Optimization of Multivariate Hawkes Processes via Decentralized Learning Automata: Towards Learning How to Boost Factual Information on Social Networks

Ahmed Abdulrahem Abouzeid

UiA University of Agder

Doctoral Dissertations at the University of Agder 439

Optimization of Multivariate Hawkes Processes via Decentralized Learning Automata

Ahmed Abdulrahem Abouzeid

Optimization of Multivariate Hawkes Processes via Decentralized Learning Automata

Towards Learning How to Boost Factual Information on Social Networks

Doctoral Dissertation for the Degree *Philosophiae Doctor (Ph.D.)*
at the Faculty of Engineering and Science, Specialization in Information and Communication Technology, Artificial Intelligence

University of Agder
Faculty of Engineering and Science
2023

Doctoral Dissertations at the University of Agder 439
ISSN: 1504-9272
ISBN: 978-82-8427-156-9

Printed by Make!Graphics
Kristiansand

To my parents, my sister, and my brother

...

Preface

"The phenomenology of spirit is exposition of the coming to be of knowledge. This is explicated through a necessary self-origination and dissolution of the various shapes of spirit as stations on the way through which spirit becomes pure knowledge"

-Georg Wilhelm Friedrich Hegel

The author of this Ph.D. dissertation, Ahmed Abdulrahem Abouzeid, comes from Egypt, where he had received his Bachelor's in Computer Science in 2006. Then, he practiced Software Engineering in the industry for more than eight years. In 2017, he received his Master's degree in Informatics from Eötvös Loránd University, Budapest, Hungary.

The author believes that the central structure of an experience is its intentionality, and hence being directed towards such an experience with a particular belief concluded from past experiences. Therefore, this dissertation's motivation can be viewed as a reflection of a life's journey — since the Egyptian revolution occurred in 2011 and the chance to witness the Egyptian people's dreams as they pursued liberty and equality. However, the people's movement challenged political manipulation, which led to the neutralization of their movement. Nevertheless, since then, a spirit, and a mind have been shaped and evolved, committing to the matter of Truth revealing.

The author believes that Artificial Intelligence (AI) is yet to be matured in societal issues such as hate speech detection, fake news detection, and political manipulation on Social Media (SM). One of the primary reasons for that is the lack of enough collaboration between the AI community and the Social Science community. That causes AI models to learn from biased or limited representations of the tackled problem. Therefore, the primary consideration of this study was to understand the learning problem before understanding the learning methods.

The following research study focused on representations and realistically designing the learning problem in addition to the learning methods. The work could interest readers from AI and Computational Social Sciences communities, where topics such as Stochastic Optimization, Non-sequential Reinforcement Learning (RL) for decision making, and SM users' behavior modeling were studied.

Acknowledgments

The author would like to thank his principal Ph.D. supervisor, Professor Ole-Christoffer Granmo, for the four-years journey which was full of patience, guidance, insightful discussions, and academic freedom. The author would also like to thank the co-supervisors, Professor Morten Goodwin and Professor Christian Webersik, for all the words of advice they provided.

The author is also thankful for colleagues in the Center for Artificial Intelligence Research (CAIR) and the Center for Integrated Emergency Management (CIEM) at the University of Agder, Norway, where this research was conducted and inspired by a lot of brainstorming and inspired conversations. The author would also like to express his gratitude and thank the colleagues he met and shared a lot of philosophical conversations about world-facing issues at the Digital Media Research Center (DMRC) at the Queensland University of Technology, Australia.

The author would also like to express his deepest appreciation and gratitude to two remarkable colleagues, Emma and Tonje, whose unwavering support and friendship have played an integral role in his journey in Norway and the University of Agder. Their exceptional welcome and the sense of belonging they fostered made an immeasurable impact on the successful completion of this dissertation.

Lastly, the author wants to express his heartfelt appreciation to his family and friends who have supported him throughout this academic pursuit.

Ahmed Abdulrahem Abouzeid
Kristiansand, Norway
15/06/2023

Summary

The increasing amount of misleading information on Social Media (SM) platforms is problematic. The reason is that these platforms have become one of the primary sources of information due to their ease-of-use and cheap cost of information acquisition. One example is that misleading information can disturb the social order and recovery from emergencies, recently actualized by the infodemic of COVID-19 and the Russian-Ukrainian information war. As the amount of misleading information on SM increases, we risk that consumers start mistrusting even reliable information sources. To this end, a wide range of Artificial Intelligence (AI)-based solutions were proposed to combat such an issue. One common approach is the intervention-based misinformation mitigation on SM, where the task is to mitigate the exposure to misinformation by alternatively boosting the exposure to factual information. Traditionally, the exposure to a particular information type for each user is defined as the count of propagated content of that information type by the adjacents of that user, e.g., followees on Twitter. To boost exposure to online factual content, SM users are incentivized to propagate these facts first, and their network adjacents such as their followers on Twitter can then reach and interact with these facts. Hence, in this context, intervening with users means incentivizing them to change their information dissemination behavior by propagating more factual information.

Because each user has a different number of adjacent users and each user has a different level of exposure to misinformation, the individual incentives should be determined differently according to the following. On one hand, (A) how much misinformation exposure a user has? On the other hand, (B) how much exposure to misinformation do adjacent users have? In all cases, (C) How likely are a user or the adjacent users will accept or be influenced by the determined incentivization?

Traditionally, the learning of individual incentives is facilitated through the Reinforcement Learning (RL) framework. In the latter, the dynamics of SM users' online engagements are modeled through a simulated social network environment from which the RL agents can learn about users' behavior. However, there have been relatively few intakes on how to learn and evaluate the optimal individual incentivization required to achieve optimal mitigation outcomes. For instance, existing criterion functions and representations mainly focused on quantities of misinformation and factual information which each user is exposed to, without considering the root causes that drive these exposures. Hence, the answer to question (C) was not investigated adequately. We believe the latter is a noticeable drawback in the proposed solutions because the simulated network and the incentivization procedure

should be conducted over the best possible representation and criterion function that reflect real-world dynamics on SM.

In this research, we propose a novel approach utilizing RL, specifically Learning Automaton (LA). Our method combines the principles of RL with the adaptive decision-making capabilities of LA to address the challenges of user needs-based incentivization learning. Further, we propose a novel simulation-based optimization framework with novel users' activity representation to model the task of intervention-based misinformation mitigation. Driven by the novel activity representation, we propose a novel criterion function that considers the key factors that influence information propagation on SM instead of only calculating quantities of misinformation and factual information exposures. These key factors were proposed in recent Social Science literature and illustrated what dictates misinformation spread on today's SM platforms. In that manner, we propose temporal activities of societal bias, content engagement, and the propagation patterns of both misinformation and factual information. Thus, we do not assume incentives to be assigned based on the quantities of exposure to misinformation only but rather evaluated and assigned based on the probability of agreeing with a content or an opinion that has a particular bias, in addition to the probability of engaging with it in the first place. Further, the study proposes preliminary algorithms to help verify and self-learning of SM activity categories such as political bias and information type. Finally, our empirical results show three main significant properties. First, they demonstrate how our novel mitigation algorithms perform better in most of the scenarios when compared to traditional RL algorithms. Second, the results indicate how our proposed criterion functions are robust to different network statistics in terms of different percentages of misinformation exposure among users. Third, our novel activity representation is more transparent and extended the analytical capacity to a misinformation mitigation solution. The latter is recognized in the provided capabilities of tracing the change in probabilities of societal bias and content engagement as a consequence of the intervention.

In all brevity, this research investigates questions and problem variables beyond the existing misinformation benchmark datasets and their underlying representations. The study gathered additional comprehensive and crucial data to create a well-developed learning setting and standard for the proposed LA agent. Our research aims to connect the realms of AI and Social Science by examining pertinent theoretical studies concerning the issue of misinformation spreading on SM and the interconnected dynamics that oversee this process.

Sammendrag

Den økende andelen av villedende informasjon på sosiale medier er problematisk, da disse i økende grad har blitt en av hovedkildene til informasjon grunnet deres brukervennlighet og lave informasjonskostnad. Misvisende informasjon kan forstyrre sosial orden og hindre gjenoppretting fra nødsituasjoner, som nylig ble aktualisert gjennom «infodemien» av COVID-19 og den russisk-ukrainske informasjonskrigen. Økningen i villedende informasjon kan føre til at brukere mister tillit til også pålitelige informasjonskilder. Flere kunstig intelligens løsninger har blitt foreslått for å bekjempe dette problemet. En vanlig tilnærming er en bekjempelses-strategi der man demper spredningen av villedende informasjon ved å øke eksponeringen for faktisk «korrekt» informasjon. Tradisjonelt sett blir eksponeringen av en spesifikk informasjonstype for hver bruker definert som antall ganger innhold av den informasjonstypen er blitt delt i brukerens nettverk, for eksempel av personer brukeren følger på Twitter. Altså, for å øke eksponeringen for faktisk «korrekt» innhold på nettet, blir sosiale mediebrukere oppmuntret til å spre det korrekte innholdet først, og deres nettverkskontakter, som deres følgere på Twitter, kan deretter nå og samhandle med denne «korrekte» informasjonen. Måten man påvirker brukerne i denne sammenhengen er ved å motivere dem til å endre sin atferd når det gjelder spredning av misinformasjon ved å spre mer «korrekt» informasjon.

Fordi hver bruker har et ulikt antall tilstøtende brukere, og hver bruker har ulik grad av eksponering for villedende informasjon, bør individuelle insentiver bestemmes forskjellig i henhold til følgende. På den ene siden, (A) hvor mye eksponering for villedende informasjon en bruker har? På den annen side, (B) hvor mye eksponering for villedende informasjon har tilstøtende brukere? I alle tilfeller (C), hvor sannsynlig er det at en bruker eller en tilstøtende bruker vil akseptere eller påvirkes av det bestemte insentivet?

Tradisjonelt sett brukes forsterkningslæring for å finne de beste insentivene for hver enkelt bruker i kampen mot misvisende informasjon. Med forsterkningslæringsmetoder, blir dynamikken til de sosiale medie-brukernes online-engasjement modellert gjennom et simulert sosialt nettverksmiljø, der forsterkningslærings-algoritmene kan lære fra brukeradferden i det simulerte miljøet. Imidlertid har det vært relativt få forsøk på å lære og evaluere de optimale individuelle insentivene som kreves for å oppnå en optimal demping av misvisende informasjon. For eksempel har nåværende kriteriefunksjoner og -representasjoner hovedsakelig fokusert på hvilke mengder av villedende og faktisk «korrekt» informasjon som hver bruker har blitt eksponert for, uten at det har blitt tatt hensyn til de grunnleggende årsakene som driver disse ek-

sponeringene. Altså er ikke spørsmål (C), slik nevnt ovenfor, tilstrekkelig undersøkt i nåværende forskning. Vi mener at dette gir en ulempe i eksisterende løsninger, fordi den simulerte nettverksmodellen og incentiviserings-prosedyren bør gjennomføres basert på den beste mulige representasjonen og kriteriefunksjonen som reflekterer virkeligheten på sosiale medier.

I denne forskningen presenterer vi en ny tilnærming ved å bruke forsterkningslæring, og mer spesifikt, en såkalt læringsenhet referert til som Learning Automaton (LA). Metoden vår kombinerer prinsippene i forsterkningslæring med de adaptive beslutningsevnene i LA for å håndtere utfordringene med læring av insentiver basert på brukerbehov. Videre foreslår vi en ny simuleringsbasert optimaliseringsmodell med en innovativ representasjon av brukeraktivitet for å modellere oppgaven med intervensjonsbasert demping av misvisende informasjon. Med utgangspunkt i den nye aktivitetsrepresentasjonen, har vi utviklet en ny kriteriefunksjon som tar hensyn til nøkkelfaktorene som påvirker informasjonsspredning på sosiale medier, i stedet for å kun beregne mengder av eksponering for villedende og faktisk «korrekt» informasjon. Disse nøkkelfaktorene ble foreslått i nyere samfunnsvitenskapelige studier og illustrerer hva som styrer spredningen av misvisende informasjon på dagens sosiale medieplattformer. På den måten har vi foreslått temporære aktiviteter som samfunnsmessig skjevhet, engasjement med innhold og spredningsmønsteret til både villedende og faktisk «korrekt» informasjon. Derfor antar vi ikke at insentiver blir tildelt bare basert på mengden eksponering for villedende informasjon, men at insentivet heller blir vurdert og tildelt basert på sannsynlighet for å være enig med et bestemt innhold eller mening som har en spesiell skjevhet. I tillegg veier sannsynligheten for at brukeren engasjerer seg med innholdet i utgangspunktet inn. Videre foreslår studien algoritmer for å hjelpe til med å verifisere og selv lære om aktivitetskategorier på sosiale medier, som for eksempel politisk skjevhet og informasjonstype. Til slutt viser våre empiriske resultater tre hovedegenskaper. For det første viser de hvordan våre nye misvisende informasjon-dempingsalgoritmer er bedre i de fleste scenarier sammenlignet med tradisjonelle forsterkningslærings-algoritmer. For det andre indikerer resultatene hvordan våre foreslåtte kriteriefunksjoner er robuste mot forskjellige nettverksstatistikker med ulike prosentandeler av villedende informasjonseksponering blant brukere. For det tredje er vår nye aktivitetsrepresentasjon mer forklarbar og utvider den analytiske kapasiteten til en misvisende informasjon-dempingsløsning. Det sistnevnte funnet anerkjennes gjennom evnen til å spore endringer i sannsynligheten for samfunnsmessig skjevhet og engasjement med innhold som en konsekvens av intervensjon.

I korte trekk undersøker denne forskningen spørsmål og problemstillinger som går utover eksisterende referansedatasett for desinformasjon og deres underliggende representasjoner. Studien har samlet inn omfattende og avgjørende data for å skape en godt utviklet læringssetting og standard for den foreslåtte forsterkningslæringsagenten. Målet med vår forskning er å koble sammen fagområdene kunstig intelligens og samfunnsvitenskap ved å undersøke relevante teoretiske studier om problemet med spredning av misvisende informasjon på sosiale medier og de sammenhengende dynamikkene som styrer denne prosessen.

Publications

All the papers listed below are an outcome of the research investigation carried out by the author of this dissertation. Including one submitted (**Paper G**) and six published papers.

Paper A: Abouzeid, Ahmed, et al. "Causality-based Social Media Analysis for Normal Users Credibility Assessment in a Political Crisis." 2019 25th Conference of Open Innovations Association (FRUCT). IEEE, 2019.

Paper B: Abouzeid, Ahmed, et al. "Learning Automata-based Misinformation Mitigation via Hawkes Processes." Information Systems Frontiers 23.5 (2021): 1169-1188.

Paper C: Abouzeid, Ahmed, Ole-Christoffer Granmo, and Morten Goodwin. "Modelling Emotion Dynamics in Chatbots with Neural Hawkes Processes." International Conference on Innovative Techniques and Applications of Artificial Intelligence. Springer, Cham, 2021.

Paper D: Abouzeid, Ahmed, et al. "Socially Fair Mitigation of Misinformation on Social Networks via Constraint Stochastic Optimization." AAAI. 2022.

Paper E: Abouzeid, Ahmed, and Ole-Christoffer Granmo. "MMSS: A Storytelling Simulation Software to Mitigate Misinformation on Social Media." Software Impacts 13 (2022): 100341.

Paper F: Abouzeid, Ahmed, et al. "Label-Critic Tsetlin Machine: A Novel Self-supervised Learning Scheme for Interpretable Clustering." 2022 International Symposium on the Tsetlin Machine (ISTM). IEEE, 2022.

Paper G: Abouzeid, Ahmed, et al. "Novel Users' Activity Representation for Modeling Societal Acceptance Towards Misinformation Mitigation on Social Media." Journal of Computational Social Science (2023).

Contents

List of Figures

List of Tables

Abbreviations

AI	Artificial Intelligence
BN	Bayesian Network
BNs	Bayesian Networks
CPTs	Conditional Probability Tables
DAG	Directed Asyclic Graph
DBN	Dynamic Bayesian Network
HP	Hawkes Process
HPs	Hawkes Processes
ICPTs	Importance Conditional Probability Tables
KA	Krinsky Automaton
LA	Learning Automaton
LAKG	Learning Automata Knapsack Game
LAs	Learning Automata
LSTM	Long-Short Term Memory
MCMHP	Multiplex-Controlled Multivariate Hawkes Processes
MHP	Multivariate Hawkes Process
MHPs	Multivariate Hawkes Processes
MKP	Multidimensional Knapsack Problem
MMSS	Misinformation Mitigation Storytelling Simulation
NEHP	Neural Emotion Hawkes Process
RL	Reinforcement Learning
RNN	Recurrent Neural Network
RNNs	Recurrent Neural Networks
SM	Social Media
TA	Tsetlin Automaton
TAs	Tsetlin Automata
TM	Tsetlin Machine
TMs	Tsetlin Machines

Part I

Dissertation Overview, Relevance, and Findings

Chapter 1

Introduction

This chapter gives an introduction to the study conducted in this dissertation. The chapter highlights the scientific and societal motivation for the research focus and provides an overview of the key topics and research questions covered in the study.

