority Observation Process

To form beliefs about which agents are participants, the authority observes a random subset of agents at regular time intervals. The number of agents per time step (which can be less than 1) represents the processing power and storage capacity available to the authority. Specifically, the interpretation is that as the authority's processing power and data storage capacities increase, it is able to collect and analyze data on a wider set of individuals in a given time period. The authority's observation of a given agent consists of two components. First, there is a *node observation process* where the authority observes a noisy signal of whether a given agent is a participant. The amount of noise captures the quality of the authority's surveillance technology, where a low-noise signal represents a more sophisticated data collection and analysis technology.

Second, for each agent the authority observes, there is an *edge observation process* in which the authority observes a subset of the agent's personal network. Specifically, the authority has a fixed probability of observing each of the agent's existing edges. This probability represents the quality of the authority's technology allowing network structure surveillance. A low probability may mean that the authority can observe connections only under very strict conditions, for example, when two agents are physically co-located. A high probability could, for example, mean that the authority can observe connections despite the mode of communication chosen by the participants, such as social network data mining. Notably, the node and edge observation technologies are governed by separate parameters. This allows us to disentangle the impacts of increased individual surveillance versus increased network surveillance, our key comparative statics when analyzing the authority.

A fully rational authority would have a prior belief and use the data from its node observation process and its edge observation process to compute a Bayesian probability of each agent being a participant. However, this exact probability is analytically intractable due to the underlying model of network formation and would require significant Monte Carlo experiments to estimate. This is exacerbated in models of social influence where agents can switch from being non-participants to participants. Since it is unlikely that a real-world authority would be able to compute these

probabilities accurately, our implementation of an authority adopts a heuristic that, with infinite samples, can separate participants from non-participants. Intuitively, the heuristic uses the fact that, on average, participants have more connections than non-participants (because participants are connected to one another as well as with those with a similar identity) and thus combines the degree of a node with the node observation data to assign a numerical value (not a proper probability) to a node being a participant. With infinite samples, the score of all participants would be strictly positive, while the score of all non-participants would approach 0. We show how well this classification process works using a receiver operator curve (ROC) in the results section of the paper.

Results

This section details the main results. We begin with the simplest scenario and examine the dynamics of the participant network formation without social influence. In these initial simulations, we focus on the *connectivity* among participants since the number of participants is fixed. We then introduce the authority and show how the accuracy of its beliefs depends on the parameters governing its observation process. We then repeat the exercise with social influence. Since the number of participants changes over time, these results focus on the evolution of the *number* of participants. Finally, we examine how the social movement growth influences the authority's ability to track participants accurately.

Table 1 presents the parameters we vary in the simulations. To ease presentation, in the subsequent analysis, we refer to the parameter values as "low", "medium", and "high" instead of their precise numerical values as specified in the Appendix.

Table 1 Parameters

Characteristics	Parameter
Underlying characteristics of the population	Initial seed of movement p
	Inclusivity c
	Social influence w
Technological capabilities of citizens	Outreach technology L_1
	Networking technology L_2
Surveillance capacity of the authority	Sampling frequency k
Surveillance accuracy of the authority	Noise in node observation δ
	Probability of link observation γ

Participant Dynamics without Social Influence

To understand how technology impacts participant connectivity, we show how the degree of participants evolves over time when there is no social influence, and thus the number of participants is fixed (i.e., $w = \infty$). Figure 1 shows the average degree of a participant and a non-participant for 40 simulations at the same parameter values. The key insight is that the average degree of participants (1) converges, (2) is stochastic and (3) is higher than non-participants. The fact that the quantity is convergent and higher for participants is not surprising. With infinite samples and without social influence, each participant will be connected to every other participant *and* every non-participant with a similar identity, while non-participants connect exclusively on the identity. The stochastic nature is also not surprising. Since each agent's identity is drawn from a uniform distribution at the start of the simulation, the number of agents another participant connects to based on identity is random.

Due to the stochastic nature of the participant's degree, we use Monte Carlo simulations to compute the average degree when holding the parameters fixed. Therefore, when we later use the term "average degree," we refer to the average over *all participants* and *40 Monte Carlo simulations* for a fixed set of parameters.

Another baseline result is that the final average degree of a participant depends on two key parameters, the initial number of participants, p and inclusivity, c. Figure 2 plots the average participant degree for each of the different parameter sets at the end of the simulation. For each parameter value, the results group around three values corresponding to three levels of the other parameter. For example, results for the parameter sets with low inclusivity—depicted in the most left stack—group around 13, 18, 22.5 for the sets with low, medium, and high initial seed, respectively. The plots show that the parameter values are such that increasing the initial number of participants by one category (low to medium, for example) has roughly the same effect as increasing inclusivity by the same category (also low to medium, for example). This is important because it establishes that c and p are on a similar scale, and differences in scale are not driving the results.

In this baseline model, we were particularly interested in the speed of social movement network formation, i.e., we only investigate connections between participants. To understand the relative impacts of increasing network and communication technologies on participant connectivity, we investigate how much faster the participants' network converges to its limits as we increase one technology endowment versus another.

This result is presented in Fig. 3. The horizontal axis represents the proportion of links among participants relative to the network's limit (all participants connected to one another). The vertical axis represents how much faster that level of convergence is reached due to an increase in outreach versus an increase in networking. So, for example, in the left plot, when p is low and c in medium (blue annotated line), increasing outreach (L_1) from low to medium means that the proportion of connections reaches 50% almost 500 time steps sooner than if networking (L_2) were to increase from low to medium. These graphs provide two high-level insights.

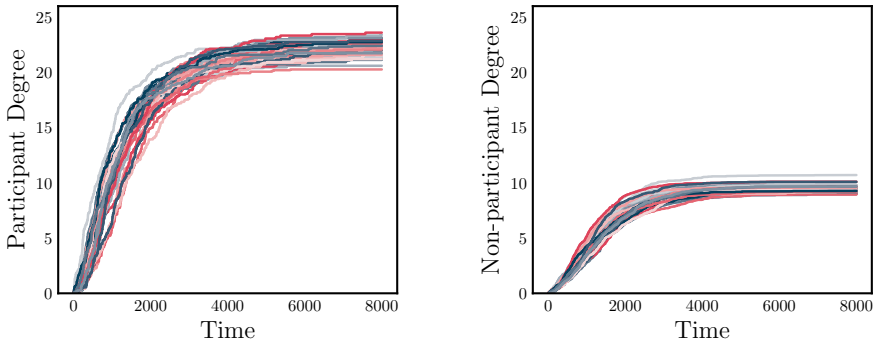


Fig. 1 Convergence of average participant and non-participant degree for fixed parameters ($L_1 = L_2$: low, p: high, c: low)

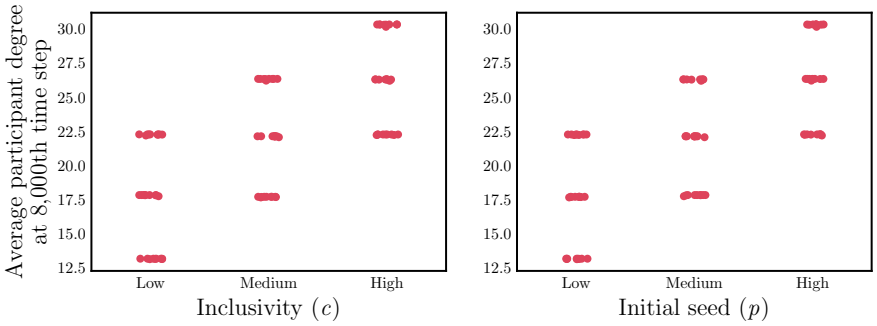


Fig. 2 Distribution of average participants degree at the limit, grouped by c and p

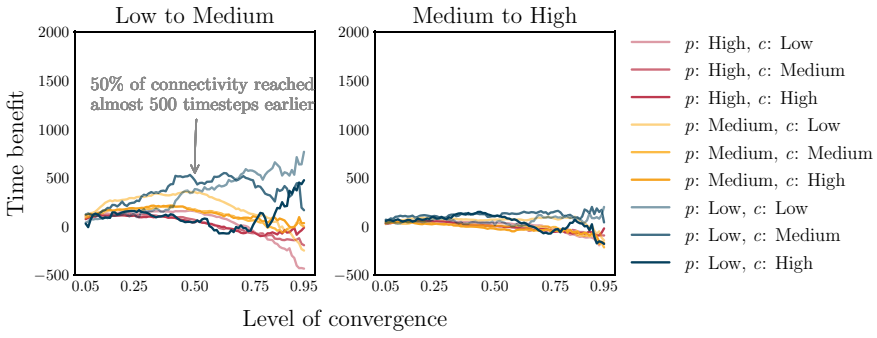


Fig. 3 Time benefits of increasing outreach versus networking (as a function of the level of convergence)

First, the relative benefits of outreach and networking depend on the initial size of the social movement. With a small number of participants, increasing outreach leads to faster link formation for any given level of connectivity. However, with a larger number of participants, increasing networking can lead to faster convergence to the ultimate network state. Intuitively, if there are very few participants, most of the connections are formed based on identity. Therefore, it is unlikely that a participant would find another participant through networking. However, when there are more participants in the population, more links are initially formed among participants, and the higher the chance of meeting other participants via one's existing connections. In total, this means that in all but the smallest movement size, increasing outreach is more apt to "ignite" connectivity among participants, whereas increasing networking is more effective at quickly bringing the participant network to its ultimate potential.

Secondly, the plots show that the relative impact of increasing outreach versus networking is nonlinear, which is evident from the difference in the height of the lines in the left and right plots. Specifically, the relative difference in convergence time from increasing outreach versus networking from low to medium is much more than an increase from medium to high. This suggests that only in technologically limited cases do increases in networking and outreach technologies have asymmetric impacts on network formation. However, as both technologies reach an advanced state, their impact on participant connectivity becomes indistinguishable.

Surveillance without Social Influence

This section investigates how improvements in different surveillance technologies impact an authority's ability to discern participants from non-participants. Since each participant receives a score from the authority, a key quantity is what participant score cutoff the authority uses to classify citizens as participants. This score would ultimately depend on the authority's cost and benefits of correct and incorrect classifications. In order to gauge the efficacy of the authority's classification process in the general sense without reference to the authority's payoff, we use a Receiver Operator Curve (ROC) to understand the trade-off between the true positives (benefits) and false positives (costs).

An example of the ROC graph from our analysis is shown in Fig. 4. The black line represents a ROC at the beginning of a simulation ($t = 50$)—it almost overlaps with the diagonal (gray dashed line), representing a random guess. As the simulation progresses, the curves shift towards the top left corner, which means that the authority's ability to classify and detect participants improves.

Figure 5 shows how the surveillance ability of the authority changes over time (here for 10 Monte Carlo simulations). Specifically, the figure plots the area under the curve (AUC) at various time steps. Just like in the participant case, the AUC at any timestep is stochastic and thus necessitates Monte Carlo simulations. Therefore, when we use the term "mean AUC" we refer to the average over *40 Monte Carlo simulations* for a fixed set of parameters.

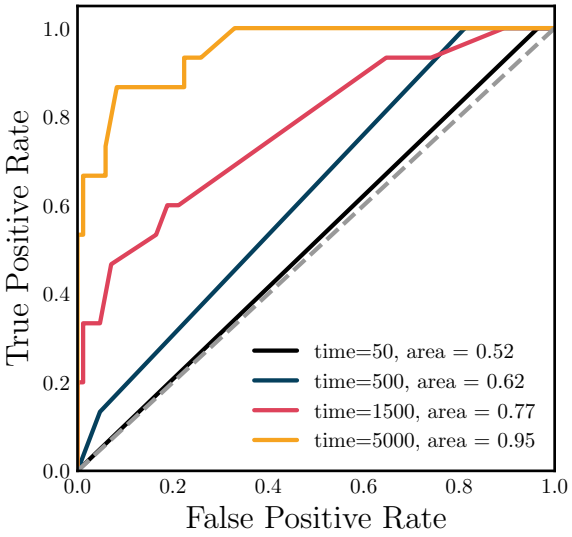


Fig. 4 Receiver Operating Characteristic (ROC) curves at various time steps for fixed parameters ($L_1 = L_2$: low, p: high, c: low, k: low, δ: medium, γ: low)

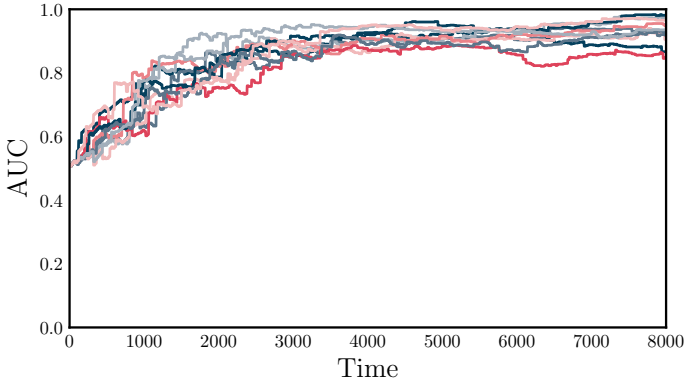


Fig. 5 AUC at various time steps for fixed parameters ($L_1 = L_2$: low, p: high, c: low, k: low, δ: medium, γ: low) in 10 Monte Carlo simulations

Noise in Node Observation (δ) and Probability of Link Observation (γ)

Figure 6 presents results of the AUC over time for different configurations of the participant parameters and link and node observation parameters while sampling frequency is held constant at 1 agent observation per 10 time steps. The figure illustrates that even when the authority's node observation process is very noisy, it still performs relatively well as long as it observes links with a high probability. This is best illustrated in the right column where γ is high. In all but the first row, the red

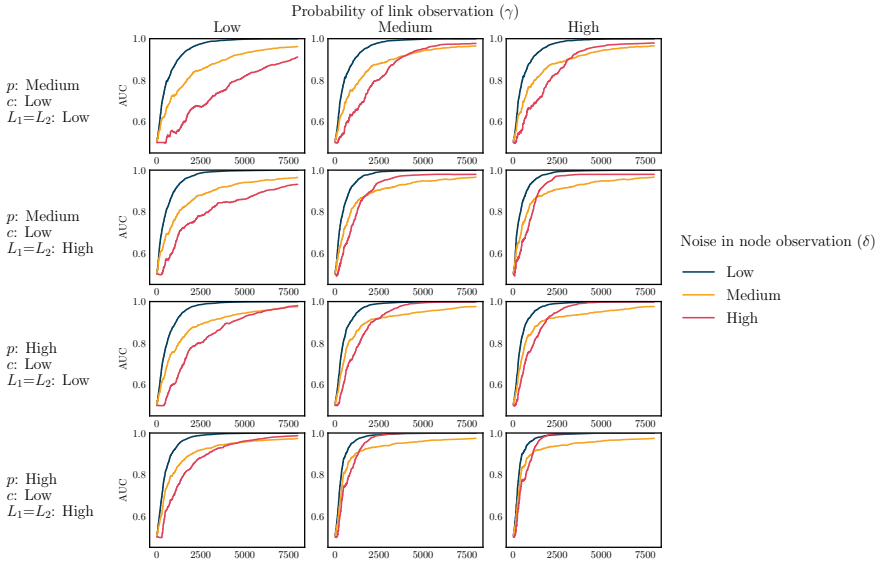


Fig. 6 AUC at various time steps for δ (low, medium, high) over different γ and 40 Monte Carlo simulation (k: medium)

line—representing high noise in the node observation—is relatively close to the blue line—representing low noise in the node observation process. However, this is *not* the case with a low number of initial participants and low communication technologies, which is the scenario in the first row of plots in Fig. 6. In this case, participants connect so slowly that network information provides very little information to the authority, and the authority's beliefs are driven primarily by its observation of nodes. Thus when the node observation process is noisy, the authority cannot "fall back" on the edge observation process to form its beliefs.

Another curious result arising from these plots is that the authority may sometimes have a higher AUC with *higher* noise in the node observation process, which is the case when the red line is higher than the yellow line. This could be caused by the node observation process being effectively ignored when it is too noisy. While this is undoubtedly an artifact of the heuristic score function, it shows the importance of correctly aggregating participant data. Of course, this is true from a modeling perspective but is also true in practice. Data come from disparate sources during real protests, and with constraints on data sharing, modeling, and processing, authorities are often required to use heuristics. This result shows how seemingly benign and innocuous heuristics can lead to counter-intuitive results and reinforces the importance of an authority taking a principled approach to data analysis.