1.1 Motivation

The increasing amount of misleading information on Social Media (SM) platforms is problematic [1]. The reason is that these platforms have become one of the primary sources of information due to their ease-of-use and cheap cost of information acquisition [2]. One example is that misleading information can disturb the social order and recovery from emergencies [3], recently actualized by the infodemic of COVID-19 [4] and the Russian-Ukrainian information war [5]. As the amount of misleading information on SM increases, we risk that consumers start mistrusting even reliable information sources [6]. Therefore, increasing the veracity of the information on SM platforms is critically required [7].

Unfortunately, SM platforms such as Facebook and Twitter are geared towards maximizing user engagement over similar preferences instead of optimizing for information veracity [8]. The latter causes the so-called Echo Chambers Effect [9], which traps users inside their comfort zones and blocks them from exploring other systems of beliefs and opinions. Hence, narrow-minded societal bubbles emanate [10], intensifying polarization and extremism.

Polarization and extremism on SM platforms lead to another harmful phenomenon [11]: the so-called Confirmation Bias. The latter occurs when users search for, interpret, and favor information that satisfies their prior beliefs. Hence, Confirmation Bias could have a crucial impact on how online users would perceive online content [12] — for instance, the political climate in society is likely to become vulnerable to manipulative online political campaigns [13, 14, 15].

The effect of Echo Chambers and Confirmation Bias can also be observed in how the social network's extreme polarization extends the lifetime of manipulative online political campaigns [1]. In this manner, the network's societal bubbles that embrace the manipulated contents will turn the latter into opinions [16]. That leads to the

problem of the unintentional spread of false content [16], a.k.a. "misinformation".

The problem of misinformation on SM is very challenging since misinformation by nature is resilient towards fact-checking attempts [17]. This is because a machine learning-based fact-checking classifier is trained on particular linguistic features — but, when a random person expresses a deceptive opinion with a random text, that could introduce many outliers that lower the classification accuracy of the model [18]. Further, fact-checking methods are judgmental in the sense that they provide a definite classification of users' contents, hence, any error in their outcomes will violate the freedom of speech, which is an ethical concern about machine learning classifiers [19].

Thus, the above challenges led us to investigate novel machine learning approaches and methods that learn how to boost factual information on SM by incentivizing online users to share verified content, a.k.a. "intervention-based misinformation mitigation". Unlike machine learning classifiers for misinformation detection, the adopted approach in this study allows for ideological debates, the democratization of combating misinformation, and reducing the risk of violating human rights.

1.2 Scope

This section introduces the approaches that sum up the scope of this dissertation. In subsection 1.2.1, we give an overview of the approach of misinformation mitigation. In subsection 1.2.2, we briefly demonstrate information diffusion modeling, which is a fundamental technique in SM analysis to predict online activities. Additionally, subsection 1.2.3 briefly explains how controlling a diffusion model could achieve misinformation mitigation. In subsection 1.2.4, we give a quick overview of the knapsack optimization problem, which has been used to define a wide range of optimization tasks. Figure 1.1 illustrates the flow of information and interaction between the methods under the scope of our study. In brief:

- An information diffusion model predicts the users' temporal activities through a parametric prediction function.

- The prediction function parameters are adjusted through optimized values by a control model. The latter should then reshape the social network toward optimal temporal user activities that serve our purpose of mitigating misinformation exposure. The exposure to a particular information type for each user in a given period of time is traditionally defined [20] as the count of propagated content of that information type by the neighbors of that user (e.g., followees on Twitter).

- Repeat the above in the same order until convergence to optimal knapsack items' values, where the latter represents the prediction function parameters adjustments within a constrained capacity of a knapsack [21].

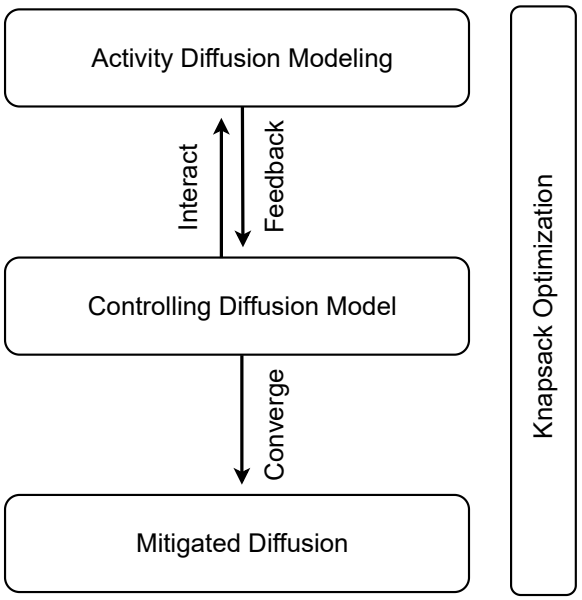


Figure 1.1: The flow of information and interaction between the main elements of the scope of this dissertation

1.2.1 Misinformation Mitigation

As discussed above, misinformation can persist on the network as it becomes a point of view. Thus, approaches such as fake news detection [22] and troll detection [23] are insufficient. That is because these approaches are usually offline solutions, and topics such as polarized opinions and societal denial of other points of view are not in the scope of these models. In contrast, a misinformation mitigation approach can focus on the network dynamics that influence the ongoing process of misleading content and proposes an online intervention-based resolution [20].

Further, since fake news and fake accounts detection methods are judgmental by providing a definite classification of users' contents or profiles, any error in their outcomes will violate the freedom of speech [24]. Alternatively, the misinformation mitigation approach aspires to democratize and intensify other viewpoints by considering the online content as opinions to interact with and possibly be changed.

One common technique in the misinformation mitigation approach and also what we adopted in this study is to learn about the incentives required for each user to support and contribute to a truth[1] campaign to neutralize the misinformation on the network [25, 20].

The idea behind incentivizing users to share the other point of view is to make them introduce a diversity of information to the network. However, the latter also creates a paradox of diversity and polarization — because in a very heterogeneous

[1]In this dissertation, we are using the word "truth" without capitalization — except for titles since we cannot argue that our truth campaign's factual information means a transcendent idea in the platonic sense. Hence, a truth campaign is just a method to propagate the other point of view that is believed and hoped to be true. Thus, we would also like to highlight our recommendation regarding utilizing our method under the supervision of society and its authentic institutions to protect against the misuse of technology.

society, the polarization can be sustained under intense pressures [26], and people become more protective of their ideas when others oppose their beliefs. Such paradox makes accomplishing misinformation mitigation very challenging.

1.2.2 Information Diffusion Modeling

Learning about users and their potential information dissemination patterns requires observing their temporal activities on the network [27]. In an intervention-based truth campaign [25], further intervention with users aims to reshape their network activities to circulate factual information. Hence, more challenges arise since real-time intervention with the social network is infeasible. The latter challenge occurs because an external algorithm cannot reach, intervene, and observe changes in all users on the SM platform. Thus, one way to observe and evaluate different intervention strategies is by modeling the information diffusion process on the network.

The idea of an information diffusion model [28] is to simulate users' activities and predict future information dissemination patterns and consequences of different intervention strategies. For example, interventions that incentivize users to propagate a particular type of information. Therefore, there is a critical need for obtaining realistic representations and a trustworthy simulated network [29] that reflects real-world dissemination patterns.

Figure 1.2 demonstrates an example of a social network and its temporal users' activities prediction from an information diffusion model, where the latter predicts the temporal dissemination patterns of harmful political propaganda and non-propaganda content types in a political context.

1.2.3 Diffusion Model Control

Traditionally, Policy-based Reinforcement Learning (RL) methods [30, 25, 20] enabled the intervention with the simulated social network to learn about campaign incentives. These methods control the diffusion model, where the latter estimates the temporal users' activities before and during the intervention. The RL agents', where each is associated with a user on the network — are tasked to learn a policy of the optimal incentive values required for each user during an intervention. The intervention is in the form of assigning some hypothetical incentive values based on a predefined criterion that characterizes users' needs in terms of being vulnerable to misinformation, and hence, each user would need more factual information exposure.

The RL agents can learn an optimal incentivization policy through a designed criterion reward function that evaluates how good an incentive was for each user. After assigning such incentives, the reward function evaluates the predicted consequent temporal activities from the diffusion model. If an agent receives a positive reward, it commits to the current assigned incentive value. Otherwise, it rolls back to a previous incentive value, given that an initial value of an incentive is 0. Iteratively, the agents perform state transitions that characterize the assignment of different incentive values until they converge to optimal values of individual incentives.

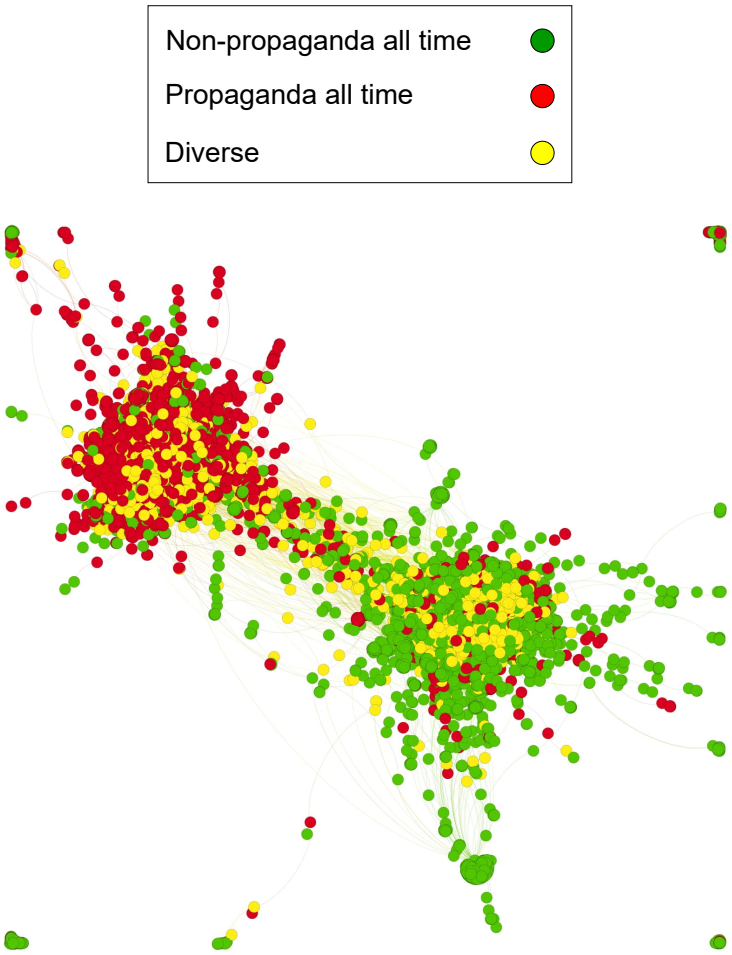


Figure 1.2: A colored graph showing information dissemination patterns inferred from a diffusion prediction function, where nodes and edges represent users and their following relationships, respectively. Colors represent the dissemination over time of a particular information type

1.2.4 Knapsack Optimization

Knapsack optimization aims to maximize the outcome of some utilized items that can fit into a knapsack without exceeding a maximum capacity constraint [21]. Therefore, the choice of the particular items with the specific amount of each item should be optimal to maximize the benefit within the capacity boundary. With an analogy to the task of misinformation mitigation via incentive learning, the knapsack items can be considered the individual users' incentives whose determined values influence the optimality of the knapsack. Therefore, the users' incentive values should be optimal and add up to a value within an incentivization budget constraint.

The advantage of defining the problem of misinformation mitigation as a knapsack optimization problem is how easy it becomes to understand and analyze the learned users' incentives with their relevance to the optimal mitigation results we seek. Moreover, the knapsack optimization will provide more insights learned from previous optimization algorithms over the knapsack. For example, whether or not greedy algorithms could be better for solving a particular knapsack problem [31].

1.3 Research Gaps

The existing limitations in intervention-based misinformation mitigation can be divided into three primary categories: (1) the mitigation learning algorithms, (2) the representation of the problem variables, which is reflected in the configuration of the information diffusion model, and (3) the perspective of Artificial Intelligence (AI) trustworthiness. We provide an overview of some of these limitations in this section.

1.3.1 The Mitigation Learning Algorithms

The work proposed by Farajtabar et al. (2017) [20] has established a novel truth campaign approach by reshaping the simulated network activities to boost factual information during misinformation circulation on SM. That was achieved through policy-based RL to optimize an information diffusion model. The successively proposed learning algorithms for the same task have also utilized the same framework [25, 30] to learn the mitigation policy. Since the truth campaign optimization task on SM is relatively new, there is an expansive scope of contribution to novel methods other than the existing policy-based RL. That drives the need and curiosity to investigate more lightweight algorithms since the nature of the problem is time-critical.

Further, the existing criterion reward functions for the RL algorithms [20, 25, 30] mainly assumed that evaluating an incentive value for its optimality should be conducted based on how much a user needs to be exposed to more factual information. One common technique to quantify the latter was by calculating the difference between misinformation and factual information exposures [20]. However, we believe more research should investigate what else could be governing the dynamics of such a process and then replicate that to contemporary criterion reward functions.

1.3.2 Representation of the Problem Variables

Intervention-based misinformation mitigation algorithms on SM generally depend on how information circulates and users' activities occur. In these scenarios, a control model intervenes with an information diffusion model, where the latter characterizes the social network dynamics [20]. These diffusion models require temporal historical network data to learn interdependencies of network activities for predicting future temporal propagation. Unfortunately, the currently existing data sources do not provide a comprehensive representation of the problem variables. For instance, the currently available datasets [32, 33, 34, 35, 36, 17] focused mainly on information veracity labels and the timestamps of true or false contents, ignoring other interdependent time-varying contexts, such as the individual political bias and the biased community engagements on SM [37]. Therefore, a more realistic diffusion model needs to learn from diverse time-tagged samples on the various categories of the problem variables rather than just one category of information.

Moreover, the above is also connected to the existing RL criterion reward functions [20, 25, 30] which as a result suffered from such limited representation. In this manner, existing criterion reward functions returned a reward signal based on naive heuristics such as a stochastic signal from a single-modality diffusion model with only timestamps information about the dissemination of either true or false information.

1.3.3 Trustworthiness of the Method

The currently proposed truth campaign optimization methods must be verifiable to suggest the learned users' incentives in the real world. Hence, one fundamental question is: what do these incentives mean in the real world?

The answer to the above question will be significant in the field of intervention-based misinformation mitigation research. Unfortunately, the currently proposed mitigation methods did not fill that research gap. However, interpreting and verifying the converged incentive values is essential since these incentives are being evaluated within a simulation (i.e., the diffusion model) and not the real world. Further, it is challenging for a diffusion prediction model to accurately predict human behavior, because the human brain's decision-making process is highly challenging [38].

Another aspect of any proposed solution's trustworthiness is how the misinformation mitigation approach depends on a machine learning classifier to detect and categorize the targetted information categories for the diffusion model. These classifiers are usually a black box and cannot provide interpretability of their results [39], which is a big disadvantage when judging people's authenticity.

1.4 Contributions

1.4.1 Major Contribution

This dissertation's major research focused on utilizing Learning Automaton (LA) [40] as an incentive control model for a knapsack optimization problem. The latter was applied to solve a truth campaign optimization for misinformation mitigation via controlling an information diffusion model. We utilized Multivariate Hawkes Processes (MHPs) to characterize the network diffusion volumes [41].

The main diffusion volumes we eventually sought to optimize were the counts of temporal users' activities regarding their engagement with factual information. The latter was then evaluated with regard to the knapsack optimality and the predefined criterion for that. To the best of our knowledge, our research is the first attempt to utilize LA to control a MHP-based knapsack for the problem of misinformation on SM, where our proposed network of LAs solved another non-convex optimization problem [42].

Our choice of the LA as a diffusion control model was due to how easily implementable and lightweight an automaton is. Therefore, we believe our research significantly widens the application of LA in the literature and opens the venue for more decentralized and easily implementable methods for time-critical issues such as emergency responses to an infodemic on SM [43].

To this end, our coupling of LAs and MHPs allowed for the below major contributions:

- The study proposed novel temporal user activity representations to provide close-to-reality SM dynamics by modeling and predicting the interdependence between multiple temporal activity categories.

- Driven by the above point, we modeled the recently proposed theoretical strategy of intensifying the societal acceptance on SM [44]. Hence, this dissertation provided an example of integrating Social Science and AI for a typical societal problem and provided a well-established criterion reward function for the proposed LA mitigation algorithm.

- The societal acceptance awareness in our study yielded a novel optimization loss function where societal acceptance awareness was constructed through novel function domain variables.

- The study proposed easy-to-implement and lightweight novel LA-based mitigation algorithms with sampling techniques for size-scalable social networks. Further, the work offered mitigation software that facilitates interactive visualization of consequent diffusion scenarios to provide an informative evaluation of different mitigation strategy parameters.

- All the above contributions resulted in finding a novel architecture and mathematical model. Thus, this dissertation proposed the Multiplex-Controlled Mul-

tivariate Hawkes Processes (MCMHP) — a unified model to define parametric-MHP control tasks formally. As one possible application of the proposed MCMHP, the latter was evaluated for the integration of multiple interdependent diffusion volumes categories to optimize. The latter functionality extended the analytical capacity of existing intervention-based misinformation mitigation models. Hence, instead of only optimizing and tracing the intensities of factual information engagement, we also traced and optimized other network dynamics that derive the engagement with factual information.

1.4.2 Minor Contribution

As a minor contribution, we proposed a preliminary self-supervised machine learning scheme and architecture based on the so-called Tsetlin Machine (TM) [45] to self-learn categories without the existence of the ground truths. Our proposed architecture allows for interpretable [46] clustering of such categories, which is crucial when judging the authenticity of users' generated content on SM. Further, the self-learning characteristics facilitate an independent and transparent misinformation mitigation pipeline since no manual work is needed to annotate the data before training the MHP. Additionally, we proposed a causal graph-based mechanism to verify users' trustworthiness on SM, which could also be fine-tuned to verify the learned truth campaign incentives.

1.4.3 Availability of Collected Datasets and Source Code

The collected datasets and developed source code for all experiments done for this dissertation; were made public[2].

1.4.4 Research Questions

This dissertation investigated the following five research questions (RQ) to fill some relevant research gaps for the required sub-tasks for the problem of intervention-based misinformation mitigation on SM. While we deeply investigated some of these questions and believed we made a major contribution, we contributed with some preliminary methods to initially answer the other questions.

- Major - RQ1: How to design a lightweight control model and misinformation mitigation algorithm using LA for truth campaign optimization?
- Major - RQ2: How to introduce ethical considerations, like fairness, into the criterion and optimization loss functions, while maintaining effective mitigation?
- Major - RQ3: How can contemporary information diffusion models be enriched to capture community behavior, and how can the resulting enriched models lead to improved mitigation loss functions for our LA design?

[2]https://github.com/Ahmed-Abouzeid?tab=repositories

- Minor - RQ4: How can the misinformation mitigation strategy's learned incentives be verified for correctness in real-world practices?

- Minor - RQ5: What could be the best approach to provide transparency and independence for a misinformation mitigation pipeline? For example, how to self-learn the temporal activities of online users with justification on the classification results?

1.5 Dissertation Outline

This dissertation is organized into two parts. First, Part I gives an idea of the motivation, scope, contributions, research questions, theoretical background, and proposed methods. Part I has four chapters, chapter 1 demonstrates the motivation and scope of the study. chapter 2 provides more details on the relevant theoretical background work and positions this dissertation's main focus. chapter 3 explains the proposed methods. Eventually, chapter 4 concludes the dissertation and provides some insights on future directions. Second, Part II provides the full text of all submitted or published research papers as an outcome of this study. We demonstrate below the primary outcome of each paper with its relevance to the dissertation and the other included papers.

- **Paper A (see Appendix A):** This paper was our first attempt to investigate the problem of misinformation on SM. The paper proposed a probabilistic causal model as a theoretical view on the problem of normal users' credibility on SM. The work introduced a Causal Bayesian Network (BN) [47] inspired by the potential main entities that would be part of the misinformation process dynamics. The paper's methodology examined the problem solution in a causal manner which considered the task of misinformation detection as a question of cause and effect rather than just a classification task.

  The causality-based approach provided a practical road map for some sub-problems in real-world scenarios, such as sensitivity analysis [48] and verification of the learned truth campaign incentives. For example, the LAs decided incentives could be represented as causal parent nodes with a misinformation state as a child node, and polarization with societal engagement could be represented by other interconnected nodes. Moreover, the causality approach facilitates intervention simulations which would unveil a more casual analysis of the diffusion patterns on the network. Hence, this paper outcome contributed partially to RQ4 as highlighted in subsection 1.4.4.

- **Paper B (see Appendix B):** This paper proposed a novel lightweight intervention-based misinformation mitigation framework using decentralized LAs to control the MHP simulated network. Each automaton was associated with a single user to learn how effective the latter is when gets involved in a truth campaign to circulate factual information.

The approach conducted in this paper showed fast convergence and increased the factual information exposure on the simulated network. These results persisted independently from network structure, including networks with central nodes, where the latter could be the root of misinformation. Further, the LAs obtained these results in a decentralized manner, facilitating distributed deployment in real-life scenarios.

This work answered RQ1, where we evaluated the LA on the problem of stochastic knapsack optimization for misinformation mitigation. Thus, this paper established the basis to answer RQ2 and RQ3 in subsection 1.4.4, and its extended work is proposed in **Papers D, E,** and **G**.

- **Paper C (see Appendix C):** This preliminary study investigated how an AI model can provide excitation for the other human partner during a dyadic text-based conversation. As a first step, we proposed a Neural Emotion Hawkes Process (NEHP) based on the Neural Hawkes Process [49] for predicting the future emotional dynamics of the other conversation partner. Moreover, we hypothesized that NEHP could facilitate learning of different intervention-based emotional consequences of different excitation strategies. Thus it would allow for goal-directed excitation behavior by integrating with emotional chatbot agents [50].

  Our preliminary results in this paper showed promising emotion prediction accuracy over future conversation turns. Furthermore, the proposed model captured meaningful excitation without training on explicit excitation ground truths, unlike what was proposed in recent studies [51]. The outcome of such excitation simulation could facilitate a learning environment of emotion dynamics on SM conversations. The latter can construct another learning layer to mitigate polarization through positive and factual-based emotional intervention by LA-based chatbots where the automaton learns optimal emotional expressions that would mitigate the negative emotions in a conversation sequence. Therefore, this paper can contribute to mitigating the harshness in the diffusion model environment, which can be further controlled to mitigate misinformation by the algorithms and architectures proposed in **Papers B, D, E** and **G**.

- **Paper D (see Appendix D):** This paper extended the work done in **Paper B**. It proposed a generic misinformation mitigation algorithm that is robust to different social networks' misinformation statistics, allowing a promising impact in real-world scenarios. Further, a novel loss function was proposed to ensure fair mitigation among users.

  The paper further answered RQ1, RQ2, and RQ3 in subsection 1.4.4 and also has relevance with **Papers B**, **E** and **G**. The paper extended the work done in **Paper B**, established the basis for the proposed software in **Paper E** and the further improvement made in **Paper G**.

- **Paper E (see Appendix E):** This paper proposed a Misinformation Mitigation Storytelling Simulation software. The software has an interactive information diffusion visualization technique to facilitate the informative visual evaluation of different consequences of applying different incentives to a truth campaign. The latter could be beneficial to emergency responders to help in the decision-making process during an infodemic that threatens the societal order [52].

  The work done in this paper provided more practical implementations for the tasks tackled in RQ1, RQ2, and RQ3 as demonstrated in subsection 1.4.4, and has relevance to **Papers B, D,** and **G**.

- **Paper F (see Appendix F):** This paper proposed a self-supervised learning scheme inspired by the self-correction and interpretability of a standard TM [45]. The proposed architecture used a twin of Label-Critic Tsetlin Automaton (TA) [53]. The Label-TA learns the individual samples' correct labels guided by a self-corrected TM logical clause. At the same time, the Critic-TA validates the learning and approves the Label-TA reward.

  The empirical results on synthetic and real data showed promising capabilities for self-supervised learning and interpretable clustering. Hence, this paper proposed a preliminary model that can help self-learn annotations of categories such as SM temporal activities used to train and estimate the parameters of a MHP. The interpretability and self-learning capabilities provide transparency in judging users' generated content and allow for an independent misinformation migration pipeline. Therefore, this paper partially answered RQ5 as introduced in subsection 1.4.4.

- **Paper G (see Appendix G):** This paper concluded our dissertation and proposed a novel representation of temporal users' activity on SM. We further embed these in a knapsack-based mitigation optimization approach. The optimization task was to find ways to mitigate political manipulation by incentivizing users to propagate factual information.

  In this study, we have created PEGYPT, a novel Twitter dataset to train a novel multiplex diffusion model with political bias, societal engagement, and propaganda events.

  The collected dataset and the novel approach align with recent theoretical findings on the importance of societal acceptance in information circulation on SM as proposed by Olan et al. (2022) [44]. The paper empirical results showed significant differences from traditional representations, where the latter assume users' exposure to misinformation can be mitigated despite their political bias and societal acceptance. Hence, This paper opened venues for more realistic misinformation mitigation with extended analytical capacity.

  The work is considered an extension to the earlier proposed methods in **Papers B, D,** and **E** and answered more maturely RQ1, RQ2, and RQ3.

Chapter 2

Theoretical Background

In chapter 1, we gave an overview of the scope and contributions of this study. We introduced the concept of controlling an information diffusion model in a simulation-based optimization framework to find the optimal incentives for a truth campaign on Social Media (SM).

In this chapter, we go further and provide more details on the theoretical background on the relevant topics in the scope of this dissertation. section 2.1 gives an overview for the problem of deception on SM and the different approaches adopted in the literature, emphasizing the truth campaign as the followed approach in this study. section 2.2 explains the different techniques of modeling information diffusion on SM, where more details are given on the Multivariate Hawkes Process (MHP) as our utilized method for diffusion modeling. Eventually, in section 2.3, we demonstrate the fundamentals and relevant learning schemes for the Learning Automaton (LA) as our adopted control model approach.

2.1 Misinformation on Social Media

2.1.1 Key Terminologies and Approaches

SM enables users to be connected and interact with anyone, anywhere, and anytime. Moreover, user engagement over information such as online news articles, posts, and comments carries implicit judgments from the users to the propagated information. Unfortunately, not all judgments are fair on such online platforms [54], which opens a window for persistent manipulative content that people might believe and further propagate. That motivated researchers to observe users' activities on an unprecedented scale to investigate the harmful effect of unauthentic content on SM.

The definition of unauthentic content on SM has evolved due to the increased complexity of such social problems and how recent technological efforts have progressed. To this end, there are different definitions and demonstrations of this phenomenon. For instance, the spread of fake news on SM has been considered the intentional dissemination of false information in online news articles [55]. The latter issue is also known as disinformation [56]. Hence, a common approach is disinfor-

mation detection, where the goal is to classify false information early by evaluating the truthfulness of the claims mentioned in a news article or shared online content.

Online content can have multiple modalities, such as text, image, and video. The latter features help create what is known as content-based disinformation detection [57]. That technique employs supervised machine learning models, which are trained on features from single or combined modalities [58] to be able to classify future content. Alternatively, context-based detection models [59] utilize co-occurrence-based word embeddings [60] for the content, and user-driven societal engagements [61].

Unsupervised machine learning techniques were also integrated with supervised machine learning classifiers for propagation-based detection models [62]. The latter approach takes advantage of the echo chamber cycle on SM and investigates network-based structural features such as propagation trees [63].

Other research work gave attention to a broader scope of the problem, such as rumor detection [64], SM fake accounts detection [65, 66], and the unintentional spread of false content, a.k.a "misinformation" [56]. In rumor detection models, the task is to identify rumors which are statements whose veracity is not quickly or ever confirmed. As in disinformation detection models — content-based, context-based, and propagation-based features were investigated for the problem of rumor detection. However, since the nature of rumors makes them have no reliable source to verify their veracity, it is usually harder to collect annotated data to train supervised machine learning models for rumor detection [67].

Another potential cause of incorrect or inconsiderable information is SM fake accounts. The latter can be categorized into three categories: Trolls [68], Cyborgs [69], and Social Bots [70]. Trolls are deceptive accounts run by a human who aims to motivate others to react emotionally to agree with deceptive content. Cyborgs are semi-automated accounts that objectively try to spread fake information. Social Bots are usually run by a computer program and used in many cases, like advertising and fake news circulation. In general, fake accounts detection models investigate SM users' online engagement and profile features such as the number of followers, friends, and mentions [71].

The criticality of misinformation and disinformation on SM is relative to the context of such an issue. For instance, the large-scale manipulation carried out during political events is one of the greatest threats to social justice and democracy [72]. That was observed in the so-called Cyber Army of Russian Trolls attack on the U.S.A 2016 presidential elections [65]. Moreover, the problem of misinformation is a more persistent issue. That is because a biased group would confirm and embrace false content that has been intentionally circulating at some point, and they unintentionally spread the underlying idea on the network [56]. Thus, this study focused on the problem of misinformation in the political context.

2.1.2 The Truth Campaign Approach

The risk of violating freedom of speech is considered an ethical concern [19] about machine learning classifiers (i.e., the misinformation detection approach) since one false positive violates human rights if the associated account or content is suspended accordingly. Therefore, alternative approaches were encouraged. In this manner, recent studies showed that exposing social network users to factual information would significantly mitigate the effect of misinformation [20, 30, 73]. Hence, a mitigation approach that allows for ideological debates, democratization of combating misinformation, and reducing the risk of violating human rights, arose.

In common mitigation methods, the task is to boost the exposure to valid information. There are different techniques to achieve such a boost on SM [74]. For instance, influence blocking [75] is a technique to target a subset of users on the network to block their access to misinformation. In the latter technique, this subset of users is prioritized to have minimal access to misinformation, which is assumed to minimize the risk of misinformation exposure over the whole network.

Truth campaigning is another common mitigation technique, where the task is to learn an optimal mitigation strategy by incentivizing a group of users to circulate factual information. An example of incentivization can be learning and then delivering personalized verified news articles to suit users' reading preferences [76].

Another example of incentivization is to learn about the number of incentives per user that would acquire the latter to accept propagating the verified information on the network [20]. Intervention-based techniques are followed to learn such optimal incentives. That is, by intervening with users and learning from their online activities such as their responses to a particular content type, the mitigation model ensures a maximal delivery of authentic content to the network by targeting suitable users. The individual learned incentives could then characterize the user readiness to help boost the factual information when asked by authorities or platform moderators, for instance.

Typically, in an intervention-based truth campaigning approach, an intervention with the network users is modeled through the Reinforcement Learning (RL) framework to learn the optimal mitigation policy [20]. The intervention procedure allows the RL agents to learn about the users' online activity. The latter is simulated with an information diffusion model where Point Processes are commonly utilized [77].

This study adopted the intervention-based truth campaigning approach and proposed novel RL algorithms and architectures to learn the campaign's incentives. section 2.2 and section 2.3 give more details on the theoretical background of information diffusion modeling and our proposed RL method, respectively.

2.2 Information Diffusion Modeling

2.2.1 Definition and Techniques

Diffusion processes have received interest from various scientific fields due to their ability to characterize a broad range of real-world scenarios. For instance, information dissemination in both offline [78] and online [79] situations, the transfer of stress in earthquakes [80], and the spread of disease in a population [81].

Information diffusion modeling on SM simulates the dynamics of a social network by how the information propagates and users engage and influence each other. According to the relevant literature [82], there are three main components in a diffusion process on SM: information, users' activities, and social vectors. These three main components influence each other in the process. For instance, the social vector represents the influence between the adjacent users on the network and then dictates the patterns of users' activities. Consequently, the information will propagate on the network according to these influence relationships. However, other external factors also dictate the propagation — for example, the motivation from the physical world relationships and users' biological and political perspectives, such as gender and political preferences [83].

Further, the time factor is essential to information propagation since the moment a particular user makes a social action on the network (e.g., retweet, comment), cascaded information is constituted [84] (e.g., further comments, retweets).

There are five main steps to building an information diffusion model on SM [82]. We explain below each step and its level of importance to obtain a realistic diffusion model.

2.2.1.1 Data Collection

The first step to obtaining a diffusion model is to get sufficient representative data. The data can be collected from SM platforms through their provided APIs (e.g., Twitter API [85]). An alternative channel for obtaining social network data is through published and available datasets [32, 33, 34, 35, 36, 17]. The collected data can be processed into three categories: cascades, social topology, and corresponding user information.

The cascades are the propagation paths of information through different users. That data category acts as the ground truth for any diffusion model since part of the data could be separated for evaluation purposes. The users' topologies represent users' relationships on the network, such as friendship and following. Topologies are also fundamental to any diffusion model as they affect the speed of the information diffusion [86], and they are crucial features to help predict the diffusion. Corresponding users information such as their profile information and generated content semantics are also essential to filter candidate users based on learning their content preferences. Thus, the combination of these three categories improves prediction accuracy, mainly when utilized in state-of-the-art popularity prediction diffusion models based on Deep Learning and Natural Language Processing architectures

[87].

As mentioned earlier in subsection 2.2.1, the temporal information of social network data should be considered. Therefore, at least cascade data must be indexed with a finite set of discrete time windows to infer the diffusion speed and predict the correct time windows of a particular cascade.

There are a variety of published SM datasets which include some of the mentioned data categories [32, 33, 34, 35, 36, 17]. However, as introduced in section 1.3, these available datasets are not enough to establish a realistic diffusion model for the problem of misinformation mitigation.

2.2.1.2 Diffusion Mechanism Analysis

Analyzing the mechanism of a diffusion process is crucial to understanding the essential features needed for modeling such diffusion. The analysis can be viewed as three main focus points. First is the social vector, where network structures and user relationships play a fundamental role in predicting information diffusion. An example of the latter is a link prediction model which simulates the future incidents of users' associated social vectors [88].

The second focus feature is users' activities (e.g., retweeting, commenting). The latter feature is very important since it tells about the societal activities on the network. However, users' activities are complicated to learn and predict since the human brain's decision-making process is highly challenging [38]. Therefore, modeling these activities in a realistic manner is one of the main goals of our study.

Finally, the third feature to analyze is the scope of information to represent in the modeled diffusion. That involves semantic analysis to filter out the content type that should be predicted [89].