Sampling Frequency (k)

We can compare how changes in the sampling frequency k impact the authority's beliefs at different probabilities of link observation and participant parameters. These results are summarized succinctly in Fig. 7. This depicts an expected trend of a quicker convergence to AUC = 1 when the sampling frequency increases from sampling a node every 20 time steps (low sampling frequency) to every 5 time steps (high sampling frequency).

Finally, we can compare how changes in the sampling frequency k impact the authority's beliefs relative to changes in the noise in its node and link observation process. These results are summarized succinctly in Fig. 8. This figure shows gains to AUC over time (averaged for each parameter over all other parameter combinations) for changes in each of the three authority parameters. The plot shows that the impact of changing the sampling frequency is independent of the underlying participant dynamics. This is demonstrated by the relatively small changes in the height of the lines in the first column of the figure. This is in contrast to noise in the link observation process, where improvements have the largest impact in large social movements with fast connectivity (middle column plots are highest in the bottom row) and noise in the node observation process, where improvements have the largest impact in small social movements with slow connectivity (third column plots are highest in the top row).

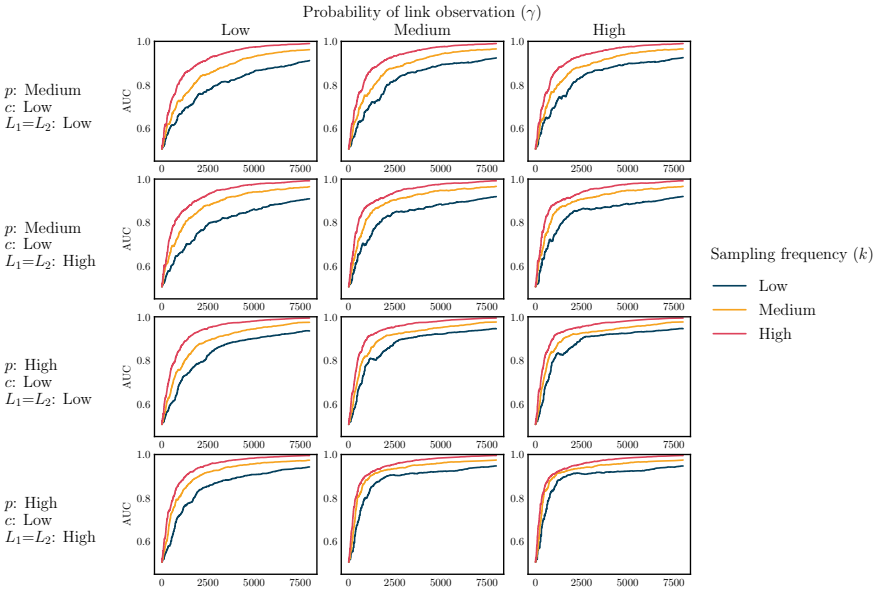


Fig. 7 AUC at various time steps for k (low, medium, high) over different γ and 40 Monte Carlo simulation (δ: medium)

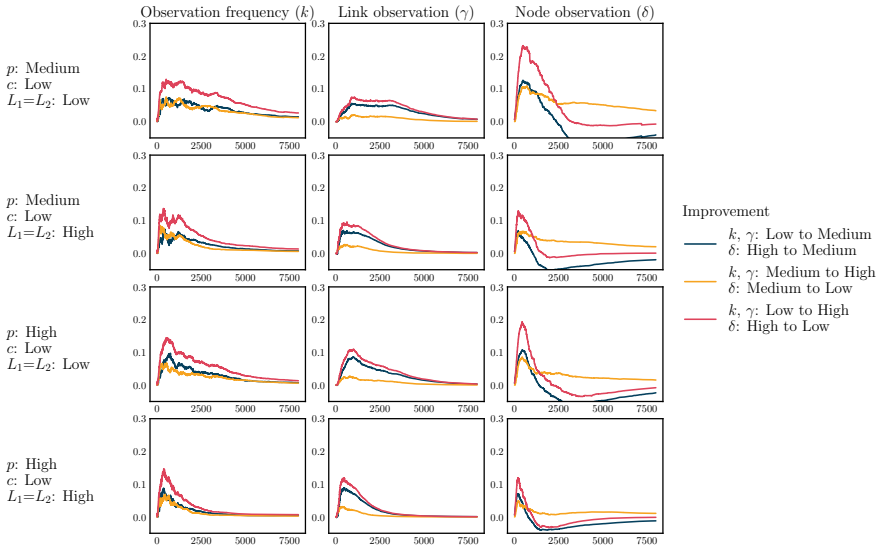


Fig. 8 Average AUC improvement over time for improvements in k, γ, and δ

Participant Dynamics with Social Influence

In this section, we re-investigate the participant dynamics, allowing non-participants to become participants via social influence. Specifically, if a non-participant is connected to at least w participants, it becomes a participant. Importantly, social influence is a one direction mechanism—participants cannot become non-participants. In the previous sections without social influence, the number of participants was fixed, so we focused on the connectivity among participants. With social influence, the number of participants varies over time, and therefore, we focus our analysis on the size of the social movement.

Once again, the number of participants is stochastic and convergent. This is illustrated in Fig. 9, which shows the change in the number of participants over time for 40 simulations at the same parameter values. In this particular case, all 40 simulations started with 10 participants; however, the final number of participants varies between 40 and 100.

Similar to the term "average degree", the term "average number of participants" refers to the number of participants averaged over 40 Monte Carlo simulations.

Figure 10 shows the distribution of the average number of participants at the 8,000th time step grouped by the inclusivity and the initial number of participants parameters. For example, the most left stack presents results for the parameter sets with low inclusivity, where colors signify sets with different levels of social influence; dots in each color cluster around three different values corresponding to three levels of the initial number of participants. The plots show that even with a high degree of social

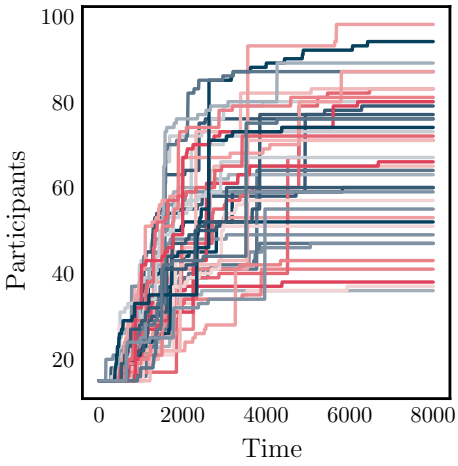


Fig. 9 Convergence of the number of participant for fixed parameters ($L_1 = L_2$: low, p: high, $c$low, w: high)

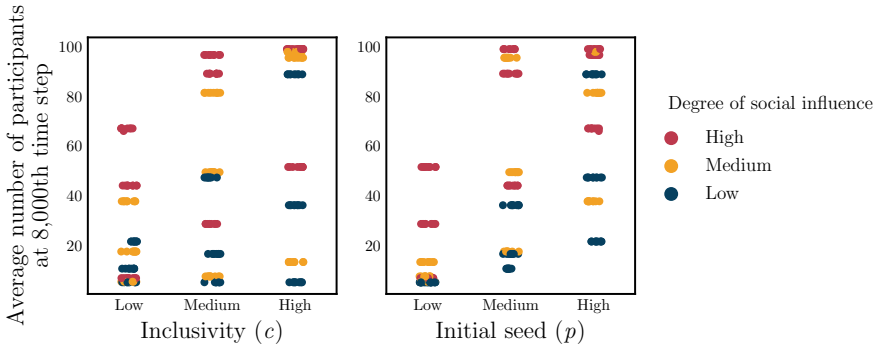


Fig. 10 Distribution of the average number of participants at the limit, grouped by c and p

influence (red dots), it is not guaranteed that everyone will become a participant. For instance, when both the inclusivity and the initial number of participants are low, the final number of participants remains low, which is represented on the plot by the red dots clustered around 5 in the most left stack. This is because even if social influence is high if c and p are low, non-participants do not have the opportunity to connect to participants and subsequently become participants themselves.

In addition to the number of participants, another relevant quantity is the variance in the number of participants, as this captures the uncertainty associated with how large a social movement might ultimately get. To gain insight into this feature, Fig. 11 presents the standard deviation of the number of participants at the limit for three different levels of social influence as a function of the inclusivity and the initial number of participants. As we can see at each level of social influence in this pop-

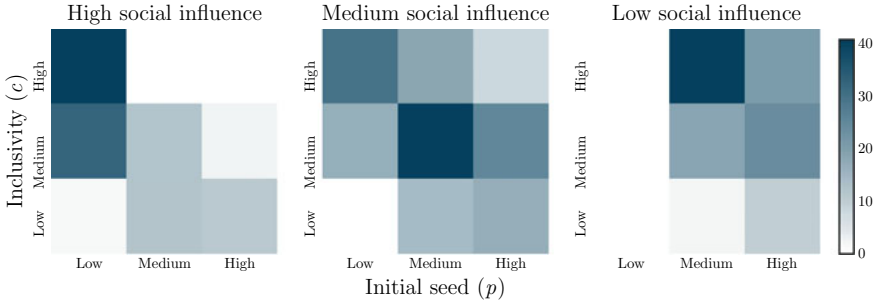


Fig. 11 Standard deviation of the number of participants at the limit

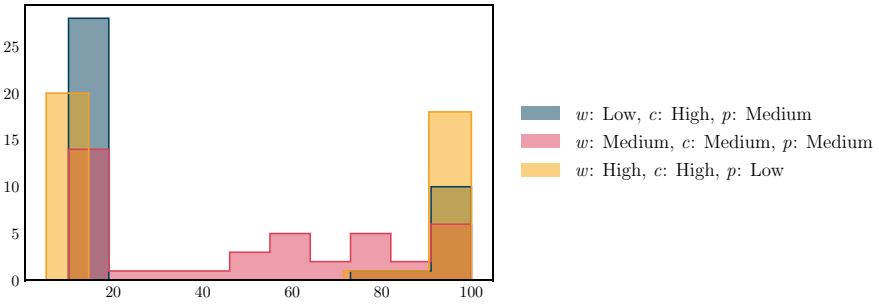


Fig. 12 Distribution of participants number at the limit for the highest variance parameter sets

ulation of 100 agents, the standard deviation of the ultimate number of participants ranges from 0 to 40.

The plot shows that the standard deviation is smallest either when almost every agent becomes a participant or when almost none does. The former is possible with high social influence in a population of high inclusivity and considerable size of initial movement seed (left graph, top row). The latter happens when both social influence and initial movement size are low (right graph, left column). High uncertainty is in turn associated with an overlap of these conditions. For example, in a society that encourages connections (high threshold of c), even a small number of participants can spark a social movement growth since the society is well-connected. We can see this effect in the left graph with high social influence, where we observe the highest standard deviation when the inclusivity is high, but the initial number of participants is low.

Interestingly, the cases with the highest variance are, in fact, cases of multimodal distribution, where either no one becomes a participant or everyone. This result is presented in Fig. 12. It is especially true for the highest variance parameter sets with high and low social influence (blue and yellow histograms). With medium influence (red histogram), the steady-state number of participants can be anywhere between the initial number of participants (10) and the maximum number of participants (100). This speaks to the difficulty of predicting the pace of propagation of a social

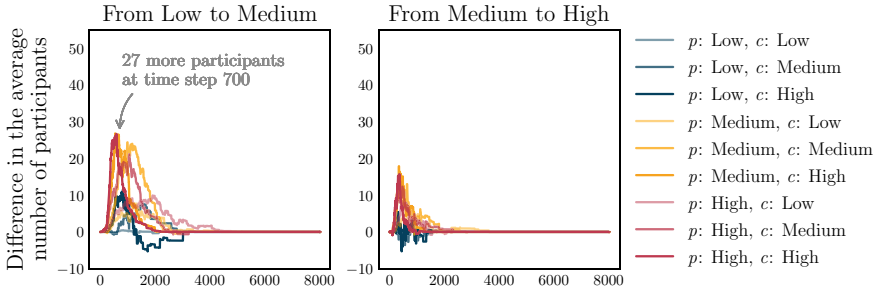


Fig. 13 The difference in the average number of participants per timestep resulting from increasing outreach versus networking (high social influence)

movement. The same characteristics of a movement at the outset can lead to vastly different popularity over time.

Another interesting question is how fast the social movement grows as a function of communication and network technologies. As increasing either networking or outreach technologies will unambiguously speed up the growth of the movement, the key question is how much does an increase in networking technology speed up the growth of the movement relative to an increase in outreach. Figure 13 demonstrates this by showing at each time step how many more participants there are (on average) by increasing outreach instead of networking when the degree of social influence is high. For example, the red annotated line in the left plot of Fig. 13 says that when p and c are high, increasing outreach ability from low to medium will result in an average of 27 more participants at time step 700 than if networking were to increase from low to medium.

Since most of the lines are weakly greater than 0, the plot shows that increasing outreach appears to almost always lead to a higher number of participants at any given timestep. That pace of movement growth depends on the underlying characteristics of the population—the higher the inclusivity, the higher the pace. Increasing networking leads to faster growth in the number of participants only when the initial number of participants is low (dark blue line on the left graph). This is due to the fact that when the number of participants is low, it is very unlikely to randomly sample a non-participant that is already connected to w participants. The advantage of increasing outreach instead of networking in all other cases indicates that a social movement would likely grow faster in societies endowed with better outreach technologies. The networking, on the other hand, becomes important later in the process, when members of the movement connect, and thus, the movement recognizes its strength, as the model without social influence indicated.

Surveillance with Social Influence

In this section, we return to the analysis of authority's surveillance capabilities allowing for social influence. As described in the model section in the appendix, the computation of the authority's belief of citizen i is based on all previous observations of the citizen node, even though now each citizen's participant status can change over time. While results are similar to the case of surveillance without social influence, there are some important distinctions highlighted below.

Noise in Node Observation (δ) and Probability of Link Observation (γ)

Figure 14 presents results of the AUC over time for different configurations of the participant parameters and probability of link observation parameter while sampling frequency is held constant at the medium level. As was the case when there was no social influence, even when the authority's node observation process is very noisy, it still performs well as long as the link is observed with high probability. In fact, in most cases, it performs better than the configuration with the least noise in the node observation process. Moreover, an important distinction is that in this case, this result holds even with a relatively low probability link observation process. This is likely because the noise in the node observation process is exacerbated when agents can turn into participants. For similar reasons as those outlined in the case without

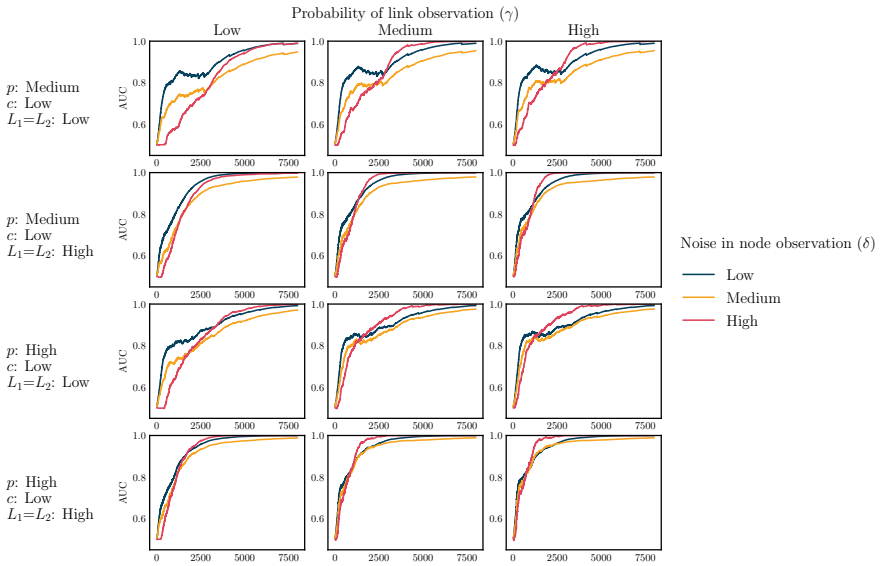


Fig. 14 AUC at various time steps for δ (low, medium, high) over different γ and 40 Monte Carlo simulations (k: medium)

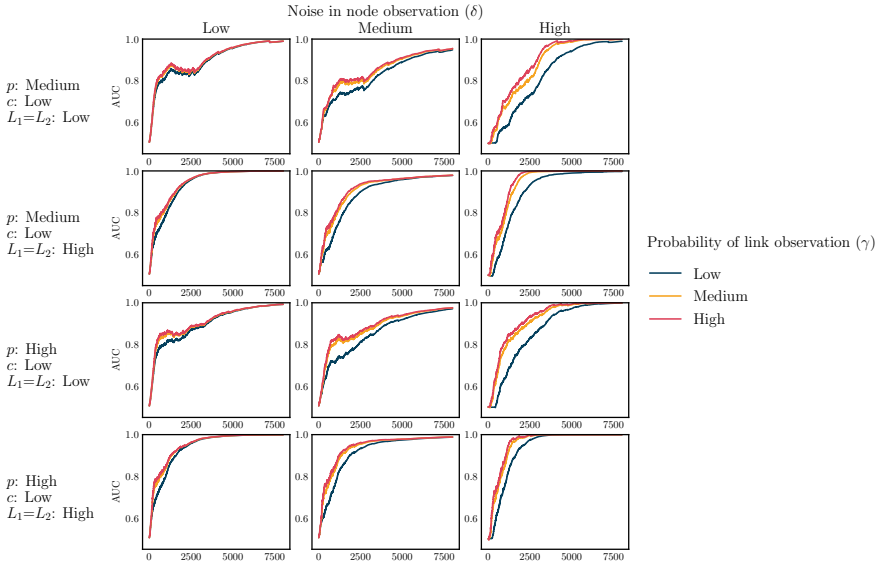


Fig. 15 AUC at various time steps for γ (low, medium, high) over different δ and 40 Monte Carlo simulations (k: medium)

social influence, counterintuitively, the authority may perform better when the noise in node observation is higher in this case as well.