2.2.1.3 Feature Extraction

As a result of the diffusion mechanism analysis step, there will be critical needs for more specific features. For example, in an online content popularity prediction, content-based features such as statistical insights about used terms frequency [90], and deep semantics of the content such as sentimental and emotional features [91, 92] are crucial. While for link prediction diffusion models (e.g., predicting the new relationships between two users), structural features are essential. Some examples of such structural features are degree distribution [93], density [94], and cascade structures which refer to propagation trees and diffusion hops [95]. Some other techniques utilize user-related features, such as predicting users' content preferences [96].

Further, as a common foundation for all diffusion models, temporal features such as the time windows between activities or user actions' timestamps are required. Temporal features can be constructed from two main types: sequence-based features and statistical-based features. Sequences of counts of particular event type (e.g., retweets, comments on a particular hashtag) over defined time windows can construct the temporal sequential features [95]. Statistical features are secondary

processing of sequential features and aim to find potential diffusion laws to provide mathematical expressions for a diffusion model [97]. In this dissertation, we focused on sequence-based and statistical-based features in our diffusion processes. In subsection 2.2.3, we discuss in more detail the theoretical foundations of our diffusion method.

2.2.1.4 Establishing the Model

An established information diffusion model can be described in terms of its approach. The different extracted features and targeted diffusion mechanisms define the particular approach. Therefore, the two main approaches to information diffusion modeling can be divided into time-series, and data-driven approaches [82]. Time-series approaches provide diffusion models in terms of mathematical expressions and hypothetical diffusion laws driven by the sequential and statistical elements of the time-series features. In that approach, likelihood maximization and simulations are studied over the time-series sequence to provide estimated parameters and prediction functions [98], respectively. In data-driven, one does not need to give explicit expressions or fixed parameter values since data-driven approaches take advantage of machine learning algorithms to predict the diffusion by automatically learning from features [99].

There are three typical possible outputs from diffusion models [82]. For instance, the model can output volumes, individual information, and propagation relationships. Volume models only capture the change in the diffusion volumes over time [100]. Individual models investigate, and output details about influenced users [101]. The propagation relationships models demonstrate the information paths within the inferred influence from individuals [102].

For the scope of this dissertation, we utilized a volume-based information diffusion model since our main focus was to boost the amount of true information to combat the false ones. In subsection 2.2.3, we give more details on the adopted volume-based information diffusion approach.

2.2.1.5 Model Evaluation

The evaluation of information diffusion models can be understood in the scope of how typical data-driven machine learning models are evaluated. For example, the task of a diffusion model can be either a classification or a regression task [82]. Therefore, we can utilize the convenient evaluation metric according to the desired output. However, evaluating an information diffusion model is considered a challenging task [82]. chapter 3 demonstrates how this study contributed to a more convenient evaluation of an information diffusion model for the task of misinformation mitigation on SM.

2.2.2 Point Processes as Diffusion Models

Point Processes [103] are stochastic processes that can be utilized to construct the volume-based diffusion models described in subsubsection 2.2.1.4. The core feature of such diffusion models is the temporal aspect of the volumes of information propagation caused by users' activities. Therefore, the diffusion process can be considered a stochastic process with a Poisson distribution [104] where the discrete probability distribution calculates the probabilities of a particular volume (i.e., count) of independent activities occurring in discrete-space partitions or discrete-time intervals.

In the context of SM, non-exhaustive examples of users' activities can be tweets, mentions, and comments — with their associated discrete-time intervals $\mathbf{T}$ of occurrence. In this manner, a point process can estimate the diffusion contribution (i.e., counts) made by each user's activities.

A Poisson process is the simplest form of a point process. In the latter, the probability of a random variable k that represents a diffusion count is calculated inclusively within the discrete-time intervals $\mathbf{T} = \{t_1, t_2, ..., t_r\}$. Moreover, the diffusion activities are independent of each other, and it is assumed that there is no external influence on them.

The Poisson probability distribution of an activity x occurs within a discrete-time interval t_s can be calculated as given in Equation 2.1. Where k is a particular activity count we calculate for, and λ is the expected count. For r discrete-time intervals in $\mathbf{T}$, Equation 2.1 is applied for each $t_s \in \mathbf{T}$.

$$P_x^{\mathbf{t_s}}(k) = \frac{\lambda^k \exp^{-\lambda}}{k!}. \tag{2.1}$$

λ is also called the intensity rate, and when the intensity rate is constant over different time intervals, the process is called a homogeneous Poisson process. Alternatively, a non-homogeneous Poisson process has a variable intensity rate λ over time [105]. Figure 2.1[1] demonstrates an example of three independent activities with their calculated Poisson distribution over a discrete-time interval t_s.

Point Processes have been utilized as diffusion models for a wide range of applications [103] since they realize any generic bursty behavior with a sequence of point events on the real line, either in time or space. To characterize more complicated real-world scenarios, an extension to Poisson Point Processes was proposed [106]. In subsection 2.2.3, we demonstrate how the non-homogeneous Poisson Processes constructed the fundamentals of Hawkes Processes (HPs), where more realistic physical world patterns have been considered [106].

2.2.3 Multivariate Hawkes Processes

According to Equation 2.1, Poisson Processes do not consider the external motivation and historical occurrence of activities when estimating the probability of their counts. Alternatively, HPs offer the excitation mechanism where the history of activities propagation patterns influence the future occurrence of the latter [106].

[1] The figure was legally downloaded from https://commons.wikimedia.org/w/index.php?curid=9447142

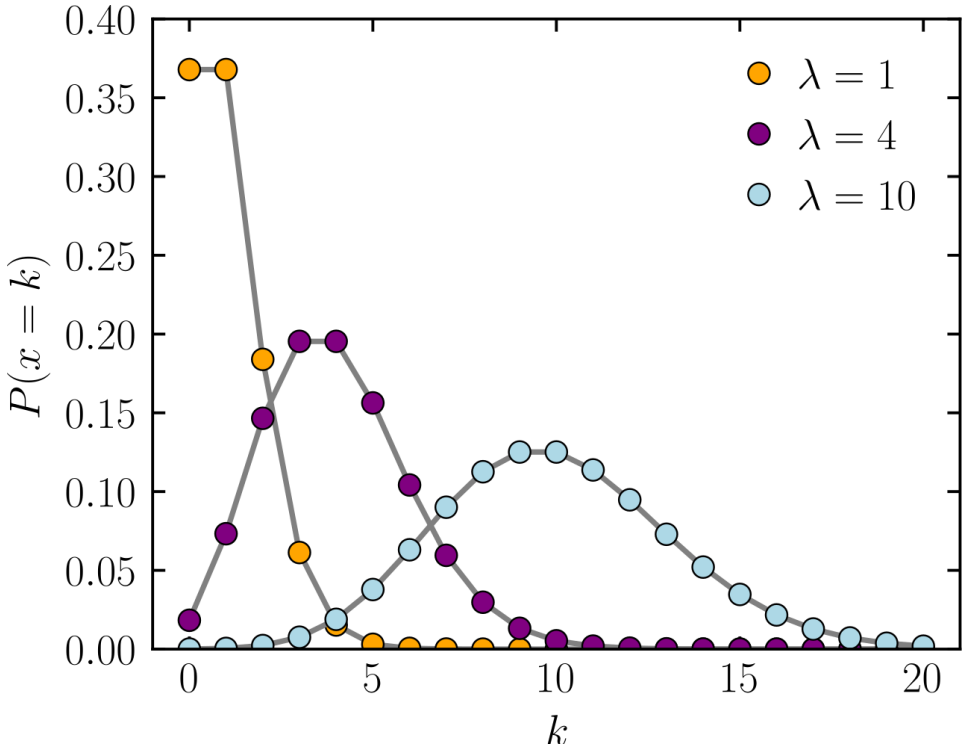


Figure 2.1: An example of three independent activities with their calculated Poisson distribution over a discrete-time interval t_s

Furthermore, HPs estimate an external influence parameter to model external motivation of activities which is more realistic.

Let i denotes a particular activity of user i on SM, for each user i, let $\lambda_i(\mathbf{T})$ denotes the intensity of a single HP within the discrete-time intervals $\mathbf{T} = \{t_1, t_2, ..., t_r\}$. Similarly, let $\mu_i(\mathbf{T})$ denotes an external influence parameter over user i to do the activity, and $g_i(\mathbf{T})$ be a kernel function that estimates temporal influence between user's activity i and other n users' activities, where the temporal influence is estimated from the history of all activities inside the social network that occurred before t_1. We refer to the historical time intervals by $H^{\mathbf{T}}$. To this end, a HP can be defined by the intensity $\lambda_i(\mathbf{T})$ which is conditioned over $H^{\mathbf{T}}$ and estimates the count for activity i within $\mathbf{T}$.

The above mechanism is also called the mutual-excitation Point Process and is defined according to Equation 2.2 and Equation 2.3. The mutual excitation represents the estimated internal influence relationships inferred from the activities history and denoted by $A \in \mathbb{R}^{n \times n}$ as the internal influence matrix, where the matrix entry $a_{ij} > 0$ indicates if activity j excites activity i with the excitation weight value from a_{ij}, and $a_{ij} = 0$ indicates no influence at all. Finally, $0 \leq w \leq 1$ is an exponential decay factor that characterizes the decay of the influence while going forward on time over the discrete-time intervals $\mathbf{T}$. The optimal value w can be estimated through a grid search.

$$\lambda_i(\mathbf{T}|H^{\mathbf{T}}) := \mu_i(\mathbf{T}) + g_i(\mathbf{T}, A), \tag{2.2}$$

$$g_i(\mathbf{T}, A) = \sum_{s=1}^{r} \sum_{j=1}^{n} a_{ij}\ w\ \exp^{-wt_s}. \tag{2.3}$$

To predict diffusion intensities for n users on the social network for a particular activity, n HPs to be established. One significant advantage of HPs is how they provide flexibility when defining the representations of the activities. That can be observed in a wide range of HP-based proposed diffusion models where activities could be of the same type or marked for different interdependent types [107, 108].

We demonstrate in chapter 3 how we proposed a diverse range of activity categories and simulated each category as a group (i.e., HPs) of activities where the latter relate to the same information category and have mutual excitation over user level as described in Equation 2.3. Then, we proposed a higher level of excitation between these groups. For example, HPs that modeled misinformation activities from all network users can be excited by HPs that modeled political bias activities from the same network users.

When multiple activities exist, the process is called Multivariate Hawkes Process (MHP) [41]. Figure 2.2 demonstrates a MHP for a network of three users (i.e., three activities) and discrete-time intervals $\mathbf{T} = \{t_1, t_2, t_3, t_4, t_5, t_6\}$, where each interval has a window of 5-days. The three colors of circles represent three marks, where each color is associated with an individual user, and each circle is an activity occurrence. As observed in Figure 2.2, for each user, intensities vary over time. In the latter example, these variations represent the time-decayed predicted counts of the information propagation based on the inferred parameters for the MHP, namely the time-decay scalar w, internal influence matrix $A \in \mathbb{R}^{3\times3}$, and the external influence vector $\mu \in \mathbb{R}^3$. In such parametric MHP, these parameters are constant over $\mathbf{T}$.

2.2.3.1 Excitation Between Activities

In a simple one-dimension HP, a historical sequence of activities from the same single type influences the occurrence of these activities over time [106]. When that happens, the process is called self-excitation HP. The self-excitation phenomenon has been studied in the literature on diverse application domains [109]. There are many self-excited phenomena. For instance, many musical instruments have self-excited vibrating [109], and our human body is self-excited as in human-structure interaction when footsteps are altered subtly, giving a net damping effect on the body structure [110]. Further, social and political activities can be explained as self-excited vibrations [109]. However, more complex real-world problems are usually multi-dimensional. In this manner, in a MHP, the influence and dependency through the activity's history have mutual excitation rather than self-excitation. The mutual excitation exists between the dimensions of the MHP. In the context of this dissertation and the application of SM, these dimensions can represent the social network users.

As discussed in subsubsection 2.2.1.3, the statistical-based extracted features for the MHP mean finding and establishing diffusion laws such as the mutual excitation matrix A. One more critical aspect of the diffusion of temporal information is how

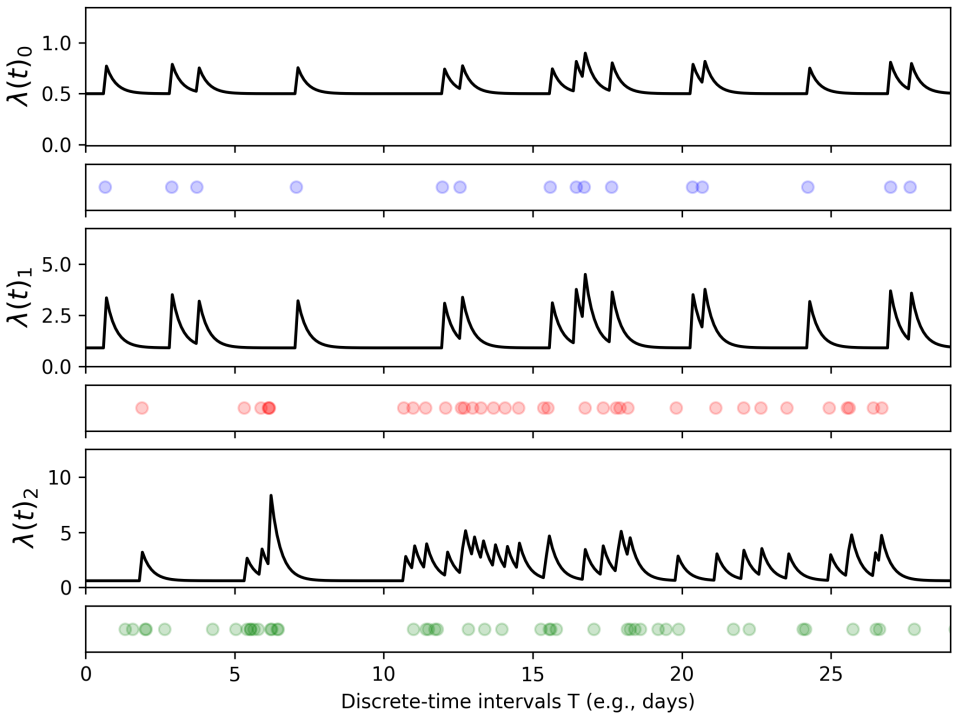


Figure 2.2: A MHP with the dimensionality $6 \times \mathbb{R}^3$

the mutual excitation and its consequent propagation decay over time. For instance, as time elapses, online content loses its popularity in SM [111]. Another example is epidemics, where a pathogen loses its infectiousness over time [112].

There are different time relaxation functions to model the time-decay characteristic. Four primary time relaxation functions [103] exist for Point Processes-based diffusion: exponential [113], power-law [114], lognormal [115], and gamma decay [116]. Exponential decay defines the process of reducing an amount by a consistent percentage rate over a period of time. For SM dynamics, the exponential decay function is the most commonly utilized relaxation function and was utilized in our study.

2.2.3.2 MHP Fitting and Evaluation

One of the common approaches to establish the MHP to predict counts of activities in future discrete-time intervals is through parametric mathematical expressions as demonstrated in Equation 2.2 and Equation 2.3. That means we first must estimate the parameters μ as the vector of baseline intensities across the users' space and A as the squared influence matrix of all users. Further, a good value for the decay factor w should be explored to ensure realistic temporal influence decay behavior.

One of the most traditional and established approaches to estimating the MHP parameters is what Ogata (1988) proposed [117]. The latter work calculated the maximum log-likelihood over the parameter space $\theta = \{(\mu_1, A_1), (\mu_2, A_2), ...\}$, which

was obtained numerically by a standard nonlinear optimization technique. It was then possible to evaluate which of θ possible values provide the best fit for the process data. Finally, the expression in Equation 2.4 was proposed to define the conditional likelihood.

$$\log L(\theta) = \sum_{s=1}^{r} \log \lambda(t_s; \theta) - \int_1^r \lambda(t; \theta) dt. \tag{2.4}$$

Alternatively, there were different proposed techniques for MHP predictions where a non-parametric approach was adopted. For example, Mei and Eisner (2017) [49] modeled streams of discrete events in continuous time by forming a neurally self-modulating MHP. In the latter method, the intensities of multiple event types evolved according to a novel continuous-time Long short-term memory neural network [118].

The advantage of utilizing a non-parametric approach to diffusion models is extending the model's capacity so that no constant parameters control the diffusion, which could be more realistic. However, when the required task is to control a diffusion model to optimize some propagation patterns, a parametric diffusion model would be more beneficial since it is easier to control and evaluate different simulated scenarios through its parameters. The evaluation of the different approaches to MHP predictions is not the scope of this dissertation, and we adopted the maximum log-likelihood approach for a parametric MHP in our research study.

Figure 2.3 demonstrates a generic structure of representing temporal activities category's associated samples from a SM collected data over a number of days and hours — to feed in the MHP. As observed in Figure 2.3, the temporal activities per user are grouped into ordered discrete-time intervals. Then, the grouped timestamps per interval and user are fed into a MHP parameter estimation based on Equation 2.4. The parameter estimation means inferring the influence and diffusion laws that govern these historical user-generated counts.

Since the final target is to predict diffusion volumes, the MHP intensities calculated through Equation 2.2 and Equation 2.3 are interpreted as predicted activities count through the MHP modified thinning simulation algorithm introduced by Ogata (1981) [119]. We adopted the same diffusion volumes simulation technique in our study.

To evaluate the MHP predictions, the predicted counts for an activity for each user are compared to the relevant real counts from the test dataset. Therefore, and as shown in Equation 2.5, the absolute difference average error ϵ is usually calculated [25] to measure how close to reality a MHP simulation is. Where n is the number of users and $N^{\mathcal{H}}$, $N^{\mathcal{R}}$ represent the counts of the arrived activities from MHP predictions and real data, respectively. The evaluation is made independently per each discrete-time interval t_s.

$$\mathcal{E}_{t_s} = \frac{1}{n} \sum_{i=1}^{n} |N_i^{\mathcal{H}}(t_s) - N_i^{\mathcal{R}}(t_s)|. \tag{2.5}$$

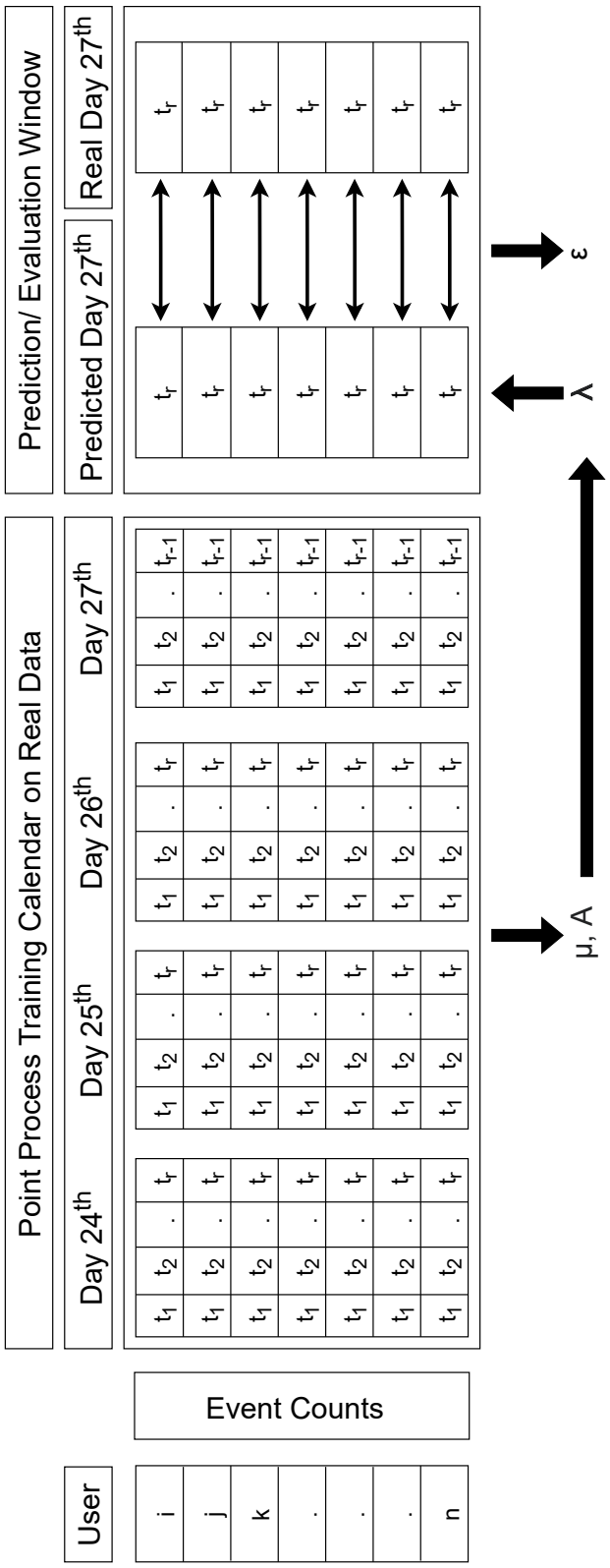


Figure 2.3: Feeding the MHP with temporal activities' associate samples from a SM dataset. The same structure applies to other activity categories

2.2.4 Multiplex Multivariate Hawkes Processes

So far, we previewed MHP as a model that characterizes the information diffusion dynamics over multiple users and their activities on SM. However, the interactions between the social network nodes (i.e., users) encompass more diverse and complex hidden propagation and excitation patterns. Thus, the inference of these hidden influence patterns becomes a challenging task. We believe a typical mutual-excitation MHP is not enough for the problem of misinformation propagation analysis. Therefore, this dissertation extended the traditionally adopted MHP-based control approach to multiplex MHPs-based control.

A multiplex diffusion model can be viewed as different interdependent groups of MHPs. These groups are the different information categories, and inside each group, there is a diffusion model that predicts the generated activity counts of the network users for the associated information category. For instance, Starnini et al. (2016) [120] measured and modeled the characterization of inter-group correlations to investigate how some activities in one information group excite societal interaction in another. Further, Sun et al. (2018) [121] proposed a multiplex diffusion model incorporating multiple parametric MHPs for different topic models groups. In the latter approach, they connected the diffusion between the topic groups so that a change in one topic propagation count would excite the propagation of another topic. Therefore, we followed a similar approach in this study and extended these proposed methods for the MHP control scenario. In this manner, we studied how a single control (e.g., incentivize for a truth campaign) of a user's activity over one information group will excite other users on all information category groups.

chapter 3 explores in more detail the proposed Multiplex-Controlled Multivariate Hawkes Processes (MCMHP) for the interdependent problems of societal bias, societal acceptance, and information veracity as our proposed three main diffusion groups on SM.

2.3 Learning Automaton

2.3.1 The Learning Problem

According to psychologists, a learning system is a system that changes its performance desiring the accomplishment of a specific goal [122]. Such change characterizes the ability to improve behavior over time by observing past experiences and learning from them [40]. A learning task has two parts: the learning system and the environment.

There are three main approaches to designing an artificial learning system: Supervised Learning [123], Unsupervised Learning [124], and RL [25]. In the latter approach, the learning system learns from interacting with its environment through trial and error by waiting for either a positive or a negative reward (a.k.a "penalty") as a consequence of the system's trial [125]. RL problems can be classified into two categories [126]: sequential and non-sequential. In sequential problems, the task is

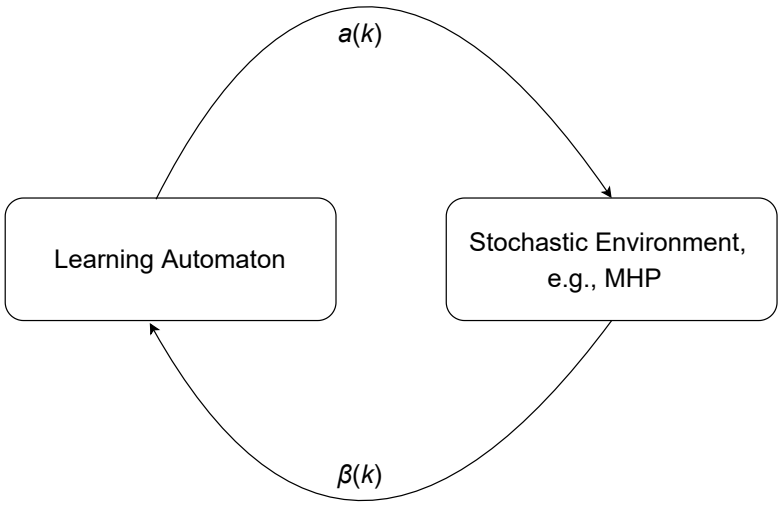


Figure 2.4: The interaction of a LA and an environment. For instance, a MHP

to maximize the total reward of sequential actions made by the system [127]. On non-sequential RL, the task is to learn the optimal action or system state among the action or state space, respectively.

The non-sequential RL has been studied extensively in the field of LA [128]. LA is suitable when the environment dynamics are unknown [128] to the learning system (e.g., a stochastic process). Further, LA is an easy-to-implement algorithm compared to Q-learning algorithms [129], where the latter is a well-known technique to learn from a stochastic signal. Therefore, LA is the learning system adopted in this dissertation to learn optimal incentives for a misinformation mitigation strategy over an information diffusion model as its environment (e.g., MHP).

Figure 2.4 shows the interaction of a LA and a MHP-based environment, where a and β are the selected action and its consequent reward signal at interaction step k, respectively.

2.3.2 The Learning Environment

The mathematical definition for a LA environment can be expressed in the form of the quintuple $\langle \mathcal{X}, \mathcal{A}, \underline{\beta}, F, Q \rangle$, where $\mathcal{X}$ is a set of context vectors, $\mathcal{A}$ denotes the set of inputs, $\underline{\beta}$ is the set of reinforcement signal values, $\{F = f(a, \mathbf{x}) | a \in \mathcal{A}, \mathbf{x} \in \mathcal{X}\}$ represents the set of probability distributions over $\underline{\beta}$, and Q denotes the probability distribution over $\mathcal{X}$.

Both F and Q are assumed to be unknown, where the main task of an automaton toward its environment is to eventually estimate F by learning the optimal input a for the associated context $\mathbf{x}$. $a(k)$ and $\beta(k)$ define an automaton's action (or state) and its consequent reinforcement signal in the interaction step k, respectively. k is a step in the discrete-time interaction indices $\{0, 1, 2, ...\}$.

The set of probability distributions F dictates if the environment is stationary [130] or non-stationary [131]. For example, in a stationary environment, the probabilities in F are fixed over time, while in the non-stationary, it varies with time.

Further, the nature of $\underline{\beta}$ determines three possible types of the environment [132]: P-, Q-, and S-models. In a P-model environment, $\beta(k)$ could have two possible discrete output values that represent a reward or a penalty (e.g., {0, 1}). In a Q-model environment, $\beta(k)$ output is finite discrete categorical symbols (e.g., {-1, 0, 1}). In S-model environments, $\beta(k)$ outputs a finite number of values in the interval $[0, 1]$.

In this dissertation, we evaluated both a P-model environment for the proposed LA interaction algorithms, where the environment was a non-stationary MHP.

2.3.3 The Automaton

A state-based automaton can be defined by the quadruple $\langle \mathcal{X}, \underline{\beta}, \Phi, \mathcal{F} \rangle$, where $\mathcal{X}$ is a set of context vectors, $\underline{\beta}$ is the set of reinforcement signal values, Φ denotes the set of internal states of the automaton and equivalent to the input set $\mathcal{A}$ of the environment (see subsection 2.3.2). Finally, $\mathcal{F} : \mathcal{X} \times \Phi \times \underline{\beta} \rightarrow \Phi$ denotes a state transition function, where the automaton's state $\phi(k+1) \in \Phi$ is determined based on the context vector $\mathbf{x}$, step signal $\beta(k)$, and the automaton's state $\phi(k) \in \Phi$.

In a structure of an automaton that is represented by its state space, the automaton task is to learn the optimal state value that will ensure an environment positive feedback β when the automaton is in that state with the context $\mathbf{x}$. In general, an automaton is considered to be deterministic if its state transition probabilities when either rewarded or penalized are equal to One. Otherwise, the automaton is considered stochastic. There have been various proposed learning schemes and architectures to implement the transition function $\mathcal{F}$. We highlight the most common methods to our study in subsection 2.3.4 and subsection 2.3.5. In general, the utilization of LAs has the below benefits to many applications [126].

1. The automaton can learn optimal decisions without any prior knowledge of its environment.

2. When a network of LAs is utilized, it can facilitate a multi-agent and distributed systems framework for a more complicated task.

3. In general, an automaton's structure can be simple and easy to implement, besides its few mathematical operations, making it a practical, lightweight system for real-time applications.

4. It has been shown that LAs are optimal in single, hierarchical, and distributed structures.

2.3.4 Common Architectures and Learning Schemes

In the following, we demonstrate the current LA-based techniques relevant to our proposed LA-based architectures and algorithms. These existing techniques and architectures can be considered fundamentals to the contribution of our study, which we discuss further in chapter 3. We gradually begin with the most straightforward architecture and advance to more complicated ones.

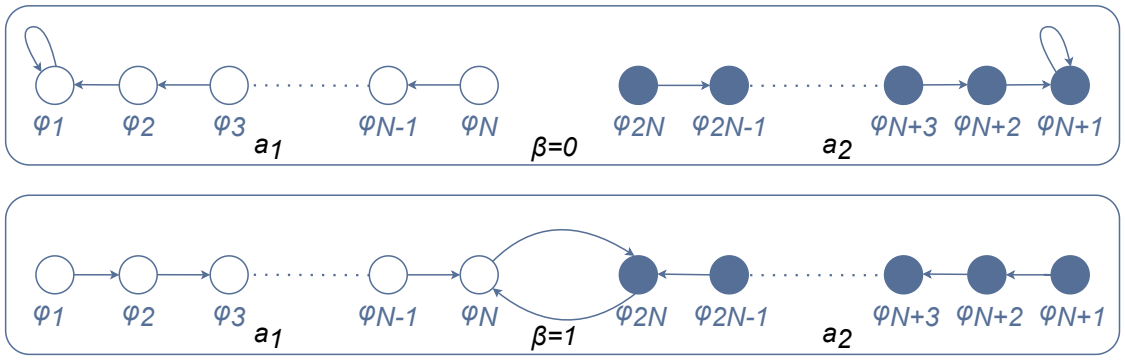


Figure 2.5: The state transition graph for a TA

2.3.4.1 Tsetlin Automaton

The Tsetlin Automaton (TA) [133] operates in a P-model environment (see subsection 2.3.2) and has a simple structure with two actions, each having N associated states. Hence, the overall number of states is $2N$, and the actions can be denoted as a_1 and a_2.

Figure 2.5 shows the structure of a TA, where the state space Φ is divided into two partitions, where each of the latter is associated with one action. The automaton selects its action $a(k)$ at step k based on the current state space partition it is in, and then the state transition is made based on the environment feedback $\beta(k)$.

The initial state of the automaton is selected according to a uniform probability distribution over the states $\{\phi_N, \phi_{2N}\}$, to guarantee a fair final action selection by starting in the middle of the state space structure.

In general, an automaton's state structure acts as a memory to remember the best action selection after some interaction steps. If a current state is far away from the middle states, it means that the current state is more favorable. Hence, the associated action with that state partition is more likely to be the optimal action. For instance, if the state ϕ_N is a current state, the automaton will select the action a_1. Then, the environment evaluates the selected action and sends the feedback signal β. If the signal was a reward (e.g., $\beta = 0$), the next state will be ϕ_{N-1}, and a_1 will remain the action to select in the next interaction step. Otherwise (i.e., a penalty where $\beta = 1$), the automaton moves toward the direction of the opposite state partition, and the next state will be ϕ_{2N}, giving a chance to explore the other action a_2. Figure 2.5 shows such state transition scenarios on both reward and penalty, when being in any state.

In a learning scheme of an automaton, the state transition is usually probabilistic to give the automaton the capacity to explore non-visited states or actions. In this manner, the TA is categorized as a fixed-structure-LA [134], which means the state transition probabilities among its states are not varying over the interaction steps.

Fixed-structure-LAs are helpful to avoid unnecessary complexity in the learning algorithm, for instance, when an action is rewarded, and it is unnecessary to explore other actions [135]. Furthermore, Fixed-structure-LAs are also helpful when learning in environments that dictate consistent relationships between the automaton's

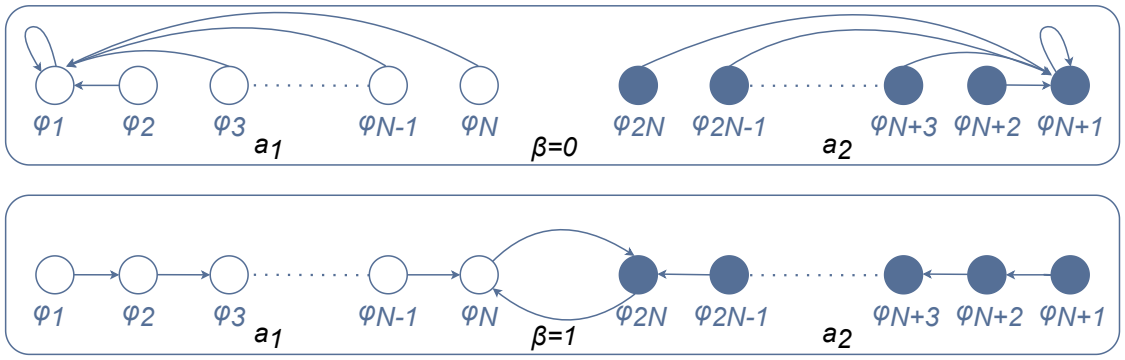


Figure 2.6: The state transition graph for a KA

states. The latter can be demonstrated when applying Fixed-structure-LAs on the so-called Static Mapping Problem [136]. On the contrary, in chapter 3, we discuss how states can have inconsistent relationships with each other in a truth campaign optimization, where a state ϕ_3 would be better than its preceding state ϕ_2 in the interaction step k_i, while ϕ_2 could become better than ϕ_3 in another interaction step k_m. The latter usually occurs because of the dynamics of information cascades on SM [137] where learning one optimal high-value state for one user could make having other high-values of states for other adjacent users unnecessary.