The convergence of the AUC is also not as monotonic as is in the case without social influence due to agents turning into participants. This is especially true when it is easier for "neutral" agents to become participants (rows 1 and 3). Figure 15 presents results of the AUC over time for different configurations of the participant parameters and noise in node observation parameter while sampling frequency is held constant at 1 agent observation per 10 time steps. We find that the main result from the no social influence case holds here—there are still diminishing returns to the improvements made to the AUC by improving the link observation process. However, what is different is that the biggest improvements in AUC are made at the beginning before new agents start becoming participants. Further, if the noise in the node observation process is low and the networking and outreach parameters are low, then the improvement in AUC over time is non-monotonic. The AUC starts to dip in the period of dynamic movement growth and then improves again once the number of participants stabilizes (Fig. 16).

Sampling Frequency (k)

Changing sampling frequency from every 5 to every 10 time steps (which corresponds to an average of 16 and 8 observations per node) seems to have a similar effect on

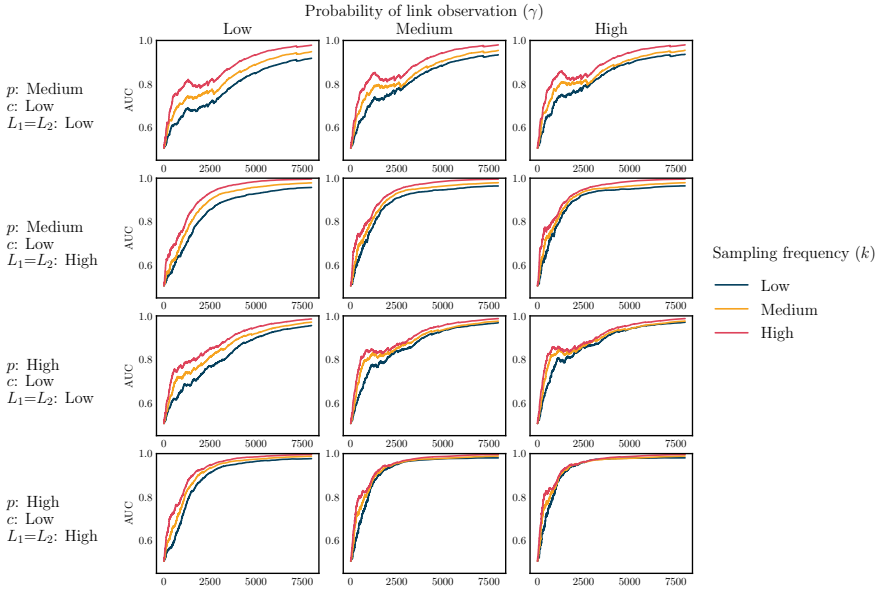


Fig. 16 AUC at various time steps for k (low, medium, high) over different γ and 40 Monte Carlo simulations (δ: medium)

increasing time to convergence as changing it from every 10 to every 20 time steps (which corresponds to an average of 8 and 4 observations per node), just like in the case of no social influence. The only difference in the trend is due to the dynamic movement growth, which makes the convergence less monotonic than the no social influence case.

The results are summarized succinctly in Fig. 17. This figure shows gains to AUC over time for various changes in each of the three authority parameters. The results are similar to the case without social influence. The gains from improving technology consistently show diminishing returns (especially over longer time horizons) except in the case of reducing noise in the node observation where we see increasing returns (i.e., reducing the node observation noise from medium to low shows more improvement in AUC than reducing it from high to medium).

Summary and Future Work

Communication technologies are not intrinsically good or bad for civil society. They can be used by violent extremists to coordinate terrorist attacks or by advocates of just social causes to organize legitimate protests. Governments can use communication technologies to track and suppress activists or to track and neutralize terrorists. Rather than focusing on participants' or authorities' objectives, this work aims to understand

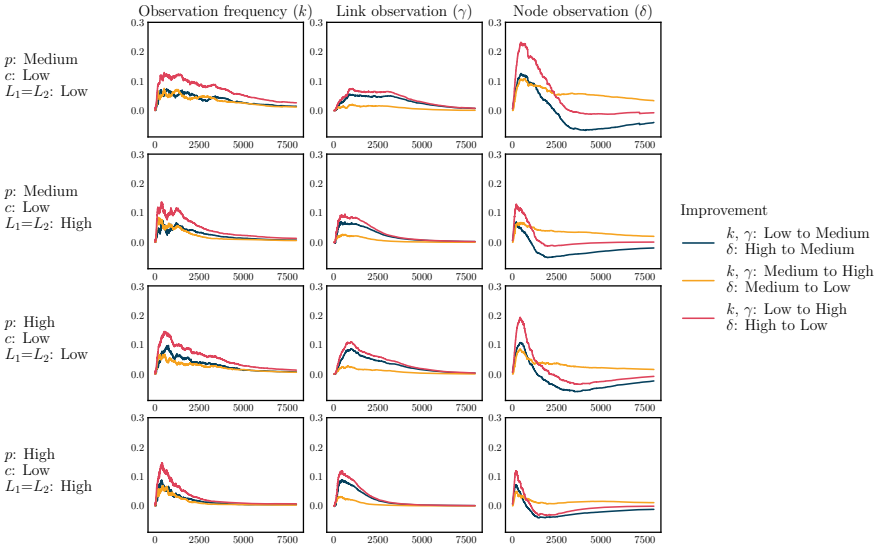


Fig. 17 AUC improvement over time for improvements in k, γ, and δ

better the dynamics of technological facilitation of both sides' goal, whatever they may be. One of the assumptions of our model is that participants need to recognize themselves as a group, build a network to support their actions, and recruit new supporters. We also assume that effective and efficient actions on the authority side require first recognizing who is a member of the protest and who is only a bystander. Consequently, we study the impact of communication technologies on the ability to perform those functions.

Our model yields some interesting results. First, the model reflects a common observation of sociologists and political scientists who try to predict how much support a given social movement might receive. Namely, depending on the underlying characteristics of the population, the final number of participants can significantly vary. Secondly, the model indicates that in the initial phase of social movement formation, technologies that enable casting a wide net accelerate both the social movement growth and the network formation more than those allowing elite, invitation-only membership. That is, open forums can lead to mobilization faster than closed chat rooms, even if members of those closed groups might be more direct about their goals. However, due to the snowball effect, the bigger the network, the bigger the utility of friend-of-friends networking technologies.

On the other hand, faster network growth enables better classification of agents by the authority. This is an artifact of the adopted link formation rule allowing participant to have more links than non-participants, making them stand out in the population. In real life, this property might be more difficult to observe as participants might deliberately limit their communication with others, precisely to stay below law enforcement radar as long as possible.

On the authority side, our model suggests that ensuring a reliable link observation process can result in reasonable accuracy even when the node observation process is noisy. The fact that our model allows the authority to perform better when the node observation process is noisier signals the importance of having a score function that correctly aggregates participant data. Further, we find that depending on the size of the social movement, the authority's belief formation is made more accurate by improving a different aspect of the surveillance process. Reducing the noise in node observation improves accuracy more effectively when there are fewer participants and lower outreach and networking. On the other hand, the improvements in accuracy made from increasing the sampling frequency are fairly independent of social movement characteristics.

As mentioned several times, one way to view our model is as a building block to more complex models of social movements that integrate participant and authority objectives. For that reason, an obvious extension is to integrate our network formation and surveillance model into a larger model of social movement. This would include adding participant and authority objectives and actions. For example, one research question might be the optimal time for which the participants decide to mobilize and act. From the authority's perspective, it might choose when to disrupt communication channels by either severing links or removing individuals they believe to be key participants (such as through incarceration).

It is also possible to build on our baseline model of participant dynamics. For example, in our model, citizens sample other citizens at random. In the real world, citizens will likely sample other citizens strategically in an effort to actively seek connections. Therefore, one extension to our model would be to incorporate more sophisticated networking and outreach behavior. Additionally, our model only allowed agents to switch from non-participants to participants and not switch in the other direction, which is, of course, possible in real social movements. This is also true of the authority; the authority may not randomly sample citizens to monitor but might strategically dedicate resources to monitoring key citizens. Furthermore, in the current model, the authority monitors links regardless of how they were formed. Strategic allocation of resources could also account for a difference in costs of monitoring open forums (outreach technologies) versus closed forums (networking technologies).

One hurdle in extending the model to include more complex and strategic behavior is determining how to modify the authority's belief process. While in this work we proposed a simple score that uses partial Bayesian updating, this score will likely be too imprecise to be relevant with more complex stochastic processes. However, the burgeoning fields of graph embeddings and graph neural networks [32] provide a promising path forward. Specifically, graph neural networks take as an input observed graph features (node features, links, etc.) and attempt to predict other features such as unobserved node characteristics. This is *precisely* what the authority wants to accomplish. Specifically, the authority observes noisy links and noisy node features and attempts to predict whether the node is a participant. Consequently, graph neural networks provide a promising direction to generalize the process in which the authority assigns scores as a function of its observations and thus allows the model to incorporate much more complex network formation dynamics.

Appendix: Model Design Concepts and Details

Design Concepts

Interaction and Emergence

Over the course of simulation, agents interact surveying each other to create links. Consequently, a network of agents emerges. To measure the connectivity among movement participants, we use their degree, i.e., the number of links to other participants.

Prediction

We are interested in an aggregate measure that captures the accuracy of the authority's belief across all agents. For this, we will use ROC curves and compute the area under the curve (AUC) [33] as an aggregate measure of the authority's beliefs. A ROC graph is a two-dimensional graph with the true positive rate plotted on the y-axis and the false positive rate plotted on the x-axis, and each point on the graph represents a different classification threshold. Specifically, the bottom left of the graph (0,0) represents a strategy of never issuing any positive classifications—so there are no false positives, but the true positive rate is also zero. Conversely, the upper right (1, 1) represents a strategy of always issuing positives. The point (0,1) represents a perfect classifier. The diagonal represents the strategy of randomly guessing a class. As a rule of thumb, one classification strategy is better than another if it yields a point that's located more northwest than the other in the ROC space.

While a ROC graph represents the performance of a classifier in two dimensions by varying the classification threshold, it may often be helpful to reduce this performance down to a single scalar value to compare the performance of multiple classifiers. One such method is to calculate the area under the curve (AUC) in a ROC graph. Since the AUC is some portion of a unit square, its value is always between 0 and 1. However, as mentioned previously, the diagonal line corresponds to a random guessing strategy, which corresponds to an AUC of 0.5. Consequently, any realistic classifier should have a value between 0.5 and 1. The AUC has an important interpretation which makes it a reasonable proxy for classifier performance. It is equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance. Tracking the AUC over time allows examining the changes in classifier quality throughout a simulation.

Initialization

The main object of the model is a time-indexed graph $\mathcal{G}_t = (V, \mathcal{E}_t)$ where V is the set of N vertices or nodes and $\mathcal{E}_t$ is the set of undirected edges at time t, and time is discrete.[2] Each vertex represents an agent and edge $e_{ij}^t \in \mathcal{E}_t$ represents a communication channel between agent i and agent j at time t.

Each agent i, (equivalently, a vertex) has two attributes, $\beta_i^t \in \{0, 1\}$ and $\theta_i \in (0, 1)$. The binary attribute β_i^t indicates whether the agent is a participant and equals 1 if the agent is a participant and 0 otherwise. The uniformly distributed attribute θ_i represents an agent's identity and captures characteristics other than its proclivity to protest. Again, this may represent the agent's socioeconomic status or a collection of interests. While β_i^t can change over time, θ_i is fixed. Initially, there are p participants in the population of N agents.

Inputs

None.

Submodels

Participant Link Formation and Evolution

At each time step t, an agent is selected at uniform random. Denote the agent selected at time t as a_t. Agent a_t can form links in two ways:

- ***Outreach*** First, agent a_t randomly queries L_1 other agents from the set $V_{a_t}^t = \{v \in V | v \neq a_t, (a_t, v) \notin \mathcal{E}_t\}$, i.e., any agent that is not already connected to a_t. Let $v \in V_{a_t}^t$ be any of the L_1 agents that agent a_t queries. Then, the edge (a_t, v) is added to $\mathcal{E}_t$ if *either* of the two conditions is met:

  1. $\beta_{a_t}^t = \beta_v^t = 1$
  2. $|\theta_{a_t} - \theta_v| \leqslant c$

  Intuitively, an edge is formed if either both agents are movement participants or their value of θ is sufficiently close.
- ***Networking*** Second, agent a_t randomly queries L_2 agents from the set $\widetilde{V}_{a_t}^t = \{v \in V | v \neq a_t, (a_t, v) \notin \mathcal{E}_t, (a_t, j) \wedge (j, v) \in \mathcal{E}$ for some $j\}$ and connects according to the same conditions as when it conducts outreach. Intuitively, the set $\widetilde{V}_{a_t}^t$ is all nodes where the shortest path between a_t and a node in $\widetilde{V}_{a_t}^t$ is 2.

[2] Since edges are undirected, as a shorthand, we use notation such that the edge $(x, y) = (y, x)$.

These two processes define the evolution of $\mathcal{E}$. Specifically, $\mathcal{E}_t$ is the union of $\mathcal{E}_{t-1}$ and the links that are formed during outreach and networking at time t.

After all of the links at time t are formed, agents update their value of β. Specifically, if agent i is connected to at least w participants and is itself not a participant, β_i^t switches from 0 to 1. Note that this is an iterative process since the effects of an individual agent turning into a participant can spread throughout the network if that agent is connected to several other agents that were previously connected to $w-1$ participants.

The key parameters are L_1, L_2, c, p, and w. Specifically, L_1 controls an agent's ability to leverage communication technology for outreach, where L_2 represents the agent's ability to leverage communication technology for networking. The parameter w represents how susceptible agents are to social influence. The parameter p defines the number of participants at $t=0$, i.e., the initial seed of the movement. The parameter c represents *inclusivity* since it represents how much agents are willing to include other agents with different identities into their network. Furthermore, c is the inverse of homophily. The lower the value of c, the more agents need to have similar identities for them to connect. Since we are interested in communication technologies, our main computational experiments explore how changing L_1 and L_2 impact the growth of the social movement and how that growth rate depends on c, w, and p, which are measures of homophily, social influence, and the initial movement size, respectively.

To measure connectivity among the participant, we repeat each computational experiment with the same parameter values 40 times, compute the average participant degree and then compute the average over each of the 40 Monte Carlo simulations. Specifically

$$\text{Average participant degree at time t, } \bar{d}_t = \frac{1}{40}\sum_{m=1}^{40}\frac{1}{p}\sum_{i=1}^{N}\beta_i^{t,m}d_i^{t,m} \tag{1}$$

where the superscript m corresponds to simulation number m.