2.3.4.2 Krinsky Automaton

The Krinsky Automaton (KA) [138] is similar to the TA except in how the former performs state transition when rewarded from the environment. Figure 2.6 shows the structure and state transition mechanisms of the KA, where the automaton's state moves directly to the edge state of the associated state partition in case of being rewarded. That means the automaton will always need N successive penalties to switch between its two actions if it was rewarded in a previous interaction step. The KA also belongs to a fixed-structure-LA family since its state transition probabilities are fixed over the interaction steps with an environment.

The different state transition mechanism demonstrated in KA is an example of how different environment settings would acquire various learning schemes to ensure the fastest convergence to the optimal action. For instance, in some cases, the environment is more confident about rewarding, but the uncertainty in giving a penalty is high. Hence, the KA can be a better option since it applies exploration on penalty and moves to an edge state (i.e., most confident state) when rewarded.

2.3.4.3 Variable-Structure Learning Automata

Variable-structured-LAs [139] are LAs with varying state transition probabilities over the interaction steps. These LAs can be represented by the quadruple $\langle \mathcal{A}, \underline{\beta}, p, \mathcal{T} \rangle$, where $\mathcal{A} = \{a_1, a_2, ..., a_r\}$ is a set of r possible actions to select, $\underline{\beta}$ denotes the set of possible feedbacks from the environment, $p = \{p_1, p_2, ..., p_r\}$ represents the action probability vector such that $p_i(k)$ is the probability of selecting the action $a_i(k)$ at the interaction step k.

$\mathcal{T}$ represents the learning algorithm that updates the action probability vector with regard to the particular environment feedback.

To understand the general idea of how a learning algorithm $\mathcal{T}$ works in variable-structured-LAs schemes, Equation 2.6 and Equation 2.7 demonstrate how the action probability vector is updated when the environment responds with a reward or with a penalty, respectively.

$$\begin{aligned} p_i(k+1) &= p_i(k) + \gamma[1 - p_i(k)], \\ p_j(k+1) &= p_j(k) - \gamma[p_j(k)] \ \forall j \in p : j \neq i. \end{aligned} \tag{2.6}$$

$$\begin{aligned} p_i(k+1) &= (1-\psi)p_i(k), \\ p_j(k+1) &= \left(\frac{\psi}{r-1}\right) + (1-\psi)p_i(k) \ \forall j \in p : j \neq i, \end{aligned} \tag{2.7}$$

where r represents the number of the automaton's possible actions, and the reward and penalty parameters $(0 < \gamma < 1)$, $(0 < \psi < 1)$ determine static increase and decrease rates of the action probabilities, respectively.

The tuning of the parameters γ and ψ has a crucial influence on how the learning algorithm works and converges. For instance, when $\gamma = \psi$, the learning algorithm is considered a linear reward-penalty algorithm [140] and is usually abbreviated as L_{R-P}, where getting a reward or penalty on a given action, will result in an equal decrease or increase among other actions probabilities, respectively.

In cases when $\gamma > \psi$, the algorithm is a linear reward-ϵpenalty with $0 < \epsilon < 1$ [141] and abbreviated as $L_{R-\epsilon P}$, where getting a penalty on a given action will result in a smaller decrease of that action probability, compared to its amount of increase when getting a reward.

A particular case is when $\psi = 0$, which means the action probabilities remain unchanged when the environment sends a penalty signal. The latter algorithm is called a linear reward-Inaction [142] and abbreviated as L_{R-I}.

The convenient selection of the parameters γ and ψ values depends on the nature of the task. Further, the smaller the value for γ, the slower the convergence. Therefore, carefully selecting the latter is vital to avoid slow convergence (e.g., γ is too small) or less accurate learning (e.g., γ is too high) [143].

variable-structured-LAs are helpful when the environment has higher uncertainty. In the latter scenario, optimal actions or states require much exploration and verification of what is optimal and what is not.

This dissertation proposed a group of novel variable-structured-LAs with a dynamic rate for updating the action selection probabilities, where that rate changes based on the frequency of actions being rewarded, unlike the fixed values of γ and ψ. chapter 3 demonstrates the proposed variable-structured-LA learning algorithm for the intervention-based misinformation mitigation task.

2.3.4.4 Non-estimator vs. Estimator Algorithms

So far, we introduced different learning schemes of a LA that depend only on the environment's recent response to update its state transition mechanism or the action probability vector. That class of algorithms is known as non-estimator LAs [126]. In the non-estimator schemes, there is no consideration of the long-term environment responses [144]. The latter characterization can be observed in how the action probability vector in Equation 2.6 and Equation 2.7 is updated without memorizing the accumulated feedback from the environment, i.e., only the feedback from the current interaction step $\beta(k)$ is considered for what should happen in next interaction step $k+1$.

Alternatively, estimator algorithms characterize the action probability vector to pursue the action that is currently estimated to be the optimal action. This is achieved by increasing the probability of selecting the action whose current estimate of being rewarded is maximal [145]. The below steps show how an estimator algorithm adds more layers of calculation and pursues the estimation of the maximal rewarded action while updating the action selection probabilities.

Step 1) at the step k, the action $a_i(k)$ is selected according to the probability distribution in p as introduced in subsubsection 2.3.4.3.

Step 2) if and only if receiving $\beta(k)$ with a reward from an environment, instead of updating p_i directly, the update takes place as in Equation 2.8 and Equation 2.9, and according to $d_i(k)$ and $d_j(k)$, $\forall d_j : j \neq i$. Where $d_i(k)$ is the estimated reward probability from choosing action i until step k. In cases when $\beta(k)$ is a penalty, the action probability vector can be updated as introduced in Equation 2.7.

$$\begin{aligned} p_i(k+1) &= p_i(k) + \gamma \sum_{j \neq i} [d_i(k) - d_j(k)] \\ &\cdot \left[G_{ij}(k) p_j(k) + G_{ij}(k) \left(\frac{p_i(k)}{r-1} \right) (1 - p_j(k)) \right], \\ p_j(k+1) &= p_j(k) - \gamma [d_i(k) - d_j(k)] \\ &\cdot \left[G_{ij}(k) p_j(k) + G_{ij}(k) \left(\frac{p_i(k)}{r-1} \right) (1 - p_j(k)) \right], \forall j : j \neq i. \end{aligned} \tag{2.8}$$

$$\begin{aligned} G_{ij}(k) &= 1 \text{ if } d_i(k) > d_j(k), \text{ and,} \\ &= 0 \text{ if } d_i(k) \leq d_j(k). \end{aligned} \tag{2.9}$$

Step 3) in all cases, the action long-term reward probability estimates vector $d(k+1)$, where the latter is cumulatively updated from previous interaction steps [146]. Equation 2.10 demonstrates an example of how to estimate whether action a_i

is an optimal action, in parallel with updating its selection probability from Equation 2.7 and Equation 2.8. Hence, the action selection probability update is pursued:

$$\begin{aligned} W_i(k+1) &= W_i(k) + (1 - \beta(k)), \\ Z_i(k+1) &= Z_i(k) + 1, \\ d_i(k+1) &= \frac{W_i(k+1)}{Z_i(k+1)}, \end{aligned} \tag{2.10}$$

where $W_i(k)$ represents the number of times the action a_i was rewarded up to step k, and $Z_i(k)$ is the number of times the same action was selected until step k. $\beta(k) = 0$ if the environment feedback is a reward. Hence, d_i at interaction step $k+1$, is the estimated probability that action a_i would be rewarded as of step $k+1$.

Compared with the linear reward-Inaction scheme L_{R-I}, estimator algorithms can converge almost an order of magnitude quicker than the former [147]. In our research study, we only considered the update of the action selection probability as in the non-estimator LAs. However, we did that cumulatively by considering the actions reward probability estimates as in Equation 2.10 to be the action selection probability update mechanism. The latter can also be interpreted as proposing a dynamic rate for updating the action selection probability, unlike in traditional variable-structured-LAs (see subsubsection 2.3.4.3). We further explain and claim the reasons behind our proposed learning schemes and their empirical results in chapter 3.

2.3.4.5 Random Walk and Knapsack Algorithms

Random walks are stochastic processes that describe a path projection in some space [148]. The possibility of being in a current location during such a walk is stochastic and depends on the previous location. An example is an object that attempts to perform stochastic moves on the real line. In a more complicated scenario, there could be n-dimensional random walks [149] where a joint probability distribution governs the joint moves.

A wide range of problems can be represented as a random walk. For example, in a stochastic knapsack optimization problem [150], the optimal amounts of items that maximize the total knapsack value can be estimated over such a random walk. Hence, the multidimensional random walk tasks can represent a Multidimensional Knapsack Problem (MKP). MKP is known to be NP-hard in operations research and has a wide range of applications in engineering and management [151].

Liu et al. (2016) [151] proposed a binary differential search method to solve MKP where a Brownian motion-like random walk guides the stochastic search. In their proposed framework, the movement of a superorganism was described by a Brownian-like random-walk model seeking the optimal migration movement gain as an optimal knapsack value. In the latter approach, they utilized a meta-heuristic method to help boost the optimal knapsack value. These heuristics can be viewed

as the context $\mathcal{X}$ described in subsection 2.3.2 when the random walk is conducted through a LA.

Figure 2.7 gives an example for a structure of a random walk LA for the knapsack problem where state transitions go either to the left or to the right direction over a discrete-state space. The latter can also be viewed as two potential actions representing two opposite moves over the state space.

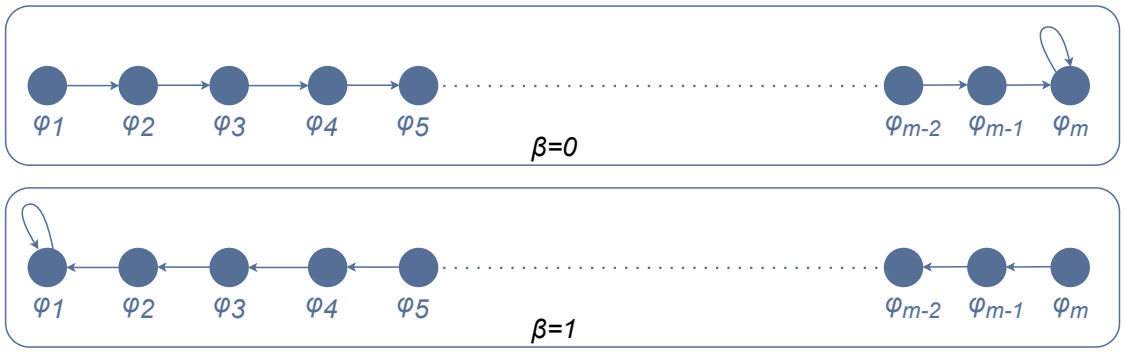


Figure 2.7: An example of a random walk LA for the knapsack problem. The LA has m discrete states and increases its state value by going to the right when rewarded ($\beta = 0$) or decreasing its state by going to the left when penalized ($\beta = 1$)

The state transitions are non-deterministic and based on a stochastic signal from a knapsack. The latter sends an evaluation signal for each conducted state transition to indicate how the state value would contribute to maximizing the knapsack's total value. With an analogy to the truth campaign's incentives (see subsection 2.1.2) investigated in this dissertation, these state values can represent the being-learned incentives during the interaction between the LA and the information diffusion environment (i.e., SM predicted dynamics), where such an interaction can be defined as a knapsack optimization problem.

In chapter 3, we discuss further how our proposed LA-based algorithms performed a multidimensional random walk, where state transitions probabilities were updated similarly to the concept in Equation 2.10.

2.3.5 Relevant Social Network Applications

LAs have been employed in many social network problems. For example, Khomami et al. (2016) [152] proposed an algorithm based on distributed LAs for community detection in complex networks. In their proposed work, they assigned each vertex on the network to an automaton, and each updates its action probability according to coordination between all the LAs. At the same time, the algorithm interactively tried to identify high-density local communities. In such an algorithm, the action probabilities were updated according to the environment feedback, where the latter was based on an evaluation of a loss function. In this manner, the loss function was evaluated in interaction step k when a particular action dictates the projection of the function output. Then, the function value at step $k-1$ was compared to its value at step k — and if the obtained value was decreased from what it was in the

previous step, the automaton was rewarded and the action probability vector was updated accordingly.

Other applications, such as link prediction, have also been studied through the framework of LAs. For example, Moradabadi and Meybodi (2017) [153] proposed a time-series link prediction-based LA. The automaton had two actions to decide whether there should be a link between two adjacent nodes on the network. In this method, the action probabilities were updated according to similarity measurements between the link candidate neighbors. The latter acted as a P-model environment feedback where $\beta = 0$ if similar and $\beta = 1$ if not. In their approach, they utilized a linear reward-penalty algorithm.

Further, Ghavipour and Meybodi (2018) [154] utilized LAs for propagating trust through trust networks by inferring the trust between two indirectly connected users. In their approach, they utilized distributed LAs to capture the dynamicity of trust during the trust propagation process and dynamically update the found reliable trust paths based on a dynamic reward parameter.

Other efforts have also been made to address the problem of influence-based propagation. For instance, Ge et al. (2017) [155] extended the confidence interval estimator-based LA [156] to model the social network environment as S-model and proposed an end-to-end approach for influence maximization. Their approach estimated the maximal rewarded actions to update the action probability vector as in estimator algorithms.

This dissertation investigated similar techniques to some of these previewed efforts on social networks. For instance, our proposed architectures assigned each automaton to each user to learn the optimal incentives for that user to join a truth campaign on SM. Furthermore, we modeled an environment signal through a loss function, where the latter encapsulated the proposed social context representations and dictated the state transition of the proposed LAs.

Chapter 3

Proposed Methodologies

In chapter 1, we introduced the idea of information diffusion modeling to simulate Social Media (SM) users' activities and predict their information dissemination patterns. Moreover, we introduced how Reinforcement Learning (RL) agents can control an information diffusion model through interventions to optimally incentivize users to propagate facts. We also discussed some research gaps in the existing solutions, such as the need for more realistic criterion functions, the limited representation of users' activities in the diffusion model, and the need for transparency and verification of the learned incentives. Further, chapter 2 provided theoretical background on Multivariate Hawkes Processes (MHPs) and Learning Automata (LAs) as the basis for our proposed Multiplex-Controlled Multivariate Hawkes Processes (MCMHP). The MCMHP can provide a mathematical framework for modeling information diffusion with LA, where the LA can be utilized as a diffusion intervention algorithm. In this manner, the LA converged state value can be interpreted as an incentivization value for tasks such as intervention-based misinformation mitigation.

In this chapter, we introduce our proposed methods and research outcome that filled some of the research gaps for the problem of intervention-based misinformation mitigation on SM. In section 3.1, we demonstrate the major contribution and finding of our study by explaining our novel approach: "societal acceptance-aware truth campaign", which is based on our novel users' activities representation encapsulated in the MCMHP framework. Therefore, we demonstrate the details of the MCMHP framework in section 3.2 and provide a general mathematical framework to control interdependent groups of MHP-based diffusion models via decentralized LAs. In the latter, the LAs are rewarded based on a proposed criterion function that traces interdependent temporal diffusion dynamics. Eventually, section 3.3 shows empirical results after evaluating the LA-based misinformation mitigation algorithms on multiple datasets. Moreover, section 3.3 provides some preliminary results of other helper methods, which we believe can improve the trustworthiness and analysis of the MCMHP framework in the future.

3.1 Societal Acceptance-aware Truth Campaign

To characterize the propagation of misinformation caused by political manipulation on SM, we utilized MHPs as our time-series information diffusion modeling approach (see subsection 2.2.3). We believed fitting a good diffusion model for online misinformation needed to interconnect the interdependent temporal features that characterize the concept of societal acceptance. The latter concept is a fundamental factor in understanding information dissemination patterns on SM [44]. Therefore, this dissertation proposed novel temporal features representation to characterize users' expressed beliefs toward others' opinions. The proposed features extended the traditional existing feature space [20, 25, 30] for the problem of intervention-based misinformation mitigation on SM, where the feature space was mostly limited to misinformation and factual information activity counts.

3.1.1 Online Users' Activity Representation

According to Olan et al. (2022) [44], societal acceptance is considered a game-changer for controlling the spread of misinformation on today's SM platforms. The reason is how societal and political beliefs control how users react to online content [12]. On SM, contents live and circulate over societal bubbles [10], where a societal bubble is a circle of biased users embracing a particular opinion or idea. Therefore, we believe modeling the temporal formulation of these circles' engagements and the nature of such engagements can reveal the patterns of societal acceptance. Some examples of these engagements can be in the forms of retweeting, commenting, and liking. The nature of the latter engagements can be agreeing, disagreeing, or being neutral to what ideas these circles are talking about.

Paper G (see Appendix G) which concluded this dissertation — proposed that a societal acceptance incident can be interpreted as a societal circle that accepts to be changed, weakened, and hence transformed into another unharmful circle. An example of the latter is when the number of users agreeing with a harmful circle's political or societal belief[1] becomes very small. In this manner, we proposed the users' activity representation through three MHP diffusion groups to simulate and predict three online activities: information veracity-related activities, societal bias levels-related activities, and societal engagements-related activities. We calculated the counts generated from the MHP diffusion group associated with societal circles' engagement to measure the change in a societal circle after applying incentivization through the LA-based misinformation mitigation algorithm. Similarly, we calculated the counts from the societal bias diffusion to measure the nature of engagements. The latter estimated how likely a user would accept or be influenced by an incentivization to propagate or deny the factual information. Overall, because the

[1]The judgment of a political or a societal belief is relative to the point of view from which we look at it. Therefore, in our research case studies, we considered a political or a societal belief harmful according to human rights reports from organizations such as Amnesty or our common sense when it is obvious. An example of the latter is when Donald Trump propagated false information about Covid-19 during the early phase of the pandemic [157].

dynamics of societal engagement and bias are strongly dictating the generated misinformation or factual information, we needed to measure the counts of generated misinformation and factual information activities. We demonstrate our proposed interdependent diffusion groups below:

- **Information Veracity**: That had two MHPs, one MHP to predict counts of misinformation activities, and another MHP to predict counts of factual information activities.

- **Societal Bias**: That had three MHPs, each was concerned with the nature of users' engagements toward a particular topic, where one MHP predicted counts of agreeing activities, and another MHP predicted counts of disagreeing activities. Finally, the third MHP predicted the counts of being neutral-related activities.

- **Societal Engagement**: That had the number of different societal or political beliefs engagements (i.e., societal circles) which discussed the particular topic. For example, a MHP predicted the counts of agreements on a topic while expressing misinformation around it. Another societal engagement MHP predicted the counts of agreements while expressing factual information. Another MHP predicted the counts of being neutral while expressing factual information. Hence, by having three possible engagements natures (i.e., societal bias), and two possible veracity levels, we ended up with six societal circles MHPs.

To obtain the above representation, we had to collect and annotate a novel dataset with the three temporal labels that illustrate the three diffusion categories. The temporal information was based on a user level, e.g., each user's tweet or retweet was indexed by its timestamp and was given three labels: its veracity, bias, and associated societal circle. Eventually, the data has branched to different temporal versions to train the different MHPs models to predict the associated activities.

Analogically to current approaches of misinformation mitigation [20, 25, 30], Figure 3.1 shows the proposed design of diffusion and control models interaction, compared to the typical existing solutions as shown in Figure 3.2. The truth campaign optimization problem introduced in subsection 2.1.2 was solved in **Paper G** according to the proposed mechanism in Figure 3.1. In this manner, we proposed interaction between a network of LAs and multiplex MHPs. The LAs network was built on top of the social network users, meaning for each network user, there was an associated LA to optimize the truth campaign incentive needed for that user. However, unlike how the optimization task was conducted in the existing methods [20, 25, 30], our proposed intervention model, the LA, was rewarded and guided through its state transitions based on a novel criterion from diverse and complex interconnected MHPs rather than just one MHP. That constructed the proposed MCMHP from our research study, where we proposed a shared LA state as a shared incentivization among the three interdependent diffusion groups as in Figure 3.1. We provide the complete details of the MCMHP framework in section 3.2.

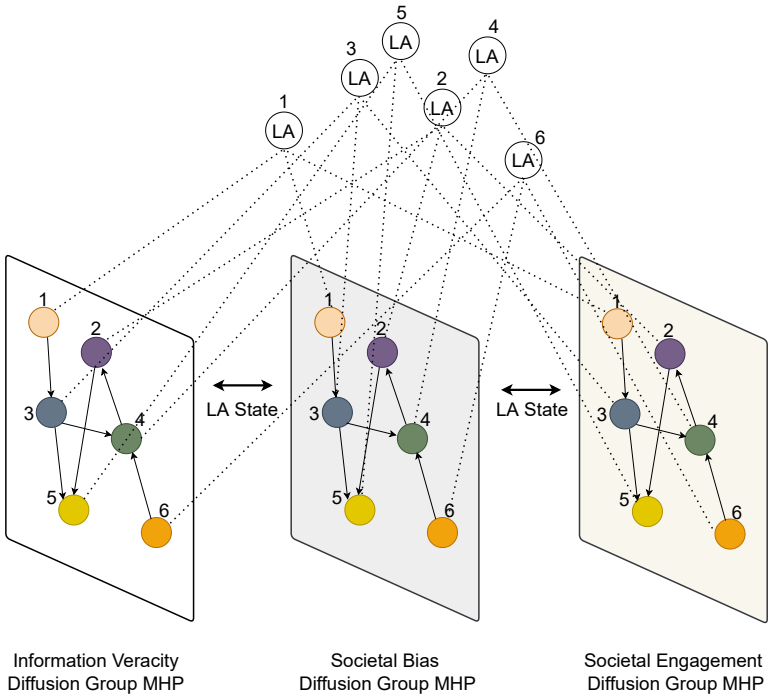


Figure 3.1: A toy example of a social network with 6 users and the proposed design of MCMHP interaction

3.1.2 Applying Truth Campaign Interventions

The shared automaton's state means that the being-learned incentive for propagating a particular activity inside a diffusion group (e.g., factual information) will be the same as for other incentivized interconnected diffusion groups (e.g., particular bias and engagements). This incentive value was shared because when incentivizing a user to circulate factual information or a particular type of information, we also incentivized the particular bias direction that information had. Similarly, we incentivized the user to be a member of the societal circle where such information and bias direction was associated with the circle's political or societal beliefs. We demonstrate below, the mechanism of applying the truth campaign incentives.

- An associated LA for each user can conduct its state transition and learning of the optimal incentive through societal acceptance-aware criterion feedback. The feedback is received after interaction with multiplex MHPs-based diffusion. The latter facilitates the prediction of users' activities over the simulated social network as demonstrated earlier in subsection 2.2.3.

- The societal acceptance awareness is achieved because rewarded incentives will be determined based on how an incentive (LA state) is applicable with regard to the estimated user's probability of accepting outsider ideas (i.e., truth campaign content). We provide the complete details about the proposed societal acceptance-aware criterion in subsubsection 3.3.1.3.

- Suppose our task is to intervene to incentivize and predict the counts of three interdependent MHPs activities: x, y, z, and only predict the counts of the

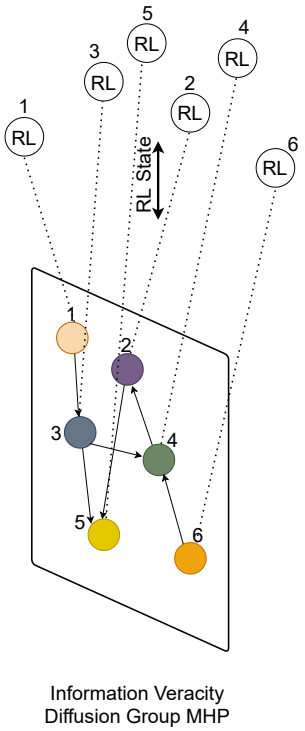


Figure 3.2: A toy example of a social network with 6 users and the typical design of MHP interaction with a control model

MHP activity w without any intervention. Then, for each user i inside the set of users U, a LA_i conducts a state transition to a particular incentive value, and the incentive is passed to the activity x of factual information MHP to boost its counts in the future. The latter is performed through Equation 3.1, where ϕ_i is a current state being-learned for user i. Further, the same incentive ϕ_i is passed to a societal bias activity y (see Equation 3.2) and a targetted societal circle activity z (see Equation 3.3), while the activity w is only predicted without any intervention (see Equation 3.4).

- Activities such as the activity w will only be predicted without any incentivization because they represent independent or undesired-to-be-incentivized diffusion with regard to the truth campaign. For instance, an activity w can be the activity of misinformation events that will not be affected by the incentivization of factual information since we only incentivize the latter and its relevant activities such as particular bias and engagements.

- Suppose the factual information diffusion incentivization resulted in increased counts of factual information but did not change the societal bias or the societal circle activity counts in terms of accepting the truth campaign beliefs. Hence, the associated automaton will not be rewarded for being in the current state since the user bias and societal circle membership remain against the truth campaign.

$$\forall i \in U : \lambda_i^x(\mathbf{T}|H_x^{\mathbf{T}}) = \phi_i + \mu_i^x(\mathbf{T}) + g_i^x(\mathbf{T}, A^x), \tag{3.1}$$

$$\forall i \in U : \lambda_i^y(\mathbf{T}|H_y^{\mathbf{T}}) = \phi_i + \mu_i^y(\mathbf{T}) + g_i^y(\mathbf{T}, A^y), \tag{3.2}$$

$$\forall i \in U : \lambda_i^z(\mathbf{T}|H_z^{\mathbf{T}}) = \phi_i + \mu_i^z(\mathbf{T}) + g_i^z(\mathbf{T}, A^z), \tag{3.3}$$

$$\forall i \in U : \lambda_i^w(\mathbf{T}|H_w^{\mathbf{T}}) := \mu_i^w(\mathbf{T}) + g_i^w(\mathbf{T}, A^w). \tag{3.4}$$

The above illustrates how the novel activity representation informed the LA intervention algorithm with the interdependencies between the different dissemination patterns. Thus, the intervention did not rely on the naive assumption that increased counts of factual information mean that the intervention succeeded. In this manner, a change in the societal bias or the societal circle activity counts will result in a change in the estimated probabilities of being in a particular bias or being a member of a particular circle, respectively. As a result of the latter, the optimization procedure can be traced at any point to see how the current incentives succeeded in reshaping the societal bias and engagements on the network and that empowered the transparency and analytical capacity of the solution. For complete details about the proposed activity representation and the fitting of the multiplex diffusion groups, see Appendix G.

3.2 The MCMHP Framework

By proposing the MCMHP framework, we contributed with the abstraction and simplification to model relevant scenarios where LAs are required to optimize a volume-based diffusion model to boost the counts of a particular activity on a social network. In this section, we define the problem statement and assumptions to set boundaries for when we believe the MCMHP framework can be relevant. Further, we provide the mathematical notion and topology of the MCMHP. Eventually, we demonstrate the learning model inside the MCMHP.

3.2.1 Problem Statement and Assumptions

Given a simulated social network where multiple interdependent activities circulate, we want to boost the circulation for some of these activities to increase exposure to them. Hence, the challenge is to learn to incentivize network elements to elevate particular activities. Below, we demonstrate the assumptions for when we believe the MCMHP framework can be applied:

- A social network element, such as a user, can be influenced to elevate some activity [41] by an unknown outsider influence and an unknown internal network influence from other elements, such as other network users.
- The unknown internal network influence is temporal (i.e., based on temporal patterns) and variable to each element and not exclusive to only directly connected elements, but a diffusion law [41] of temporal influence exists between

the elements. For instance, those exposed to the activity of others because they were active on the network at the same time of these activities.

- The unknown outsider influence is variable to each network element because of the different external physical world conditions that govern these elements [158].
- The physical world conditions and activity dynamics intensity can be characterized by a diffusion model where both outsider and internal influence can be quantified [41].
- Social network activities can be categorized into different diffusion model groups based on an information category. For example, all activities that relate to expressing misinformation or factual information can be categorized inside an information veracity diffusion group. Other activities can be considered part of another predefined diffusion groups. See Figure 3.1 as an example of three predefined diffusion groups.
- In general, influences occur when some interaction happens between two elements on the network or outside it, whether in an agreement or disagreement manner since the influence process does not necessarily indicate the immediate acceptance of an induced behavior [159].
- Internal influence is decayed over time within a particular time window [41].
- In addition to the influence over network element level, another level of influence exists over the diffusion groups, that is how some activities inside a diffusion group can trigger other activities in another diffusion group. For instance, the activity of propagating misinformation content based on a political opinion will trigger the activity of boosting the political bias of that opinion.
- A network element i can internally influence its adjacent element j (e.g., following relationship on Twitter between i and j) through the intensity of an activity generated by i. Hence, j becomes influenced by i through exposure to that activity.

3.2.2 Notations and Topology

The MCMHP framework can be expressed with the sextuple $\langle U, G, \mathcal{Z}, \mathcal{Z}^H, L, F \rangle$. Let U be a social network where $U = \{u_1, u_2, u_3, ..., u_n\}$ is a finite set of n network elements. Then, let $G = \{G_1, G_2, G_3, ..., G_l\}$ be the set of l predefined MHPs diffusion groups, where each represents a group of MHPs to predict the diffusion of different network element activities that relate to the same information category. For example, all activities that relate to expressing misinformation or factual information can be categorized inside an information veracity diffusion group.

Further, let $\mathcal{Z} = \{Z_1, Z_2, Z_3, ..., Z_n\}$ be a set of n sets, where each underlying set is associated with a network element u_i, where $Z_i = \{\lambda_1, \lambda_2, \lambda_3, ..., \lambda_k\}$ is the finite

set of the k predicted counts of network activities, i.e., the predicted counts of u_i activities from the MHP diffusion groups. $\mathcal{Z}^H = \{Z_1^H, Z_2^H, Z_3^H, ..., Z_n^H\}$ also contains underlying sets associated with network elements, but $Z_i^H = \{H_1, H_2, H_3, ..., H_k\}$ is the finite set of the k historical counts (i.e., before prediction), each associated with a network activity.

To predict the count λ^x of a particular activity x inside a MHP diffusion group, the unknown vector $\mu^x \in \mathbb{R}^n$ that represents the outsider influence on n network elements must be estimated first from the historical data of activity x. Similarly, the internal influence matrix $A^x \in \mathbb{R}^{n \times n}$ is also to be estimated to measure the internal influence between any two network elements. For example, the matrix entry a_{ij} quantifies how much the element i influences the element j and that guides a diffusion law during the propagation of activity x. See subsection 2.2.3 for complete details about how activity counts are predicted and evaluated for MHPs.

Further, let $L = \{LA_1, LA_2, LA_3, ..., LA_n\}$ be the set of n LAs associated with n network elements, where LA_i observes u_i activities over the MHP diffusion groups. Therefore, L establishes a network of LAs that learns the unknown set $\Phi^* = \{\phi_1^*, \phi_2^*, \phi_3^*, ..., \phi_n^*\}$, where ϕ_i^* is the desired optimal LA_i discrete state value for the element u_i. The latter is interpreted as an optimal incentive value to boost the circulation of some activities for element u_i. In this manner, an optimization loss function $F(\mathcal{Z}, \mathcal{Z}^H, \Phi)$ is defined to send a multiplex MHP-based environment feedback to each LA_i based on evaluating its current state value with regard to the predicted diffusion counts after assigning the state value (i.e., incentive), seeking the LAs network to eventually converge to an equilibrium [160]. $F(\mathcal{Z}, \mathcal{Z}^H, \Phi)$ is optimized over the current state values set Φ of all LAs and both the predicted and historical counts of all predefined network activities. The use of $\mathcal{Z}^H$ is important in the MCMHP to evaluate how the intervention (i.e., incentivization) with the MHP-based environment improved the diffusion with regard to the initially calculated diffusion from the historical counts. Moreover, part of the counts in $\mathcal{Z}^H$ is used as a test dataset to evaluate the MHPs prediction.

Figure 3.3 demonstrates the topology of the MCMHP entities discussed above, with an example of three interdependent activities counts to be incentivized and predicted: λ^x, λ^y, λ^z and one activity λ^w to be only predicted. Three predefined MHP diffusion groups: G_1, G_2, G_3 are illustrated in the example.