Similarly, the netwrok size is measured as

$$\text{Average number of participants at time t, } \bar{p}_t = \frac{1}{40}\sum_{m=1}^{40}\sum_{i=1}^{N}\beta_i^{t,m} \tag{2}$$

Authority Observation

The authority agent dynamically observes the network as it forms, but not necessarily at every time step. Specifically, every k time-step, the authority observes m citizens. The parameter k represents how fast an authority can process data obtained through citizens and draw insights, while m represents how much data an authority can col-

lect simultaneously. Of course, the actual data collection and processing procedures are likely asynchronous, but we assume the processes are synchronous to maintain simplicity.

For each citizen i of the m citizens sampled at time t, the authority has a **node observation process** where it observes y_{it}, which equals β_i^t with probability $1-\delta$ and $1-\beta_i^t$ with probability δ. Intuitively, the authority observes the correct value of whether or not the citizen is a participant with probability $1-\delta$ but makes an error with probability δ.[3] The parameter δ captures yet another notion of surveillance technology; the lower δ, the more accurate the surveillance. The authority keeps a record of these observations $y_i^t = \{y_{i1}, y_{i2}...y_{ik_t}\}$ where y_{ij} represents the authority's jth observation of citizen i and k_t represents the total number of times the authority observed agent i up to and including time t.

In addition to the value of β_i^t, for each agent i of the m sampled agents, the authority also completes an **edge observation process** where it observes $e_{ij}^t \in \mathcal{E}_t$, $j \in \{1, 2...N\}$ with probability γ. This process represents the authority's ability to conduct surveillance on a citizen's network. If $\gamma = 0$, then the authority cannot observe network links. If $\gamma = 1$, the authority perfectly observes a citizen's connections. The authority also keeps track of the degree of each agent i, given by

$$d_i^t = |\{e_{ij}^t | \exists t \text{ such that the authority observes } e_{ij}^t\}|$$

To quantify the authority's belief on whether a citizen is a participant, it must combine observations y_i and d_i. It does this through the function

$$s_i^t = (r_i^t)^{\alpha}(d_i^t)^{1-\alpha} \tag{3}$$

where

$$r_i^t = \frac{\prod_{t=1}^{k_t}(1-\delta)^{y_{it}}\delta^{1-y_{it}}\frac{p}{N}}{\prod_{t=1}^{k_t}(1-\delta)^{y_{it}}\delta^{1-y_{it}}\frac{p}{N} + \prod_{t=1}^{k_t}\delta^{y_{it}}(1-\delta)^{1-y_{it}}(1-\frac{p}{N})} \tag{4}$$

Intuitively, r_i^t is the authority's Bayesian posterior belief of citizen i being a participant *if it only completes a node observation process*. However, to leverage the data collected from the edge observation process, it weights this probability by the node's degree to obtain s_i^t. Again, the motivation is that a participant will, on average, have a higher degree than a non-participant, so raising an agent's degree would increase an agent's score. In the case of social influence, one could expect that if the participant status of each citizen can change over time, the authority should limit their "memory" to some number of last observations. However, the experiments conducted for the model with social influence have shown that constraining memory negatively affects the classification. Consequently, we allow the authority to form their belief based on all previous observations of the citizen node.

[3] Without loss of generality, we assume $\delta \leqslant 0.5$.

Table 2 Participants' parameters

Characteristics	Parameter	Values
Underlying characteristics of the population	Population size N	100
	Initial seed of movement p	5, 10, 15
	Inclusivity c	0.05, 0.075, 0.10
	Social influence w	∞ (no influence), 5, 4, 3
Technological capabilities	Outreach technology L_1	1, 2, 3
	Networking technology L_2	1, 2, 3

Table 3 Authority's parameters

Characteristics	Parameter	Values
Surveillance capacity	Sampling frequency k	20, 10, 5
Surveillance accuracy	Noise in node observation δ	0.1, 0.3, 0.5
	Probability of link observation γ	0.1, 0.5, 0.9

Parameters

We vary both the quality of technologies themselves and the underlying characteristics of the population, according to the parameter values in Table 2. For the authority agent, we collapsed sample size and sampling frequency into one parameter reflecting the frequency of sampling one agent. Parameters defining surveillance technologies were varied separately for node and edge observation processes, as shown in Table 3. While the exact values of the parameters are not necessarily of interest, we performed several robustness checks to ensure our experiments captured the full spectrum of the parameter space. To ease presentation, in the subsequent analysis, we refer to the parameter values as "low", "medium", and "high" instead of their precise numerical values as specified in Tables 2 and 3.

References

1. Stewart, C.J., Smith, C.A., Denton Jr., R.E.: Persuasion and Social Movements. Waveland Press (2012)
2. Tobin, J., Tobin, J.L.: Hidden in Plain View: The Secret Story of Quilts and the Underground Railroad. Anchor (2000)
3. Calhoun, B.: Shaping the public sphere: english coffeehouses and french salons and the age of the enlightenment. Colgate Acad. Rev. **3**(1), 7 (2012)
4. Lok-kei, S.: Hong Kong Reddit: how Leaderless Extradition Protests Took a Lead from Social Media. South China Morning Post
5. Tkacheva, O.: Internet Freedom and Political Space. Rand Corporation (2013)

6. McPherson, M., Smith-Lovin, L., Cook, J.M.: Birds of a feather: Homophily in social networks. Ann. Rev. Soc. **27**(1), 415–444 (2001)
7. Shrum, W., Cheek Jr, N.H., MacD, S.: Friendship in school: Gender and racial homophily. In: Sociology of Education, pp. 227–239 (1988)
8. Ruef, M., Aldrich, H.E., Carter, N.M.: The structure of founding teams: Homophily, strong ties, and isolation among us entrepreneurs. In: American Sociological Review, pp. 195–222 (2003)
9. Colleoni, E., Rozza, A., Arvidsson, A.: Echo chamber or public sphere? predicting political orientation and measuring political homophily in twitter using big data. J. Commun. **64**(2), 317–332 (2014)
10. Friedkin, N.E., Johnsen, E.C.: Social influence and opinions. J. Math. Soc. **15**(3–4), 193–206 (1990)
11. Hamari, J., Koivisto, J.: "working out for likes": an empirical study on social influence in exercise gamification. Comput. Hum. Behav. **50**, 333–347 (2015)
12. Aral, S., Walker, D.: Creating social contagion through viral product design: a randomized trial of peer influence in networks. Manag. Sci. **57**(9), 1623–1639 (2011)
13. Di Gregorio, M.: Networking in environmental movement organisation coalitions: interest, values or discourse? Environ. Polit. **21**(1), 1–25 (2012)
14. Dincelli, E., Hong, Y., DePaula, N.: Information diffusion and opinion change during the gezi park protests: Homophily or social influence? Proc. Assoc. Inf. Sci. Technol. **53**(1), 1–5 (2016)
15. Stern, P.C., Dietz, T., Abel, T., Guagnano, G.A., Kalof, L.: A value-belief-norm theory of support for social movements: the case of environmentalism. Hum. Ecol. Rev. 81–97 (1999)
16. Centola, D.: How Behavior Spreads: The Science of Complex Contagions (Introduction). Princeton University Press (2018)
17. Aggarwal, C., Subbian, K.: Evolutionary network analysis: a survey. ACM Comput. Surv. (CSUR) **47**(1), 1–36 (2014)
18. Bollobás, B., Riordan, O.M.: Mathematical results on scale-free random graphs. In: Handbook of Graphs and Networks: From the Genome to the Internet, pp. 1–34 (2003)
19. Atalay, E., Hortacsu, A., Roberts, J., Syverson, C.: Network structure of production. Proc. Natl. Acad. Sci. **108**(13), 5199–5202 (2011)
20. Guimera, R., Sales-Pardo, M., Nunes Amaral, L.A.: Modularity from fluctuations in random graphs and complex networks. Phys. Rev. E **70**(2), 025101 (2004)
21. Kim, K., Altmann, J.: Effect of homophily on network formation. Commun. Nonlinear Sci. Numer. Simul. **44**, 482–494 (2017)
22. Cioffi-Revilla, C., Rouleau, M.: Mason rebeland: an agent-based model of politics, environment, and insurgency. Int. Stud. Rev. **12**(1), 31–52 (2010)
23. Makowsky, M.D., Rubin, J.: An agent-based model of centralized institutions, social network technology, and revolution. PloS one **8**(11), e80380 (2013)
24. Epstein, J.M.: Modeling civil violence: an agent-based computational approach. Proc. Natl. Acad. Sci. **99**(suppl 3), 7243–7250 (2002)
25. Moro, A.: Understanding the dynamics of violent political revolutions in an agent-based framework. PLOS one **11**(4), e0154175 (2016)
26. Lemos, C.M.: Agent-Based Modeling of Social Conflict: From Mechanisms to Complex Behavior. Springer, Berlin (2017)
27. Downey, D.: Convergence versus emergence of youth extremism: An agent-based model of the arab spring. In: Youngman, P.A., Hadzikadic, M. (eds.) Complexity and the Human Experience, pp. 183–202. Pan Stanford, Boca Raton (2014)
28. Lemos, C., Coelho, H., Lopes, R.J., et al.: Agent-based modeling of social conflict, civil violence and revolution: State-of-the-art-review and further prospects. In: EUMAS, pp. 124–138. Toulouse (2013)
29. Lohmann, S.: The dynamics of informational cascades: the monday demonstrations in leipzig, east germany, 1989–1991. World Polit. **47**(1), 42–101 (1994)
30. Waldherr, A., Wijermans, N.: Modelling the role of social media at street protests. In: *Advances in Social Simulation 2015*, pp. 445–449. Springer, Berlin (2017)

31. Hai-hua, H., Cui, W.-T., Lin, J., Qian, Y.-J.: Icts, social connectivity, and collective action: a cultural-political perspective. J. Artif. Soc. Soc. Simul. **17**(2), 7 (2014)
32. Zhou, J., Cui, G., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., Sun, M.: Graph neural networks: a review of methods and applications (2018). arXiv:1812.08434
33. Fawcett, T.: An introduction to roc analysis. Pattern Recognit. Lett. **27**(8), 861–874 (2006)

Foraging Games: Ideal and Not

Robin Clark and Steven O. Kimbrough

1 Introduction

Distributions of populations over a landscape are of keen interest in biology and many of the social and behavioral sciences, including linguistics, economics, anthropology, and archeology. We are interested in exploring how individual preferences affect these distributions, in particular how distributions are affected by dynamic preferences built from the situation at hand. Particularly, in dynamic situations the information processing abilities and practices of the relevant agents are crucial. Small changes in their information processing abilities and practices can have large impacts on behavior and the resulting distributions of the agents.

We begin by raising the fundamental question of whether preferences are retrieved or formulated. Following that, we examine in some detail a foraging model built upon and extending the classic ideal free distribution (IFD) model. We find, through agent-based modeling, that very slight departures from the stringent IFD assumptions lead to significant changes in the resulting distribution. The upshot of the paper weighs in favor of preference construction, both in fact and for modeling.

R. Clark
Department of Linguistics, University of Pennsylvania, Philadelphia, PA 19104, USA
e-mail: rclark@sas.upenn.edu

S. O. Kimbrough (✉)
Department of Operations, Information and Decisions, University of Pennsylvania, 3730 Walnut St., Philadelphia, PA 19104, USA
e-mail: kimbrough@wharton.upenn.edu

Z. Yang and S. Núñez-Corrales (eds.), *Proceedings of the 2022 Conference of The Computational Social Science Society of the Americas*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-031-37553-8_14

2 Preferences

Under a widely received view, decisions (alias choices) are made on the basis of preferences.[1] Given a menu of alternatives, decision makers choose based on what they prefer, presumably acting so as to choose what they prefer the most. Preferences, moreover, are in this tradition generally thought to be stable, produced from memory when needed (either directly by retrieval or indirectly by functional transformation of retrieved information), and global in the sense that for all possible choices, a_i versus a_j, the decision maker prefers a_i to a_j (we write $a_i \succ a_j$) or prefers a_j to a_i ($a_j \succ a_i$) or is indifferent between the two ($a_i \sim a_j \sim a_i$).[2] Further, it is a fundamental tenant in models in economics, game theory, and decision theory that agents are rational in the sense that their preferences conform to the axioms one version of utility theory or another (e.g., [9, 13, 14, 16]). Together, the bundle of these assumption propositions constitute the core of Rational Choice Theory (RCT).

The obvious alternative to retrieval of preferences is to view them as constructed by the agent when needed (and perhaps stored in memory after creation). That is our focus in this paper. There is by now an extensive behavioral literature establishing that indeed preferences are, at least often, so constructed. See [20] for a review; see also [12]. Preferences when empirically studied are often far from stable. They are subject to priming and environmental conditions. They are (often) constructed for the occasion rather than retrieved from memory. A fortiori, they are not global; there simply are not pre-existing preferences for everything. The behavioral and psychology literature in these regards is by now compelling. That literature, however, has not developed much more than a rudimentary account of the mechanisms by which preferences are created or constructed (alias formulated).

Thus, we seek to develop and explore models of preference formation by agents. Where to begin? Benjamin Franklin's justly celebrated method, expressed in a letter to Joseph Priestly, is the kind of account we seek https://www.1000minds.com/decision-making/benjamin-franklin.

In light of empirical findings about preference construction (noted above), a natural interpretation of Franklin's method is that it is an account, both descriptive and prescriptive, of how preferences may sensibly be constructed (by listing pro and con factors and striking balanced pairs). Franklin proffers the approach as a way to circumvent cognitive limitations. When there are very many pros and cons we cannot keep proper track of them. Franklin thus proposes a physical procedure for making the comparisons without decision makers having to use more than all of their limited cognitive resources.

In his letter, Franklin offered a procedure—a "moral or prudential algebra"—by which to employ data available at the time of decision and arrive at a preferred choice. What is preferred is discovered or calculated by this procedure, not dredged up from a stable pool of memories.

---

[1] Terminology is not fully standardized. We do our best to use mainstream senses, but the reader should understand that there is other terminology floating about.

[2] This is an axiom of utility theory. See [20] for an authoritative review.

Generalizing Franklin's letter, the principle involved is one of *rule-based preference formation*. Beginning in the next section, we explore this principle for explanatory purposes in the analysis of foraging behavior and preference formation. First, however, we wish to acknowledge and address an objection to our framing and approach. In doing so we introduce some nuance, a refinement of the narrative given so far. The objection is that the distinction between retrieving preferences from memory and constructing them by recalling rules for generating them is a distinction without a real difference. "Of course," the objection goes, "no one thinks that a forager would have a fixed preference for patch A over patch B. Instead, the forager has a fixed preference for alternatives that maximize food intake. What is fixed and stable and recalled from memory is the rule to prefer more food. If you want to call this preference construction you can, but you are not introducing anything new or different. What's stable and retrieved is the rule/preference for more food."

Our response is first that maximizing food intake is a value or goal, not a preference for how to achieve it (by choosing a patch). Second, if stable and simple rules can be identified that account for the behavior in question, then the retrieval–or–construction debate would presumably be resolved in favor of retrieval. But if simple rules are not sufficient for explaining observed behavior, while rather more complex rules are sufficient, then at some point of increased complexity, a retrieval account becomes misleading in a way that a construction account is not. In the end, it matters little where exactly we draw the line. What matters is delineating correctly the mechanisms involved.

In what follows we use agent-based modeling to simulate and investigate foraging. The ideal free distribution (IFD) hypothesis, explained in the sequel, is the immediate target of our modeling. Foragers will be distributed as in the IFD under standard assumptions and Rational Choice Theory. We replicate this finding and go on to demonstrate that small departures from the IFD assumptions lead to large departures from IFD behavior. This creates opportunities for foragers to learn and use revised rules for choosing foraging patches. Whether they do so or not is an empirical question. But to the extent they do, to that extent at least a construction account of preference formation is closer to the mark.

3 Ideal Free Ducks

Food and other resources are often patchy, appearing more in one place than in another, and absent in most places. When this is the case, individuals in a community are typically faced with a strategic decision problem of where to forage for resources. Without other individuals involved, foraging at the most productive patch, net of transit and other costs, would be optimal. The individual, however, is affected by decisions made by other individuals in its community. Perhaps foraging at a less productive patch will yield an individual more in return than foraging at a more productive patch because the less productive patch is comparatively less crowded.

Thus we have a foraging game. How will it turn out in equilibrium? Fretwell was the first to investigate this framing of foraging theory. In a series of papers [2, 3, 5] and a book, [4], Stephen Fretwell investigated this question both theoretically and empirically. The literature since has blossomed.

Fretwell developed a model he called ideal free distribution (IFD) for predicting the distribution of individuals in a community across a diversity of foraging patches.[3] Assuming ideal rationality and identity of capabilities among the individuals, Fretwell posited that at equilibrium individuals would be arrayed among patches such that each individual would obtain an equal amount of resource per unit of time. This *ideal free distribution* (IFD) can be expressed as follows.

Let N_i be the number of foragers at site i and R_i be the level of resources available at site i, which the foragers at the site share equally, then the IFD is

$$\frac{N_i}{\sum_i N_i} = \frac{R_i}{\sum_i R_i} \tag{1}$$

With just two foraging patches this reduces to

$$\frac{N_1}{N_2} = \frac{R_1}{R_2} \text{ or equivalently } \frac{N_1}{R_1} = \frac{N_2}{R_2} \text{ and in general } \frac{N_i}{R_i} = \frac{N_j}{R_j} \ \forall i, j \tag{2}$$

As an aside, note that the relationship between the N_is and the R_is is exactly that posited by the matching law of [8], which is widely used in the description of animal learning. Let the B_is be the behavior alternatives, and the R_is be the associated reinforcements; then the matching law says that at convergence of learning, we have:

$$\frac{B_i}{B_1 + B_2 + \cdots + B_n} = \frac{R_i}{R_1 + R_2 + \cdots + R_n} \tag{3}$$

Although it was not designed for modeling strategic decision making, the associated hypothesis of *melioration* [8] addresses the dynamics of learning under the matching law. According to this hypothesis [8, page 77], the adapting/learning agent adjusts its B_is so that:

$$\frac{R_1}{B_1} = \frac{R_2}{B_2} = \cdots = \frac{R_n}{B_n} \tag{4}$$

Returning to Fretwell's ideal free distribution, it is assumed throughout that the sum of the R_is is fixed, at least comparatively so. In the case of two alternatives, doing R_1 at a time entails not doing R_2, for example.

[3] We focus as does the literature on conspecifics in a single habitat.

Moreover, Fretwell's basic model assumes the following properties of Ideal Free foraging:

- Animals need to forage from patches of varying qualities distributed over a region.
- Animals will tend to distribute themselves over these patches in order to maximize their individual returns.
- They will therefore tend to distribute themselves so that the resources will be distributed evenly, taking into account the fecundity of the resource patch and the number of conspecifics exploiting the patch. This is the insight behind deal free distributions [2].
- Elaborating on the above point, suppose that some animal is getting more resources than would be expected if the resources were distributed evenly. This means that some other animal is getting less than expected; the latter animal should, then, move to the patch where the other animal has more than its share. As the foragers reshuffle, they should eventually hit an equilibrium point where all animals are getting the same amount.
- In particular, if F_i is the fecundity of patch i, and n_i is the number of conspecifics at patch i, then:

$$\frac{F_i}{n_i}$$

  is the return that an individual forager can expect when occupying i.
- Finally, the basic IFD model assumes that the animals at a patch are dividing resources evenly, so that they all get the same amount. This need not be the case. For example, some animals may be more efficient at foraging, or can bully the other animals, in which case the distribution of resources will not be equal. This should effect the distribution of animals per patch.

4 IFD Modeled in NetLogo

We constructed a model in NetLogo (IFD-Foraging01.nlogo) of foraging based on the ideal free distribution with two food sources (patches). Our model is somewhat more general than the pure IFD model because it allows for the animals to transit non-instantaneously from one patch to the other. Because we based the model on experiments by [7] involving ducks wintering at a pond, we will henceforth refer to the agents (or "animals") as *ducks*. The IFD mathematical model supposes that the ducks can move instantaneously from one food source to another. Our NetLogo model permits a number of variations, which turn out to be significant, as we shall see.

4.1 The Foraging Procedure

The IFD-Foraging01.nlogo model is freely available online and gives full access to its code.[4] Here we sketch in pseudocode the main foraging procedure, occurring at each time step in the execution of the model.

1. Collect statistics on system state.
2. Ask food sources to distribute food according to their schedules
3. Ask animals to collect and record food and determine if they want to move to a different patch.
4. if `transit?` is true
   (a) Decrement the transit-counter for all animals in transit;
   (b) For animals that have completed transit, place them at their new destination food source;
   (c) Put into transit all animals that have decided to move and set their transit counters (indicating the amount of time to be in transit).
5. Ask the animals not currently in transit:
   (a) Increment the time counter for presence of the food-source they are at;
   (b) Determine whether to move to a new source immediately; this is implemented as the procedure `move-immediate` in the code (see below).
   The decision depends on the probability of moving `p-move`, the number of animals at the other site and the fecundity of the other site.
   (c) If `to-move` is `true` and `transit?` is `false`, move immediately to the other site;
   (d) If `to-move` is `true` and `transit?` is `true`, move to a special transit patch and wait.
   (e) Otherwise, the animals stay where they are and wait for more food.

Much of the central intuition of the IFD model is captured in `move-immediate` which, in essence, computes the "expected utilities" (values) of the bread[5] sources by looking at the rate at which bread is tossed out versus the number of animals at the food source. Note that `move-immediate` returns a Boolean value contingent on the expected amount of bread at each patch.

1. If the bread per animal ("duck" in the NetLogo model) is larger at the patch occupied by the duck, return False;
2. Otherwise, if the bread per duck is less at the present patch, return True with probability `p-move`; otherwise return False.

[4] https://github.com/stevenokimbrough/sokpapers/commit/3b3d5ccdaf0bce9d71ca86a34c2e05d11b391ad7.

[5] We are grateful to Annie Vo for reminding us that ducks should not be fed bread as it messes with them. Here and hereafter, when we say "bread" let us mean "units of healthy food for ducks."

4.2 Behavior of the Pure IFD Configuration

We constructed a "pure" model of the Ideal Free Distribution. This configuration is set in the NetLogo model by switching `pin-ducks?` to off on the Interface tab (see Fig. 1). The agents (the foraging animals, which we refer to as ducks) are "ideally rational" They are "free," meaning they can move instantaneously and without cost to any food patch; and they are "ideal" in the sense that they have perfect information about the fecundity of each food patch, which they use to maximize intake of resources. Furthermore, they know the number of competitors at each food patch, although they cannot predict the future of how many ducks will be at each patch. Thus, in this model the ducks should distribute themselves in such a way as to guarantee that each duck acquires the same amount of food as any other duck, in equilibrium. When we run the model as a pure IFD system, this is exactly what we see. The ducks divide themselves over the food patches in such a way that each duck gets the same amount of food as any other duck, even though the food sources are differentially productive.

In a typical case, we ran the model for 3500 time steps with an initial sample of 500 time steps, allowing the ducks to observe the rate of bread distribution at the food patches. The run had (the IFD-Foraging01.nlogo model has) two food patches. In the run there were 250 ducks, with the rate of bread distribution being 15 units

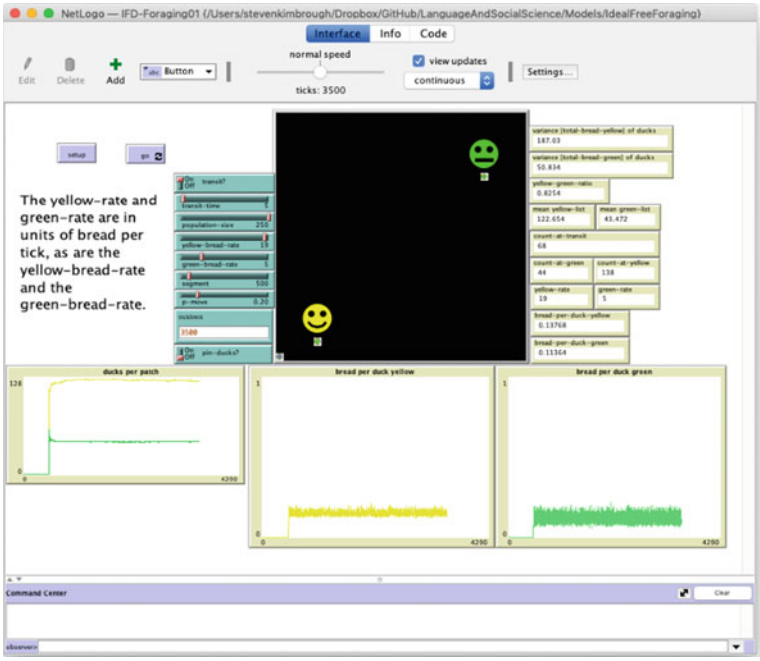


Fig. 1 Interface of the IFD-Foraging01.nlogo model

of bread per unit time versus 5 units of bread per unit time. The ducks arranged themselves in such a way as to yield 0.080 units of bread per unit of time at the more fecund food patch and 0.079 units of bread per unit time at the less fecund patch, basically identical yields of food per patch. Since the rate of bread distribution at the more fecund patch was three times the rate of the less fecund patch, we expect, then, that three times as many ducks will congregate at the first patch as the number that congregate at the less fecund patch. Again, that is exactly what we see: 187 ducks at the first patch to 63 ducks at the second patch. If we look at the mean of the total bread collected by ducks at the more fecund patch and compare it to the mean of the total bread collected at the less fecund patch, we expect the former to be three times the latter, and it is: 180 units collected at the more fecund patch to 60 at the less fecund patch. In addition, the variance of the total bread collected per food source should be nearly identical and it is: 30.06 for the more fecund patch versus 29.74 for the less fecund patch.

Within a narrow margin of statistical error (noise), the behavior of the Pure IFD NetLogo model (with `transit?` set to off) conforms exactly to the predictions of the ideal free distribution.

5 Beyond Purity

The Pure IFD model exists in a clockwork universe that is at quite a remove from the world of actual foragers, who cannot move instantaneously between food sources and who may not have perfect information about the fecundity of the patches or the number of competitors at each patch. In the next experiment, we will increase the cost of moving from one food source to the other; it will no longer be instantaneous but will rather take some amount of time. We implement this by creating a "transit patch" where ducks are confined for some number of ticks when they decide to switch patches. While confined at the transit patch, the ducks cannot acquire more bread; that is, the ducks temporarily cease foraging while traveling from patch to patch. Now, `transit?` is switched to on.

In the first configuration, when they decide to move from patch to patch, the ducks must spend 5 time ticks in the transit patch, where they receive no resources. Once again, we ran the model for 3500 time steps with an initial sample of 500 time steps for warm up, allowing the ducks to observe the rate of bread distribution at the food patches. As above, the run had two food patches, 250 ducks and the rate of bread distribution being 15 units of bread per unit time versus 5 units of bread per unit time. If we allow the probability of moving to be 1 when a duck decides its yield would be greater at the other patch, we have a striking result. After an initial apparently chaotic period, the number of ducks at the more productive patch is 2, while the number of ducks at the less productive patch is 1. The remaining 247 ducks are in transit and, therefore, not receiving bread. Thus, the ducks at the more fecund patch get 7.5 units of bread per unit time (there are 15 units of bread distributed per unit of time); the single duck at the less fecund patch receives 5 units of bread per unit time. That is,

of course, the maximum any duck could receive per unit of time at that patch. In this sense, a true ideal free distribution would be unattainable, due to the hard limit on the less fecund patch. Of course, an ideal free distribution is out of the question since the vast majority of ducks receive nothing, because they are incarcerated at the transit patch.

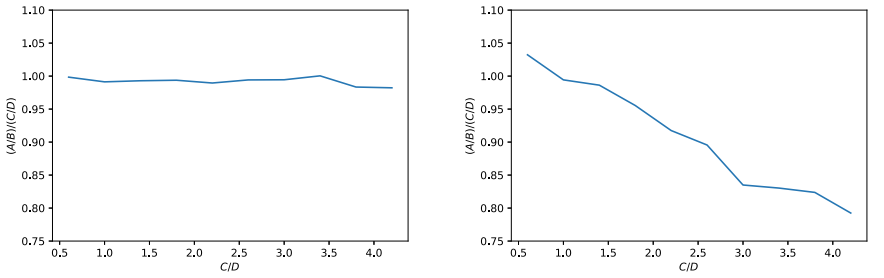
This is a largely unsurprising result, given the basic mechanics of the decision making by the foragers and the relative bread rates. A more interesting scenario is suggested by giving the ducks some commitment to their current patch. We can change the probability that they will switch patches once they perceive an asymmetry in the relative distribution of food. Suppose that the probability that a duck will switch patches is 0.2, that is they will transit only twenty percent of the time once they perceive the asymmetry. On this experiment, then, the ducks take time to move from food source to food source but are reluctant to move.[6] That is, the conditions—number of agents, food rates from the sources, and so on—are the same as the previous experiment but the probability of a duck moving when it perceives an asymmetry between the patches is just 0.2. In a typical run the results are that the ducks at the more fecund patch are getting 0.131 units of bread per time at the patch, while ducks at the less fecund patch are getting 0.102 units of bread per time. After stopping at the default of 3,500 time steps, there are 114 ducks at the more fecund patch compared to 49 ducks at the less fecund patch and 87 ducks in transit.

The results are quite different from those of the pure IFD model, due to the "friction" created by the temporal constraints on movement. The underlying decision rule the ducks are using is the same as in the pure IFD model, but the physical constraints of the world prevent them from achieving a pure ideal free distribution. Notice that the count of the number of ducks at the more profitable patch is depressed compared to the pure IFD model.

The fact that there are fewer ducks at the more profitable patch raises the bread per duck per unit time at that patch. How is this to be explained? In the mathematical model of the IFD the assumption is made that time is continuous, with upshot that each duck makes its decision at an infinitesimal interval that is unique to the duck. In our agent-based model, time is discrete and ducks at sources make their decisions independently whether to move during a finite interval and then move en masse, if they move at all. In consequence, if there's a more attractive source, while it may have more ducks in residence, more ducks will move if the less attractive source is comparatively under-populated. Conversely, at the less productive source, if it becomes less attractive compared to the other source, fewer ducks will move at any given time step. The outcome of these considerations is that the more attractive source will be relatively underpopulated, yielding more bread per duck at that source. This is exactly what we have seen, above.

In order to investigate this, we developed BehaviorSpace experiments to examine the relationship between bread rates at the sources and the number of ducks at the sources. If the distribution were a pure Ideal Free one, the ratio of the number of ducks at the higher yield patch to the number of ducks at the lower yield patch divided

[6] We tried this experiment with the pure IFD model and it made no difference.



(a) Ideal Free Distribution (b) Ideal Free Model with Transit Costs

Fig. 2 Behavioral comparison with and without transit costs. A is the count of ducks at the more fecund food source; B is the count of ducks at the less fecund food source; C is the bread rate at the more fecund food source; D is the bread rate at the less fecund food source

by the ratio of the rate of bread distribution at the higher patch to the rate of bread distribution at the lower yield patch should be equal to one. This would be what happens in the pure IFD case where the ducks distribute themselves in such a way that they all get the same amount of bread.
Let:

(a) A is the count of ducks at the more fecund food source;
(b) B is the count of ducks at the less fecund food source;
(c) C is the bread rate at the more fecund food source;
(d) D is the bread rate at the less fecund food source.

The graphs in Fig. 2 compare C/D (the x-axis) with $(A/B)/(C/D)$ (the y-axis). Under IFD the latter quantity should always be equal to 1, as happens in Fig. 2a, the Pure IFD model where transit costs are 0; there is some slight statistical noise but the ducks always arrange themselves as expected with the ducks at the two food sources getting the same amount of bread. When transit from one food source to another involves a real cost, as in Fig. 2b, we see that the distribution of bread is not equal across the food sources; in fact, since the ratio is less than 1, there are fewer ducks at the more fecund food source than we would expect under the IFD. In consequence, because there is a smaller population at the more fecund source, the bread per duck should be higher than what would be found at the less fecund source; this is what we observe. In fact, as a result of there being fewer ducks at the more fecund source, the bread per duck at that source is higher than we would expect under an ideal free distribution.

One way to think about the above is that the agent-based model occurs in discrete time so that, in effect, the agents make their decisions about moving in parallel; the agents have no information about what other ducks decide at that time step. Furthermore, once a duck is in transit, it gets no information about the ducks who were already in transit, the ducks who went in transit with the duck, and the ducks who go into transit while the duck itself is transiting; in other words, the cost of transit is to deprive the transitee of information which should condition its choice

of destination. Even if a duck in transit got information about the real distribution of ducks, the model provides no way for the duck to use that information while in transit. The pure IFD, on the other hand, takes place in continuous time (and with instantaneous movement) so the ducks always have perfect information about the other ducks and the expected returns of the food sources. In other words, discrete time imposes a form of bounded rationality by depriving the ducks of information.

The above results used a certain amount of "stickiness" on the part of the ducks; their probability of moving at any time step is 0.2. If they move immediately when they are dissatisfied with their current food source, that is the probability of moving is 1, then the results are much more chaotic. In essence, the vast majority of ducks are always in transit, landing at a food source only long enough to become dissatisfied and go back into transit. The need for reticence in moving is another illustration of the real costs of transit and reflects something that should be taken into account when analyzing real foraging data.

In order to make comparisons more systematically we conducted two BehaviorSpace experiments, each with 10 replications. In the first experiment, we used a Pure IFD configuration (`transit?` switched to off), varying the bread rate at food source A and keeping all else constant. The results are shown in Table 1. The second experiment duplicated the first, except that `transit?` was switched to on, violating the Pure IFD assumptions of no travel cost. See Table 2. Comparing the two experiments confirms what we reported for prototypical experiments of the two cases. In the Pure IFD computational model, the predicted Pure IFD obtains. As comparative bread rates change, bread per duck is equal across the two food sources, and the fraction of ducks at each food source is also as predicted by the IFD theory. Contrariwise, in the `transit?` on configuration (with transit cost > 0) and when the fecundity level of food source A is higher than B: (i) the bread per duck systematically varies between the two food sources with ducks at source A receiving more food per unit time than ducks at source B; and (ii) the number of ducks present at site A is lower than it should be under Pure IFD. The effects are robust and systematic.

When IFD conditions obtain, no duck can unilaterally do better by changing its policy for moving between food sources. This is to say that Nash equilibrium conditions rule. What happens when movement has costs? It would seem that the duck's movement policy may no longer be a Nash equilibrium policy. To investigate this conjecture we conducted another experiment, in which two ducks, one at source A and one at Source B, were "pinned." That is, two ducks were randomly picked at initialization and made to stay at one or the other of the food sources. Everything else is as set for the experiment of Table 2.

Table 3 shows the difference between the bread collected by a duck pinned at site A, whose rate of bread distribution varies from 3 units to 20 units of bread per time step, a duck pinned at site B, whose rate of bread distribution is held constant at 5 units of bread per time step, and ducks that can freely travel between A and B. The rates are compared for when temporal penalties for travel are disabled ("False") and when temporal penalties, held at 5 units of time, are enabled ("True"). We can see that the pinned ducks and the free ducks get roughly the same amount of bread when travel penalties are disabled, the Ideal Free Distribution. When temporal penalties are

Table 1 Bread rate at site B is a constant 5 units per time step. Bread rate at site A varies from 3 to 20 units per time step. Under IFD, Population at A divided by Population at B should be equal to Rate A divided by Rate B, which it is to a close approximation. Also under IFD Bread Per Duck (BPD A) should be equal to BPD B, which it is

Rate A	Population at A	Population at B	BPD A	BPD B
3	93.533	156.467	0.032	0.032
5	124.467	125.533	0.040	0.040
10	166.200	83.800	0.060	0.060
15	187.100	62.900	0.080	0.079
20	199.267	50.733	0.100	0.099

Table 2 Bread rate at site B is a constant 5 units per time step. Bread rate at site A varies from 3 to 20 units per time step. Under IFD, Population at A divided by Population at B should be equal to Rate A divided by Rate B, but in our departure from IFD due to transit time this equivalence fails. Also under IFD Bread Per Duck (BPD A) should be equal to BPD B, but transit time costs destroy this equivalence

Rate A	Population at A	Population at B	BPD A	BPD B
3	61.633	96.867	0.049	0.052
5	78.800	79.133	0.064	0.063
10	105.367	58.067	0.095	0.087
15	115.033	47.000	0.131	0.111
20	124.567	38.633	0.162	0.138

levied, we see that the pinned ducks, who cannot travel, do systematically better than the ducks that are allowed to transit. This suggests that the best strategy (the Nash equilibrium) is actually not to travel, if travel takes time. More precisely, pinning is shown to be a profitable response when all but one duck, at the other site, are transiting in the usual way. We also see that the pinned duck at site A, the site with generally higher rates of bread distribution, does better than the duck pinned at site B. Table 2 also shows this effect, we think for the same reason. In general, the more prosperous site is proportionately under-represented, although it has more ducks at it; this is because the ducks make their decisions in parallel so that more populated sites will tend to look less attractive. The underlying data shows that the number of ducks at site A is systematically less than the number at site B, when the rate of bread distribution at A is greater that that at B. The odds of this happening by chance is less that one in a billion. The Ideal Free Distribution assumes that ducks make their decisions sequentially and that transit time is zero.

6 Discussion of the Models

With any very idealized model there always lurks the danger of over abstraction and failure of robustness in the fact of plausible and realistic changes to the model. We have seen this phenomenon at play with respect to travel time and ideal free distributions. It is hardly surprising that idealized models will fail in the absence of truth in their assumptions. The most interesting subsequent questions are in regard to the nature and magnitude of the changes in model behavior in response to changes in model assumptions. To take one small example, with pinning as in Table 3 and focusing on the last row, we see that the rewards for the pinned ducks greatly exceed those of the unpinned ducks following what would be a Nash equilibrium strategy under IFD conditions. The pinned duck at source B, the less fecund of the two sources, gets 27% more food than the average unpinned duck (281/299), while the pinned duck at A gets a whopping 60% more (478/299). Exploring these kinds of effects and confronting them with data lies at the heart of strategic modeling and analysis, and the subject of our book, decision games.

The study of foraging is a rich and flourishing subfield of ecology, affording ample vistas of opportunity for modeling and data collection. Our treatment here should be taken as a point of departure into a great realm of modeling possibilities. We have demonstrated that agent-based modeling can both duplicate idealized models and serve as a basis for investigations that relax idealizations in the direction of verisimilitude.

Table 3 B source is constant at 5; A source is variable. Comparing the average bread collected by a duck pinned at A, a duck pinned at B, and unpinned ducks. Results are the means of 30 replications for each setting

(Variable Rate at A, Transit)	Duck pinned at A	Duck pinned at B	Average Unpinned
(3, False)	96.121	95.934	96.000
(3, True)	148.402	153.614	95.556
(5, False)	120.483	119.525	120.000
(5, True)	188.659	187.476	119.451
(10, False)	180.390	179.258	180.001
(10, True)	289.287	269.447	179.199
(15, False)	240.396	238.960	240.003
(15, True)	385.391	329.386	239.053
(20, False)	301.332	295.116	300.014
(20, True)	477.940	381.244	298.955

7 Future Work

There is a great deal of work that remains to be done here. Due to space limitations, we have not been able to adequately cover recent work that is quite relevant to this topic, for example [15] (we are grateful to an anonymous referee for pointing this paper out to us). There are, of course, many other papers that would have to be systematically reviewed.

Among the most interesting items for future research, identified by anonymous referees and other discussants, are the following:

- How is the order of events determined? Does order matter?
- Can the ABM model be reduced to a stochastic differential equations model? What advantages does an ABM provide versus other approaches, particularly in the pure model?
- What happens when multiple food sources of similar fecundity are introduced? What if they have different distributions of fecundity values? Would your results remain robust then?

8 Bibliographic Note

In the interest of conserving space, we note the following works related to ideal free distributions related to this paper. [18] is an early landmark development of optimal foraging theory, which focuses on individual, asocial behavior. Strategic or game theoretic modeling of foraging, which came to be called social foraging, may be said to begin with [2, 3, 5] and [4]. [7, 10, 11, 19] address important issues as the field developed. [6] synthesizes and develops social foraging models. [17] surveys the field of foraging (social and asocial) 20 years after [18]. [1] is an accessible treatment of foraging models applied to human hunter-gatherers.

Acknowledgements We are grateful for very useful comments from three anonymous referees for the conference.

References

1. Bettinger, R.L.: Hunter-Gatherer Foraging: Five Simple Models. Eliot Werner Publications Inc., Clinton Corners, NY (2009)
2. Fretwell, S., Lucas, H.: On territorial behavior and other factors influencing habitat distribution in birds: I. theoretical development. Acta Biotheoretica **19**(1), 1–36 (1969)
3. Fretwell, S.D.: On territorial behavior and other factors influencing habitat distribution in birds. III. Breeding success in a local population of field sparrows. Acta Biotheoretica **19**, 45–52 (1969)

4. Fretwell, S.D.: Populations in a seasonal environment. In: Number 5 in Monographs in Population Biology. Princeton University Press (1972)
5. Fretwelll, S.D., Calver, J.S.: On territorial behavior and other factors influencing habit disribution in birds: II. Sex ratio variation in the Dickcissel. Acta Biotheoretica **19**, 37–44 (1969)
6. Giraldeau, L.-A., Caraco, T.: Social Foraging Theory. Princeton University Press, Princeton (2000)
7. Harper, D.: Competitive foraging in mallards: "Ideal Free" ducks. Anim. Behav. **30**, 575–584 (1982)
8. Herrnstein, R.J.: The Matching Law: Papers in Psychology and Economics. Harvard University Press, Cambridge, MA (1997)
9. Jeffrey, R.C.: The Logic of Decision, 2nd edn. University of Chicago Press, Chicago, Il (1983)
10. Kennedy, M., Gray, R.D.: Can ecological theory predict the distribution of foraging animals? a critical analysis of experiments on the ideal free distribution. Oikos **68**(1), 158–166 (1993). Publisher: [Nordic Society Oikos, Wiley]
11. Kraft, J.R., Baum, W.M., Burge, M.J.: Group choice and individual choices: modeling human social behavior with the ideal free distribution. Behav. Process. **57**, 227–240 (2002)
12. Krantz, D.H., Kunreuther, H.: Goals and plans in decision making. Judgm. Decis. Mak. **2**(3), 34 (2007)
13. Krantz, D.H., Luce, R.D., Suppes, P., Tversky, A.: Foundations of Measurement, vol. I, Additive and Polynomial Representations. Academic, New York, NY (1971)
14. Luce, R.D., Raiffa, H.: Games and Decisions: Introduction and Critical Survey. Wiley, New York, NY (1957)
15. Olszewski, W.: Preferences and information processing under vague information. J. Math. Econ. **94**, 1024–61 (2021)
16. Savage, L.J.: The Foundations of Statistics. Wiley, New York, NY (1954)
17. Stephens, D.W., Brown, J.S., Ydenberg, R.C. (eds.): Foraging: Behavior and Ecology. The University of Chicago Press, Chicago (2007)
18. Stephens, D.W., Krebs, H.R.: Foraging Theory. Princeton University Press, Princeton (1986)
19. Tregenza, T.: Building on the Ideal Free Distribution. Volume 26 of Advances in Ecological Research, pp. 253–307. Academic (1995). ISSN: 0065-2504
20. Warren, C., McGraw, A.P., Van Boven, L.: Values and preferences: defining preference construction. WIREs Cogn. Sci. **2**(2):193–205 (2011). _eprint: https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/wcs.98

Entropy-Based Heuristic Approach For The Quantum-Like Generalization of Social Contagion

Ece Çiğdem Mutlu and Ozlem Ozmen Garibay

Abstract Social contagion modeling has recently attracted a great deal of attention from researchers due to its wide range of applications in network science, multi-agent systems, information science, and marketing. Since there are reinforcement effects in social contagion systems, it is necessary to consider the complexity of individuals in the system in order to understand this phenomenon. This complexity that stems from the heterogeneity of individuals and the uncertainty in their decision-making process caused the utilization of more complex social contagion modeling such as quantum-like approaches. Although these approaches are demonstrated to be able to portray this complexity and better model the social contagion process, the interference term in these models is hard to predict, causing their application very limited. To address this problem, we propose a belief-entropy-based heuristic approach to predict interference effect in quantum-like generalization of social contagion. Based on simulations of uncorrelated random regular networks (RRNs) using the proposed approach, we concluded that belief entropy is useful for detecting interference in quantum-like generalizations of social contagion models. These results should lead to increased use of quantum social contagion models in any application area without having to deal with calibration issues or time constraints.

1 Introduction

By understanding and improving the modeling of contagion dynamics in complex networks, researchers can shed light on the mechanisms that underlie the spread of diseases, microfinance activities, information, harmful emotions, and technology adoption. Thus, in addition to offering us more effective anti-pathogen strategies

E. Ç. Mutlu · O. O. Garibay (✉)
Department of Industrial Engineering and Management Systems, University of Central Florida, Florida, FL 32816, USA
e-mail: ozlem@ucf.edu

E. Ç. Mutlu
e-mail: ece.mutlu@ucf.edu

Z. Yang and S. Núñez-Corrales (eds.), *Proceedings of the 2022 Conference of The Computational Social Science Society of the Americas*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-031-37553-8_15

during epidemics, these developments also give us new theoretical foundations for predicting social behaviors, and even mitigates the spread of false information within social systems. In the context of contagion, researchers have classified these spreading dynamics in different disciplines into two main categories: (i) biological, or (ii) social contagion. Though dynamics of these spreading mechanisms have very common properties, social contagion is known for having its own inherent characteristic, known as the social reinforcement effect [21, 31, 32], compared to biological spreading. In social contagion, the reinforcement effect changes the simple mechanism of biological contagion, which assumes even one single activated source will suffice for transmission, into a more complex contagion mechanism. To describe this complexity in contagion dynamics, Markovian processes are generally used such as threshold-driven approaches, whereby adoption occurs only if a specific portion of neighboring nodes has already been adopted, contrary to biological spreading. As the dynamics of social contagion differ from epidemic contagion in terms of their complexity, and a variety of disciplines such as marketing and information science are entangled, we can argue that understanding the dynamics of social contagion is substantial and unfinished.

Traditionally, mathematical modeling of social contagion is carried out by one of the three commonly-known approaches: threshold-driven, cascading, and compartmental approaches. In Granovetter's groundbreaking study [6], where a mathematical model of social contagion was first introduced, he proposed a linear threshold model based on the assumption that individuals' behavior in a network can be influenced by their neighbors. In this receiver-centric model, individuals adopt a behavior only if a certain fraction of neighbors have already adopted the behavior. This model not only ignores the time delay between adoption and spreading, but it also has a strong but unrealistic assumption that each adopted individual is willing to spread behavior in the network. Later, Goldenberg [5] introduced a sender-centric model, called the independent cascade model, in which each adopted node has a single chance to influence one of its susceptible neighbors. Since this model is limited by the possibility of one successful influence at each time step and ignores the social-reinforcement effect, it also falls short in comprehensive modeling of social contagion dynamics. Recently, inspired by epidemic models, one of the most commonly used methods in the literature of social contagion studies is the message passing approach [10], in which individuals within the target population (or network) are divided into mutually exclusive compartments based on their current status and their future status at any time can be predicted based on the predefined rate of contact between compartments and their certain transition rates. As opposed to the conventional compartmental models, the reinforcement effect is also included with the existence of a threshold value for individuals to adopt the behavior. Therefore, the message passing approach is considered a non-Markovian process, which makes it more realistic in the application of real-world complex contagions.

Any of these approaches in the analysis of social contagion faces the challenge of modeling complexity of individuals, which arises either as a consequence of heterogeneity of the individual's threshold for adoption or the uncertainty in their decision-making process. Although the former is considered in the recent studies by utilizing

more complex threshold distributions [4, 9, 31, 34] rather than uniform, the latter has yet been widely addressed to the best of our knowledge. On the other hand, individuals may show ambiguous characteristics during behavior (information, emotion or technology) exchange [18, 24, 29, 33] due to subconscious feelings and subjective biases [16, 27] after they interact with each other. This complexity in decision-making studies has been addressed with numerous quantum-like approaches [1, 2, 7, 13, 17, 33] to explain the corresponding irrationality and existing paradoxes and fallacies. In a very recent study [19], it is demonstrated that quantum-like approaches yield better performances also in the modeling of social contagion dynamics since interference effects in these models leverage the extant social contagion analyses and thus better model its nonlinear dynamics even on critical transmission probabilities. Whereas, the interference term brings an additional complexity that decreases the time-efficiency of these models. Additionally, this effect can be determined only by calibration with a real dataset, causing its application harder. In this study, we propose a belief entropy-based heuristic for the determination of interference effect in quantum-like generalization of social contagion. Our results show that belief entropy is a promising approach in the measurement of uncertainty in quantum-like social contagion analyses and we believe that the results of this study will bring its application area more diverse in different disciplines.

2 Background

2.1 Quantum Social Contagion

A complex network, $G\langle V, E\rangle$, is graph in which set of vertices (nodes) ($V = \{v_1, v_2, ..., v_n \mid n \in N\}$) are connected with each other with edges (links) ($E_{v_i,v_j} = (v_i, v_j)$ where $(i, j \in N; i \neq j)$. In a network where each entity corresponds to a different individual, the social contagion occurs as a result of interactions of connected individuals. To exemplify the social contagion mechanism in this study, we integrate a quantum-like point of view to the classical message-passing approach [10] that generalizes the well-known susceptible-adopted-recovered (SAR) model, to fully describe the mechanisms of information (or behavior) spreading on a complex network. In this approach, each individuals of the network of N nodes and a degree distribution $P(k)$ falls into one of three states: *susceptible*, *adopted* and *recovered*. These states represent:

- An individual in a *susceptible* state (S) does not adopt the behavior yet,
- An individual in an *adopted* state observed the behavior and adopted it already and tries to transmit it to his *susceptible* neighbors,
- An individual in a *recovered* state adopted the behavior once but lose interest and will not further participate in spreading.

During social contagion, each *adopted* individual (A) tries to influence his their *susceptible* neighbors with a probability λ at each time step. Once the influence

successful, influenced *susceptible* individual updates his cumulative units of opinion, i.e. $m \Rightarrow m + 1$. Simultaneously, *adopted* individuals may lose their interest in spreading and become *recovered* with a probability γ. Furthermore, a small number of ρ_0 is defined to assign the fraction of *adopted* individuals in the network, and a steady-state is reached if all individuals in the network become *recovered* since there is no chance for individuals to change their current states afterward.

2.2 Quantum Probabilistic Approach

Regardless of using the classical or quantum approach, probabilistic approaches aim to assign marginal probabilities to each event in the sample space. More fundamentally, the two approaches differ in their representation. The classical approach uses set-theoretic representation and its sample space is defined as a set of possible events, e.g. $\{m_0, m_1\}$. On the other hand, the quantum approach uses vector space representation and its sample space is a plane space spanned by the orthogonal basis vectors, e.g. $|m_0\rangle$ and $|m_1\rangle$. This *braket* notation is also called Dirac notation and is commonly used in the representation of quantum states. The differences in classical and quantum sample space representations mainly stem from the assumptions that are used in these approaches. The classic probabilistic theory assumes a sample space in which the outcome of the events are mutually exclusive, i.e. $\Omega = \{m_0, m_1\}$. In quantum probabilistic theory, on the other hand, events are modeled as subspaces of a Hilbert space in which each orthogonal basis vector corresponds to an elementary outcome, i.e. $|M\rangle = m_0|0\rangle + m_1|1\rangle$, where

$$M = \begin{bmatrix} m_0 \\ m_1 \end{bmatrix}$$

An inner product of vector M is obtained with the multiplication of $|M\rangle$ with its complex conjugate ($|M\rangle^* = \langle M|$) as follows:

$$\begin{aligned} \langle M|M\rangle &= \begin{pmatrix} m_0 \\ m_1 \end{pmatrix} \begin{pmatrix} m_0 & m_1 \end{pmatrix} \\ &= \begin{pmatrix} |m_0|^2 & |m_0||m_1|^* \\ |m_1||m_0|^* & |m_1|^2 \end{pmatrix} = \begin{pmatrix} \psi_{m_0}\psi_{m_0}{}^* & \psi_{m_0}\psi_{m_1}{}^* \\ \psi_{m_1}\psi_{m_0}{}^* & \psi_{m_1}\psi_{m_1}{}^* \end{pmatrix} \end{aligned} \tag{1}$$

As we are familiar with, representing the set of outcomes as mutually exclusive events in classical theory enables us to easily define more complex events that require intersection, union, and/or distribution of individual events. In general, the conjunction (intersection) of two independent events is represented by $(m_0 \cap m_1)$ and the disjunction (union) is by $(m_0 \cup m_1)$. Furthermore, a distributive axiom is also applicable in classical theory, i.e. $m_0 \cap (m_1 \cup m_2) = (m_0 \cap m_1) \cup (m_0 \cap m_2)$, since it obeys a set theory.

In the quantum approach, on the other hand, mutually exclusive events are represented by orthonormal basis vectors contained in the Hilbert space. This geometric approach enables us to define a superposition state which comprises the occurrence of different events at the same time. The superposition state of happening both events of m_0 and m_1 and computed as follows:

$$|S\rangle = \frac{e^{\theta_{m_0}}}{\sqrt{2}}|m_0\rangle + \frac{e^{\theta_{m_1}}}{\sqrt{2}}|m_1\rangle = \frac{e^{\theta_{m_0}}}{\sqrt{2}}\psi_{m_0} + \frac{e^{\theta_{m_1}}}{\sqrt{2}}\psi_{m_1} \tag{2}$$

Here, the exponential term ($e^{i\theta_{m_0}}$) is called global phase factor of the quantum probability amplitude. The probability ($Pr(m_0)$) is related with a quantum probability amplitude ($e^{i\theta_{m_0}}|m_0\rangle$) which corresponds to the amplitude of a wave function, and this relation to the classical probability is obtained by multiplying this amplitude with its complex conjugate, i.e. $|e^{i\theta_{m_0}}|^2 = e^{i\theta_{m_0}}|m_0\rangle e^{-i\theta_{m_0}}|m_0\rangle^*$. This connection is obtained via Born's rule as follows:

$$Pr(m_0) = |e^{i\theta_{m_0}}\psi_{m_0}|^2 \tag{3}$$

Although the result of an individual event probability in the classical probability theory converges to that in the quantum approach, the computation of the union of mutually exclusive events differs in these two methods. The quantum-like approach yields an extra term, "*interference effect*", which does not exist in classical probability theory. To illustrate, suppose that we aim to obtain the union of three mutually exclusive events by using classical probability formula, which is given by:

$$Pr(A \cup B \cup C) = Pr(A) + Pr(B) + Pr(C) \tag{4}$$

The quantum counterpart of the classical probability of the union of three mutually exclusive events is obtained by using Born's rule in (3):

$$\begin{aligned} Pr(A \cup B \cup C) &= |e^{i\theta_A}\psi_A + e^{i\theta_B}\psi_B + e^{i\theta_C}\psi_C|^2 \\ &= e^{i\theta_A}\psi_A.e^{-i\theta_A}\psi_A + e^{i\theta_A}\psi_A.e^{-i\theta_B}\psi_B \\ &+ e^{i\theta_A}\psi_A.e^{-i\theta_C}\psi_C + e^{i\theta_B}\psi_B.e^{-i\theta_A}\psi_A \\ &+ e^{i\theta_B}\psi_B.e^{-i\theta_B}\psi_B + e^{i\theta_B}\psi_B.e^{-i\theta_C}\psi_C \\ &+ e^{i\theta_C}\psi_C.e^{-i\theta_A}\psi_A + e^{i\theta_C}\psi_C.e^{-i\theta_B}\psi_B \\ &+ e^{i\theta_C}\psi_C.e^{-i\theta_C}\psi_C \end{aligned} \tag{5}$$

Knowing that,

$$cos(\theta_1 - \theta_2) = \frac{e^{\theta_1-\theta_2} + e^{-\theta_1+\theta_2}}{2} \tag{6}$$

Equation 5 reduces to:

$$\begin{aligned} Pr(A \cup B \cup C) &= |\psi_A|^2 + |\psi_B|^2 + |\psi_C|^2 \\ &+ 2\big(|\psi_A||\psi_B|cos(\theta_A - \theta_B) + |\psi_A||\psi_C|cos(\theta_A - \theta_C) \\ &+ |\psi_B||\psi_C|cos(\theta_B - \theta_C)\big) \end{aligned} \tag{7}$$

The additional terms in (7) compared to (4) are called as "*interference terms*" which does not exist in classical probability theory [11, 12, 14, 17, 19].

2.3 Entropy-Based Uncertainty Measures of Stochastic Processes

The second law of thermodynamics states that the total entropy of an isolated system (the thermal energy per unit temperature that is unavailable for doing useful work) can never decrease. This concept of entropy was introduced by Rudolf Clausius in 1865 [23] as a term in the field of thermodynamics, later adapted into statistical physics and information theory to characterize the uncertain, ambiguous, and disordered behavior of stochastic processes [20]. After Clasius' definition of entropy as a thermodynamic concept, Shannon [26, 26] argued that this concept can be extended into different disciplines due to its probabilistic nature in defining the randomness of stochastic processes, and proposed Shannon entropy [25] as an uncertainty measure. Later, the entropy measure proposed by Renyi et al. [22], called Renyi entropy, has been applied in diverse areas including quantum information, information theory, and fractal theory. Another non-extensive measure of Tsallis entropy which is an extension of Boltzman entropy [28] has also gained a lot of attention. Recently, a new entropy named Deng entropy [3] has been proposed to solve the uncertainty of the stochastic processes based on the given evidence. [15] describes the similarities and differences of these entropy measures to better explain their use areas.

For a random variable X over a probability space Ω, Shannon entropy is defined for continuous and discrete variables as follows, respectively:

$$\begin{aligned} S(X) &= -\int_{\Omega} p(x)log_2(p(x))dx \\ S(X) &= -\sum_{x \in \Omega} p(x)log_2(p(x)) \end{aligned} \tag{8}$$

where $p(x)$ denotes the probability distribution. Although this measure performs well in the existence of finite storage capacity of transmitting channel in communication, it falls short in infinite storage capacity. To address this, Renyi [8] proposed a new measure, called Renyi entropy, which is defined for discrete variables as follows:

$$S_\alpha(X) = \frac{1}{1-\alpha} ln\left(\sum_{k=1}^{n} p_k^\alpha\right) \tag{9}$$

where $\alpha \neq 1$ and $\alpha \geq 0$. When the order of α equals to 1, Renyi entropy degenerates into Shannon entropy.

On the contrary of Shannon and Renyi entropy measures which yields exponential equilibrium distribution, Tsallis extended these definitions and proposed a new entropy measure which can be used with any non-negative real number, which yields a power-law equilibrium distribution. The formula of Tsallis entropy for a non-negative real number q is as follows:

$$S_q(X) = \frac{1 - \sum_{i=1}^{n} p_i^\alpha}{q-1} \tag{10}$$

where $q \neq 1$ and $q \geq 0$. When the order of q equals to 1, Tsallis entropy degenerates into Shannon entropy.

Belief entropy, named as Deng entropy, on the other hand, can be described as a combination of a measure of total non-specificity in the basic probability assignment indicating the degree of belief in $A_i \in P(X)$ and a measure of discord of the mass function among various focal elements. Its formula is:

$$H_d = -\sum_i m(A_i) ln \frac{m(A_i)}{2^{|A_i|-1|} - 1} \tag{11}$$

3 Methodology

In this study, we aim to mathematically describe quantum-like social contagion and show the utilization of belief entropy as a heuristic of interference term inside. Next sections will give detail explanation for these concepts.

3.1 Edge-Based Compartmental Theory

Here, we employ an edge-based compartmental theory to understand the dynamics of the quantum social contagion approach inspired by numerous studies [30, 31, 34]. For this, suppose that there are two individuals who are connected to each other in the network, e.g., u and v, $E_{u,v} \neq 0$. Among these, u, $(u \in V)$ is an individual who is in the susceptible state. Let $\theta(t)$ is the probability that the individual v has not transmitted information to the individual u by time t, the quantum probability of the same event can be calculated by using Born's rule in (3) as follows:

$$|\sqrt{\theta(t)}e^{\theta(t)}|^2 = (\sqrt{\theta(t)}e^{\theta(t)}).(\sqrt{\theta(t)}e^{-\theta(t)}) = \theta(t) \tag{12}$$

Then the probability that individual u with degree k_u has received m pieces information from his distinct neighbors by time t will be binomially distributed and computed as:

$$\tau_m(k_u, t) = \binom{k_u}{m}\theta(t)^{(k_u-m)}(1-\theta(t))^m \tag{13}$$

Individual u's state depend on the number of information he has obtained from his neighbors. If he receives enough pieces of information from his distinct neighbors to exceed his threshold (ϕ_u), i.e. $m \geq \phi_u$, then he will adopt the information and try to transmit it to his susceptible neighbors in the next time step. Otherwise, he will keep his susceptible state in the next time step. Thus, the probability of individual u with degree k_u being susceptible is:

$$\begin{aligned} s_u(k_u, t) &= \sum_{\phi_u} F(\phi_u) \sum_{m=0}^{\phi_u - 1} \tau_m(k_u, t) \\ &= \sum_{\phi_u} F(\phi_u) \sum_{m=0}^{\phi_u - 1} \binom{k_u}{m}\theta(t)^{(k_u-m)}(1-\theta(t))^m \end{aligned} \tag{14}$$

where $F(\phi_u)$ denotes the information adoption threshold function. To consider the heterogeneity among individuals, we represented $F(\phi_u)$ as a binomial distribution in which individuals may have either a relatively lower threshold ($T_A = 1$) with probability p, or a relatively higher threshold ($T_B > 1$) with probability $1 - p$.

$$F(\phi_u) = \begin{cases} T_A, & \text{with probability } p \\ T_B, & \text{with probability 1-}p \end{cases} \tag{15}$$

Combining (14) and (15), we can obtain the fraction of susceptible individuals at time t as follows:

$$\begin{aligned} S(t) &= \sum_{k_u} P(k_u) s_u(k_u, t) \\ &= \sum_{k_u} P(k_u) \sum_{\phi_u} F(\phi_u) \sum_{m=0}^{\phi_u - 1} \binom{k_u}{m}\theta(t)^{(k_u-m)}(1-\theta(t))^m \\ &= \sum_{k_u} P(k_u) \left[p\theta(t)^{(k_u)} + (1-p) \sum_{m=0}^{T_B - 1} \binom{k_u}{m}\theta(t)^{(k_u-m)}(1-\theta(t))^m \right] \end{aligned} \tag{16}$$

A similar strategy can be applied to calculate the probability of individual v with degree k_v being susceptible state. It should be noted that individual v's state is

unknown; however, it is known that individual u is in susceptible state and cannot transmit information. Thus, the individual v can receive information from his $k_v - 1$ distinct neighbors. Taking all possible values of receiving m pieces of cumulative information and ϕ_v into consideration, we obtain:

$$\begin{aligned} s_v(k_v, t) &= \sum_{\phi_v} F(\phi_v) \sum_{m=0}^{\phi_v - 1} \tau_m(k_v, t) \\ &= \sum_{\phi_v} F(\phi_v) \sum_{m=0}^{\phi_v - 1} \binom{k_v - 1}{m} \theta(t)^{(k_v - m - 1)} (1 - \theta(t))^m \\ &= p\theta(t)^{(k_v - 1)} + (1 - p) \sum_{m=0}^{T_B - 1} \binom{k_v}{m} \theta(t)^{(k_v - m - 1)} (1 - \theta(t))^m \end{aligned} \tag{17}$$

In the message-passing approach, adopted individuals may either try to transmit information or lose their interest in the transmission process and move into the recovered state. Thus, the following set of ordinary differential equations (ODEs) define the time dependence of the individuals in each compartment in the system described above.

$$\begin{aligned} \frac{dA(t)}{dt} &= -\frac{dS(t)}{dt} - \gamma A(t) \\ \frac{dR(t)}{dt} &= \gamma A(t) \end{aligned} \tag{18}$$

In edge-based compartmental theory, we have not made any assumption about the state of individual v; therefore, $\theta(t)$ may consist of three possible outcomes which are mutually exclusive in classical approach:

$$\theta(t) = \xi_S(t) + \xi_A(t) + \xi_R(t) \tag{19}$$

where $\xi_S(t)$, $\xi_A(t)$, and $\xi_R(t)$ represent the probability that a neighbor v in the susceptible, adopted, and recovered states, respectively by time t. To employ quantum probability rules, we can use Born's rule in (3) and write counterpart of (19) as follows (see (7)):

$$\begin{aligned} \theta(t) &= |e^{i\theta\psi_{\xi_S(t)}} + e^{i\theta\psi_{\xi_A(t)}} + e^{i\theta\psi_{\xi_R(t)}}|^2 \\ &= |\psi_{\xi_S(t)}|^2 + |\psi_{\xi_A(t)}|^2 + |\psi_{\xi_R(t)}|^2 + 2\Big[|\psi_{\xi_S(t)}||\psi_{\xi_A(t)}|cos(\theta_{\xi_S(t)} - \theta_{\xi_A(t)}) \\ &+ |\psi_{\xi_S(t)}||\psi_{\xi_R(t)}|cos(\theta_{\xi_S(t)} - \theta_{\xi_R(t)}) + |\psi_{\xi_A(t)}||\psi_{\xi_R(t)}|cos(\theta_{\xi_A(t)} - \theta_{\xi_R(t)})\Big] \end{aligned} \tag{20}$$

Here, the amplitude $|\psi_{\xi_S(t)}|^2$ refers to $P(\xi_S(t))$, $|\psi_{\xi_A(t)}|^2$ to $P(\xi_A(t))$ and $|\psi_{\xi_R(t)}|^2$ to $P(\xi_R(t))$. The angle $\theta_{\xi_S(t)} - \theta_{\xi_A(t)}$ corresponds to the phase of the inner product

between $|\xi_S(t)|$ and $|\xi_A(t)|$. Note that there is no direct transition from susceptible state to recovered state, so $cos(\theta_{\xi_S(t)} - \theta_{\xi_R(t)})$ will be equal to 0. By recalling inverse Born's rule again, we can finalize the relation above as:

$$\begin{aligned}\theta(t) &= \xi_S(t) + \xi_A(t) + \xi_R(t) \\ &+ \sqrt{\xi_S(t)\xi_A(t)}cos(\theta_{\xi_S(t)} - \theta_{\xi_A(t)}) + \sqrt{\xi_A(t)\xi_R(t)}cos(\theta_{\xi_A(t)} - \theta_{\xi_R(t)})\end{aligned} \tag{21}$$

Herein, the additional terms compared to (19) are called as interference terms that does not exist in classical probability theory. Later, we draw from statistical network science to make the connection between these two individuals u and v. In the case of the existence of an uncorrelated network, the probability of an edge connecting individual v with a degree k_v to one of its neighbors, e.g., individual u with degree k_u, is equal to $k_v P(k_v)/\langle k \rangle$, where $\langle k \rangle$ is the mean degree. Thus, it can be obtained that:

$$\begin{aligned}\xi_S(t) &= \frac{\sum_{k_v} k_v P(k_v) s_v(k_v, t)}{\langle k \rangle} \\ &= \frac{\sum_{k_v} k_v P(k_v) \sum_{\phi_v} F(\phi_v) \sum_{m=0}^{\phi_v - 1} \tau_m(k_v, t)}{\langle k \rangle}\end{aligned} \tag{22}$$

$\theta(t)$ is a time-dependent variable, and it will not accomplish its definition after any successful transmission. Therefore, we need to consider its time-dependence to fully understand the systems dynamics from the beginning till the steady-state. If we suppose that an adopted individual transmits behavioral information with probability λ, the decrease in $\theta(t)$ can be written as:

$$\frac{d\theta(t)}{dt} = -\lambda \xi_A(t) \tag{23}$$

At time t, the behavioral information is not transmitted with probability $1 - \lambda$ and the adopted individuals move into recovered state with probability γ, simultaneously. Then;

$$\frac{d\xi_R(t)}{dt} = \gamma(1 - \lambda)\xi_A(t) \tag{24}$$

Substituting (23) into (24) and integrating it with the initial conditions of $\theta(0) = 1$ and $\xi_R(0) = 0$, we can obtain:

$$\xi_R(t) = \frac{\gamma(1 - \lambda)[1 - \theta(t)]}{\lambda} \tag{25}$$

When $t \to \infty$, we find the final adoption size $R(\infty)$ once the degree distribution is known.

3.2 Belief Entropy as a Heuristic

In this study, we utilize belief entropy as a heuristic for the calculation of interference term in (21).

$$D_i j = cos(\theta_i - \theta_j) = -\sum_i m(x_i) ln \frac{m(x_i)}{2^{|x_i|-1} - 1} \tag{26}$$

where $|x_i|$ represents the number of possible actions which is equal to three in our example since individuals can be in one of three states. $m(.)$ denotes the function of belief mass in Dempster-Shafer evidence theory. Accordingly,

$$D_{SA} = cos(\theta_{\xi_S(t)} - \theta_{\xi_A(t)}) = -m(x_{SA}) lnm(x_{SA}) \tag{27}$$

where

$$m(x_{SA}) = \alpha \left| \xi_S(t) + \frac{\xi_S(t) - \xi_A(t)}{\xi_S(t) + \xi_A(t) - 1} \right| \tag{28}$$

It should be noted that masses of all the members of the set add up to a total of 1. Therefore, α is used as a normalization parameter.

$$\alpha \left(\left| \xi_S(t) + \frac{\xi_S(t) - \xi_A(t)}{\xi_S(t) + \xi_A(t) - 1} \right| + \left| \xi_A(t) + \frac{\xi_A(t) - \xi_R(t)}{\xi_A(t) + \xi_R(t) - 1} \right| \right) = 1 \tag{29}$$

Accordingly, (21) can be rewritten as:

$$\begin{aligned} \theta(t) &= \xi_S(t) + \xi_A(t) + \xi_R(t) + \sqrt{\xi_S(t)\xi_A(t)} D_{SA} + \sqrt{\xi_A(t)\xi_R(t)} D_{AR} \\ &= \xi_S(t) + \xi_A(t) + \xi_R(t) - \sqrt{\xi_S(t)\xi_A(t)} m(x_{SA}) lnm(x_{SA}) \\ &- \sqrt{\xi_A(t)\xi_R(t)} m(x_{AR}) lnm(x_{AR}) \end{aligned} \tag{30}$$

4 Results

Here, we aim to test the performance of belief entropy as a heuristic of interference term in (19) and (21). For this purpose, we have performance extensive numerical simulations on uncorrelated random regular networks (RRNs) with $N = 10,000$, $\langle k \rangle = 10$ and $\gamma = 1.0$.

Figure 1 shows the dependence of error between $R(\infty)$ numerical simulations and theoretical analysis by using a quantum-like approach on $cos(\theta_{\xi_S(t)} - \theta_{\xi_A(t)})$ and $cos(\theta_{\xi_A(t)} - \theta_{\xi_R(t)})$ interference terms on different initial probabilities p. Each data pint on heat map figures are calculated by taking the squared difference of $R(\infty)$ versus λ from 0.01 to 1.00 (0.01 increments) between results obtained via theoretical

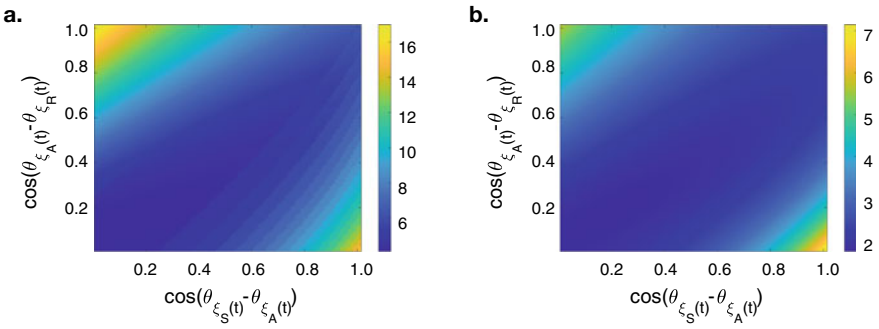


Fig. 1 The dependence of error between $R(\infty)$ numerical simulations and theoretical analysis by using a quantum-like approach on $cos(\theta_{\xi_S(t)} - \theta_{\xi_A(t)})$ and $cos(\theta_{\xi_A(t)} - \theta_{\xi_R(t)})$ interference terms when **a** $p = 0.3$, **b** $p = 0.6$

analysis and numerical simulations using a quantum-like approach. Since the origin points represents when both interference terms are equal to zero, the error degenerates to the one between numerical simulations and the theoretical analysis when classical approach is used ($e^2_{p=0.3} = 4.1707$ and $e^2_{p=0.6} = 1.7643$). Since smaller error values are obtained when $cos(\theta_{\xi_S(t)} - \theta_{\xi_A(t)}) = 0.15$ and $cos(\theta_{\xi_A(t)} - \theta_{\xi_R(t)}) = 0.16$ ($e^2_{p=0.3} = 4.1229$ and $e^2_{p=0.6} = 1.7326$), we can easily argue that quantum-like approach in edge-based compartmental model of message passing approach in the modeling of social contagion performs better compared to the classical method since it can better predict the final adoption size at close to the critical transmission probabilities.

Furthermore, we observed that the optimum value for both interference terms are obtained when $cos(\theta_{\xi_S(t)} - \theta_{\xi_A(t)}) = 0.15$ and $cos(\theta_{\xi_A(t)} - \theta_{\xi_R(t)}) = 0.16$ regardless of the changing initial probability. Using belief entropy in (30) when $t \to \infty$ gives results of $cos(\theta_{\xi_S(t)} - \theta_{\xi_A(t)}) = 0.15$ and $cos(\theta_{\xi_A(t)} - \theta_{\xi_R(t)}) = 0.17$ which shows that our proposed method can be used in the prediction of interference terms in quantum-like social contagion.

5 Discussion

Among a group of individuals, social contagion occurs when ideas, attitudes, and behaviors spread. Even though social contagion used to be regarded as a pathogen in a biological spreading, empirical studies have shown it to be a complex phenomenon due to the social, cognitive, and behavioral differences between individuals. This complexity stems from the heterogeneity of individuals in a social contagion and the uncertainty in their decision-making during the process. In the literature, the former is addressed by using more complex threshold functions to resemble the adoption thresholds of individuals and the latter is challenged with the utilization of quantum-

like functions to better model human decision-making. Although these models are able to portray the complexity of individuals and better model a social contagion process, the interference term brings an additional complexity that decreases the time-efficiency of these models. Additionally, this effect can be determined only by calibration with a real dataset, causing its application harder. In this study, we propose a belief entropy-based heuristic for the determination of interference effect in quantum-like generalization of social contagion. To test the effectiveness of the proposed approach, we have performed extensive numerical simulations on uncorrelated random regular networks (RRNs) and concluded that belief entropy can be used to compute the interference effect in the quantum-like generalization of social contagion models. We believe that these results will increase the use of quantum social contagion models in any application area without having a concern of calibration or time complexity.

References

1. Ashtiani, M., Azgomi, M.A.: A survey of quantum-like approaches to decision making and cognition. Math. Soc. Sci. **75**, 49–80 (2015)
2. Busemeyer, J.R., Bruza, P.D.: Quantum Models of Cognition and Decision. Cambridge University Press (2012)
3. Deng, Y.: Deng entropy. Chaos, Solitons Fractals **91**, 549–553 (2016)
4. Fink, C., Schmidt, A., Barash, V., Kelly, J., Cameron, C., Macy, M.: Investigating the observability of complex contagion in empirical social networks. In: Proceedings of the International AAAI Conference on Web and Social Media, vol. 10 (2016)
5. Goldenberg, J., Libai, B., Muller, E.: Talk of the network: a complex systems look at the underlying process of word-of-mouth. Mark. Lett. **12**(3), 211–223 (2001)
6. Granovetter, M.: Threshold models of collective behavior. Am. J. Sociol. **83**(6), 1420–1443 (1978)
7. Haven, E., Khrennikov, A.: Quantum Social Science. Cambridge University Press (2013)
8. Jizba, P., Arimitsu, T.: The world according to rényi: thermodynamics of multifractal systems. Ann. Phys. **312**(1), 17–59 (2004)
9. Karampourniotis, P.D., Sreenivasan, S., Szymanski, B.K., Korniss, G.: The impact of heterogeneous thresholds on social contagion with multiple initiators. PloS one **10**(11), e0143020 (2015)
10. Karrer, B., Newman, M.E.: Message passing approach for general epidemic models. Phys. Rev. E **82**(1), 016101 (2010)
11. Khrennikov, A.: Quantum-like brain: "interference of minds". BioSystems **84**(3), 225–241 (2006)
12. Khrennikov, A.: Quantum-like model of cognitive decision making and information processing. Biosystems **95**(3), 179–187 (2009)
13. Khrennikov, A., Basieva, I., Pothos, E.M., Yamato, I.: Quantum probability in decision making from quantum information representation of neuronal states. Sci. Rep. **8**(1), 1–8 (2018)
14. Khrennikov, A.Y., Haven, E.: Quantum mechanics and violations of the sure-thing principle: the use of probability interference and other concepts. J. Math. Psychol. **53**(5), 378–388 (2009)
15. Li, M., Wang, X., Gao, K., Zhang, S.: A survey on information diffusion in online social networks: models and methods. Information **8**(4), 118 (2017)
16. Loomes, G.: Variability, noise, and error in decision making under risk. Wiley Blackwell Handb. Judgm. Decis. Mak. **2**, 658–695 (2015)

17. Moreira, C., Wichert, A.: Interference effects in quantum belief networks. Appl. Soft Comput. **25**, 64–85 (2014)
18. Mutlu, E.C.: Quantum probabilistic models using feynman diagram rules for better understanding the information diffusion dynamics in online social networks. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, pp. 13730–13731 (2020)
19. Mutlu, E.C., Ozmen Garibay, O.: Quantum contagion: a quantum-like approach for the analysis of social contagion dynamics with heterogeneous adoption thresholds. Entropy **23**(5), 538 (2021)
20. Namdari, A., Li, Z.: A review of entropy measures for uncertainty quantification of stochastic processes. Adv. Mech. Eng. **11**(6), 1687814019857350 (2019)
21. Pastor-Satorras, R., Castellano, C., Van Mieghem, P., Vespignani, A.: Epidemic processes in complex networks. Rev. Mod. Phys. **87**(3), 925 (2015)
22. Rényi, A.: On measures of entropy and information. In: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics, vol. 4, pp. 547–562. University of California Press (1961)
23. Rudolf, C.: The Mechanical Theory of Heat: With Its Applications to the Steam-Engine and to the Physical Properties of Bodies (1867)
24. Savage, L.J.: The Foundations of Statistics. Courier Corporation (1972)
25. Shannon, C.E.: A mathematical theory of communication. Bell Syst. Tech. J. **27**(3), 379–423 (1948)
26. Shannon, C.E.: A mathematical theory of communication. ACM Sigmobile Mob. Comput. Commun. Rev. **5**(1), 3–55 (2001)
27. Simon, H.A.: A behavioral model of rational choice. Q. J. Econ. **69**(1), 99–118 (1955)
28. Tsallis, C.: Possible generalization of boltzmann-gibbs statistics. J. Stat. Phys. **52**(1), 479–487 (1988)
29. Tversky, A., Kahneman, D.: Judgment under uncertainty: heuristics and biases. Science **185**(4157), 1124–1131 (1974)
30. Wang, W., Chen, X.L., Zhong, L.F.: Social contagions with heterogeneous credibility. Phys. A: Stat. Mech. Its Appl. **503**, 604–610 (2018)
31. Wang, W., Tang, M., Shu, P., Wang, Z.: Dynamics of social contagions with heterogeneous adoption thresholds: crossover phenomena in phase transition. New J. Phys. **18**(1), 013029 (2016)
32. Watts, D.J.: A simple model of global cascades on random networks. Proc. Natl. Acad. Sci. **99**(9), 5766–5771 (2002)
33. Yukalov, V.I., Yukalova, E.P., Sornette, D.: Information processing by networks of quantum decision makers. Phys. A: Stat. Mech. Its Appl. **492**, 747–766 (2018)
34. Zhu, X., Wang, W., Cai, S., Stanley, H.E.: Dynamics of social contagions with local trend imitation. Sci. Rep. **8**(1), 7335 (2018)