3.2.3 Learning the Optimal Incentives

The proposed LAs network architecture to learn the optimal activity incentives within the MCMHP is demonstrated in Figure 3.4. The proposed architecture belongs to a variable-structure LA (see subsubsection 2.3.4.3) where a state transition probability can be different at each step the LA interacts with its environment. However, unlike traditional variable-structure LAs, we proposed a rate of increasing and decreasing state transition probabilities that varied with time. Further, the LAs interacted with a P-model MHP-based environment where the feedback signal $\beta = 0$ if the LA was rewarded, and $\beta = 1$ in case of a penalty. Our novel learning

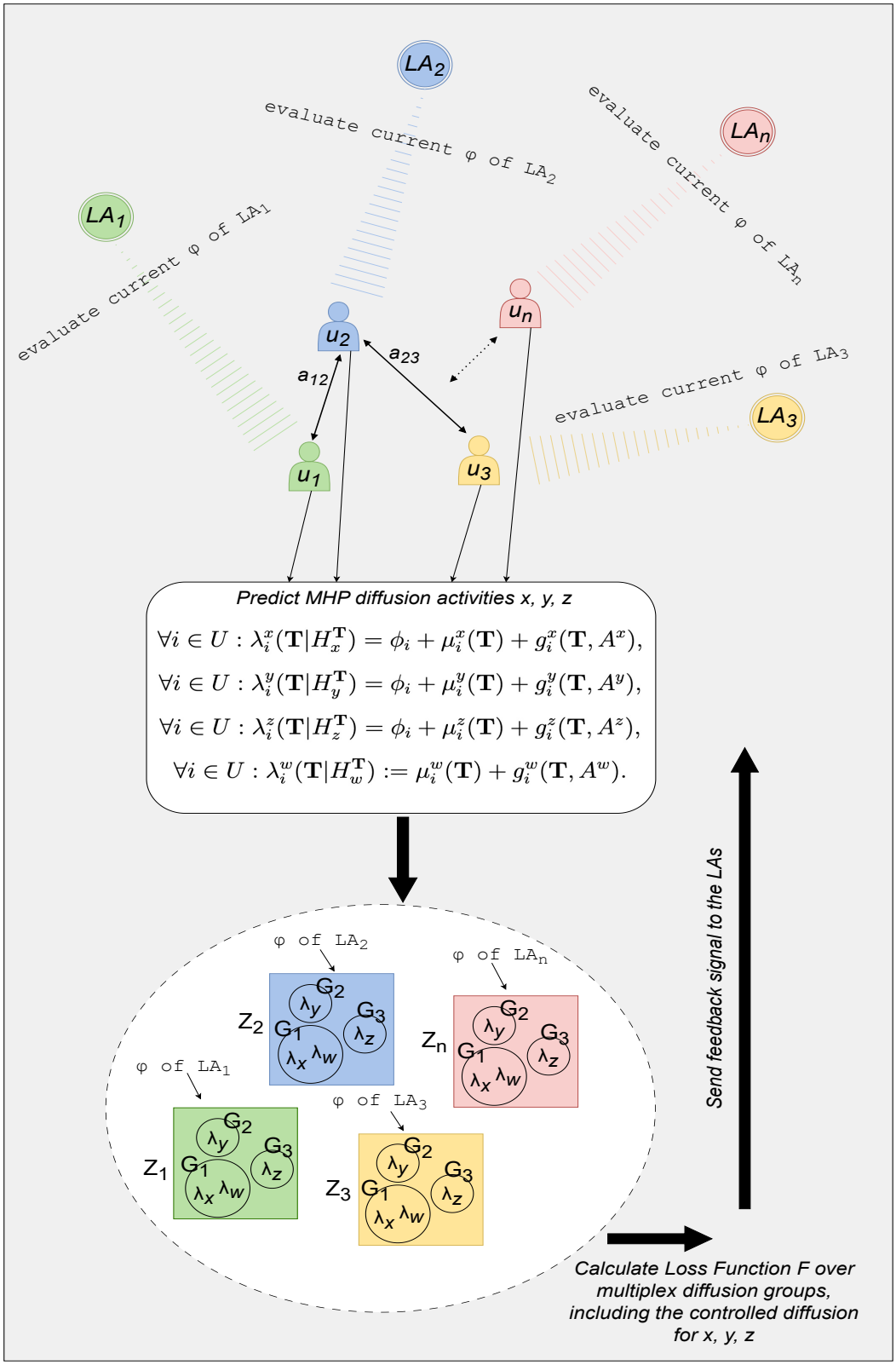


Figure 3.3: The topology of the MCMHP framework with an example of three interdependent activities counts to be boosted: λ^x, λ^y, λ^z and one activity λ^w to be only predicted. Three predefined MHP diffusion groups: G_1, G_2, G_3 are illustrated in the example

mechanism was as follows:

- An individual LA_i has a discrete state space with a memory depth m, where $m > 0$. The best value for m can be obtained from a grid search and according to how responsive the increase or decrease amounts in the parameters of the selected diffusion model are. That is because a converged state value of LA_i will be utilized to optimize the diffusion model parameter for the network element u_i, i.e., the external incentive that influences u_i.

- All LAs are initialized to start in state ϕ_1, which indicates an incentive value of 0. When a LA_i is in a state $\phi_j : m > j > 1$, then, it has three possible state transitions: $\phi_{j-1}, \phi_j, \phi_{j+1}$ indicating going to left to decrease the state value (a_2), staying at the same state, and moving to the right to increase the state value (a_1), respectively.

- Each LA_i could only have the two state transitions: ϕ_1, ϕ_2 or ϕ_{m-1}, ϕ_m if and only if its current state is ϕ_1 or ϕ_m, respectively.

- Only one LA_i can be active at a time and conducts its state transition. Then it awaits for the remaining LAs to complete their state transitions in the same manner, and after that, if not converged, LA_i becomes active again and performs another transition.

- For each LA_i, to explore different state values and eventually reach an optimal or sub-optimal state ϕ_j^*, each state in LA_i has a state transition probability distribution vector to guide the moves directions from that state.

- When being in a current state, the LA_i needs to learn a decision whether that state should be optimal or sub-optimal. Therefore, the LA_i needs to learn the probability vector of the possible transitions of that current state. The LA_i does the same for all states' probability vectors until the latter converges to a steady value. For instance, when being in an optimal state, the latter will have a probability of transition to itself closer to 1. If that probability does not change over multiple upcoming interactions with the environment, the LA_i is considered to be converged.

- Inside each automaton with m states, the above mechanism can be formally defined as overlapping m Markov chains [161] inside each LA_i, where each state has its associated Markov chain that dictates the LA_i transition, where that transition depends on the learned probabilities in the current Markov chain. That ensures the LA_i will have at least one Markov chain with a state probability of transition to itself closer to 1 and then the LA_i conducts no further transitions when it is in that state. Figure 3.5 illustrates these Markov chains and the update of the state transition probability vector for each state until convergence.

Equation 3.5 demonstrates the state transition function δ for each LA_i. The transition works as follows. At each interaction step k, the probability of the LA_i

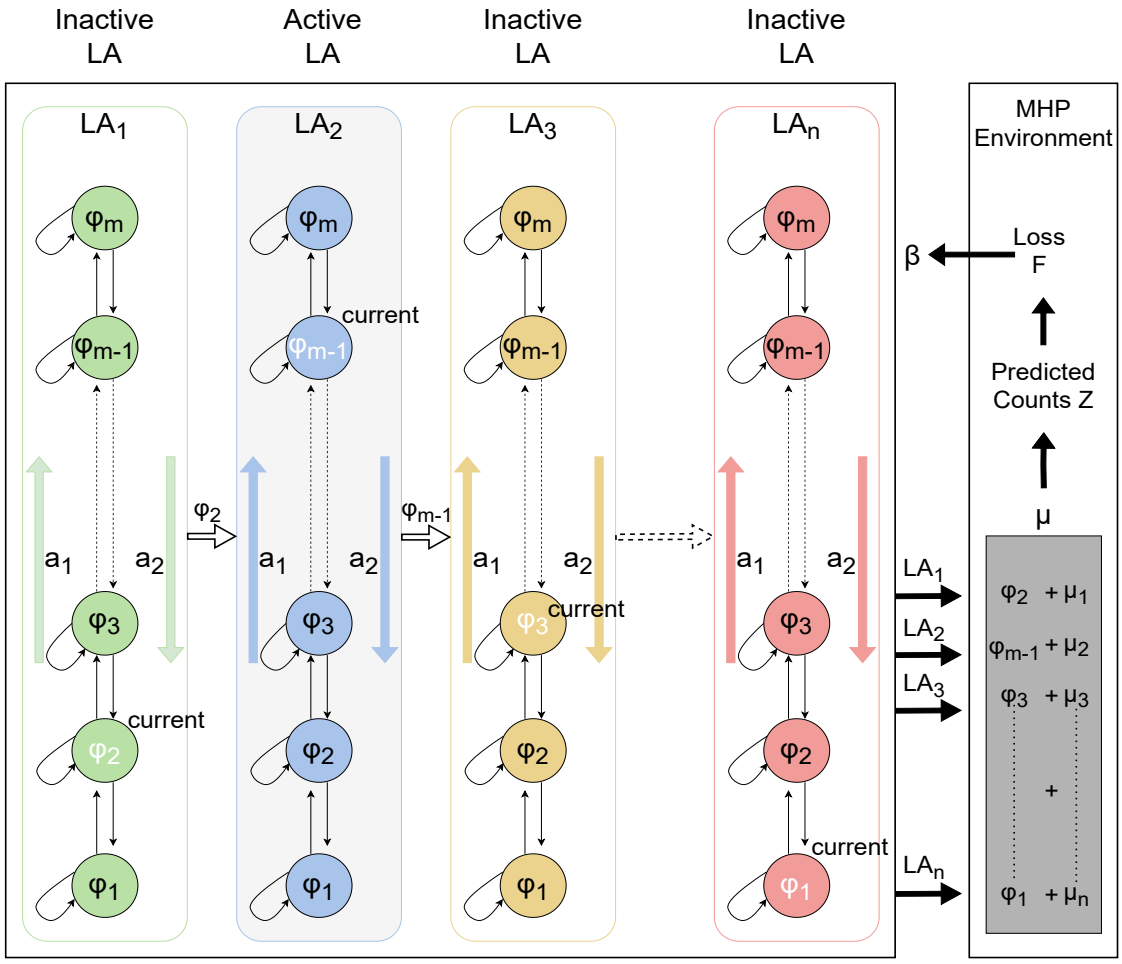


Figure 3.4: The proposed LAs network architecture, where one (active) LA conducts a state transition at a time. All LAs current states are passed to MHP-based environment to predict diffusion after modifying the base intensity μ and calculate a reward signal based on evaluating a loss function slop

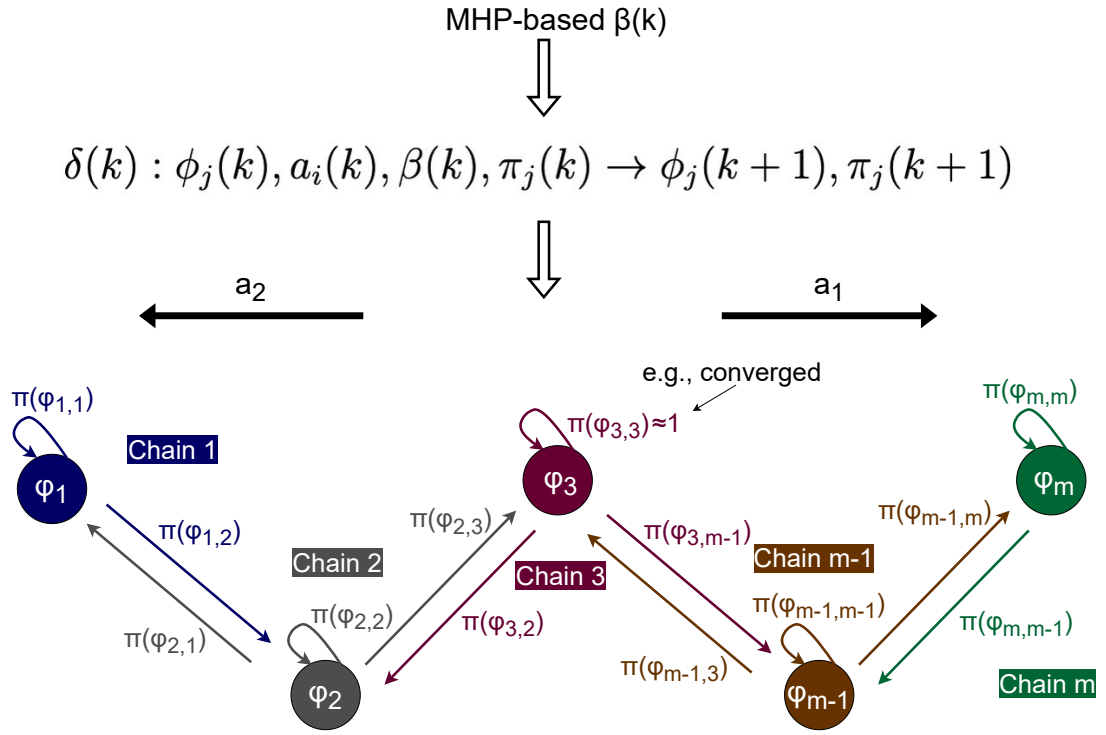


Figure 3.5: The Markov chains associated with each state ϕ_j for an automaton, where state transition probabilities for each chain are updated when being in a state that relates to that chain and receives the MHP-based environment signal β

being in the next state $\phi_j(k+1)$ depends on its state from the current step $\phi_j(k)$, the transition direction conducted at the current step $a_i(k)$, the MHP environment signal at the current step $\beta(k)$, and state transition probability vector $\pi_j(k)$ for the Markov chain associated with state ϕ_j and calculated until step k. Hence, the transition function δ determines the next state $\phi_j(k+1)$ and updates the transition probability vector $\pi_j(k+1)$ for the associated Markov chain with the state ϕ_j.

$$\delta(k) : \phi_j(k), a_i(k), \beta(k), \pi_j(k) \rightarrow \phi_j(k+1), \pi_j(k+1). \tag{3.5}$$

During the interaction with the MHP environment at a step k, all LAs conduct state transition according to Equation 3.5. However, at each interaction step k, the LAs wait except only one LA which can be active at a time. Then, the active LA_i conducts a move over its discrete state space based on a greedy selection from the state transition probabilities associated with the Markov chain of its current state. If the conducted move was rewarded with $\beta(k) = 0$, the LA_i commits such a move. Otherwise ($\beta(k) = 1$), the LA_i stays in the same recent state as before the move. In all cases, the LA_i updates its state transition probability distribution vector for the state it was in before the move, seeking at least one of the Markov chains to have a state probability of transition to itself closer to 1, so the LA_i converges and determines that such state is the suggested incentive value for the network element u_i.

The LAs probabilistic moves can be seen as a joint random walk over their

discrete state space, where the signal $\beta(k)$ is determined by the active automaton's tried state transition and also by evaluating all current states of the remaining LAs in the network. Equation 3.6 and Equation 3.7 explain how the criterion $\beta(k)$ was calculated for each possible set of LAs state values. Where $\mathbf{m}^k, \mathbf{m}^{k-1}$ were the slopes of a loss function $F(\mathcal{Z}, \mathcal{Z}^H, \Phi)$ at the interaction steps $k, k-1$, respectively. We provide more details about the optimization loss functions in subsection 3.3.1. The optimization task was defined as a constraint knapsack optimization [162] where the knapsack item amounts were represented by the state values. Therefore, $\mathbf{K}$ was a knapsack signal indicating whether the knapsack was full ($\mathbf{K} = 1$) or not yet ($\mathbf{K} = 0$).

$$\beta^k(\mathbf{m}^k, \mathbf{m}^{k-1}) := \left\{ \begin{array}{ll} 1, & \text{if } \mathbf{m}^k > \mathbf{m}^{k-1} \vee \mathbf{K} = 1 \\ 0, & \text{otherwise} \end{array} \right\}, \text{ for } \phi_{j,j''}, \tag{3.6}$$

$$\beta^k(\mathbf{m}^k, \mathbf{m}^{k-1}) := \left\{ \begin{array}{ll} 1, & \text{if } \mathbf{m}^k > \mathbf{m}^{k-1} \\ 0, & \text{otherwise} \end{array} \right\}, \text{ for } \phi_{j,j'} \text{ and } \phi_{j,j}. \tag{3.7}$$

For each LA_i, the Markov chain state transitions probability distribution vector for a current state ϕ_j was updated based on the stochastic signal from the MHP environment and calculated as Equation 3.8 and Equation 3.9.

$$\pi^{k+1}(\phi_{j,j''}) = \frac{v^k(\phi_{j,j''})}{w^k(\phi_{j,j''})}, \tag{3.8}$$

$$\begin{aligned} \pi^{k+1}(\phi_{j,j'}) &= \frac{\pi^{k+1}(\phi_{j,j''})}{2}, \\ \pi^{k+1}(\phi_{j,j}) &= \frac{\pi^{k+1}(\phi_{j,j''})}{2}, \end{aligned} \tag{3.9}$$

where $\pi(\phi_{j,j''})$ is the state transition probability of moving to the right direction (i.e., increase incentive) from a current state ϕ_j to its adjacent state $\phi_{j''}$ according to the associated Markov chain transition probability distribution vector π_j. In this manner, the value order of the three adjacent states in the chain is as follows: $\phi_{j'} < \phi_j < \phi_{j''}$. The variables v^k and w^k represented how many times a transition (i.e., the direction of a move from that current state) was rewarded and selected for an automaton up to interaction step k, respectively. Hence, that can be seen as directly pursuing the optimal transition required for the task through a dynamic rate of changing the transition probability. Finally, Equation 3.9 was applied to update other possible transition probabilities in the associated Markov chain probability distribution vector, where the elements in the latter must add to 1.

The **Papers B** (Appendix B), **D** (Appendix D), **E** (Appendix E), and **G** (Appendix G) provide the complete details of the proposed LA learning algorithms and how we extended their capabilities during our research study in terms of the MHP-based environment criterion β and practicality (check **Paper E** for practicality). In section 3.3, we show empirical results on both real-world and synthetic datasets to indicate how these algorithms performed in different scenarios.

3.3 Evaluation

In this section, we demonstrate some of the empirical and preliminary results from our published research outcome for the task of intervention-based misinformation mitigation on SM. subsection 3.3.1 presents empirical results on multiple misinformation datasets, showcasing the evolution of our state transition criterion in the LAs, as demonstrated through the optimization loss functions we have refined over time. That also illustrates how online users' activities were represented in a loss function domain and how such representation evolved over the time frame of our research study and resulted in the proposed MCMHP framework. Further, subsection 3.3.1 shows how our proposed LAs compete with previously proposed Policy Iteration algorithms on the task of intervention-based misinformation mitigation on SM. Eventually, in subsection 3.3.2, we show other preliminary models and results for some helper models that can be integrated with the major methods described in section 3.1 and section 3.2.

3.3.1 Empirical Results

3.3.1.1 Average Difference Loss Function

According to **Paper B** (see Appendix B), Figure 3.6 and Figure 3.7 show the optimization performance of three proposed variable structure LA schemes that learned how to allocate truth campaign incentives for network users: reward-penalty (RP), reward-Inaction (RI), and penalty-Inaction (PI). See subsubsection 2.3.4.3 for more details about the variable structure LA schemes. The optimization was conducted according to concepts similar to what we demonstrated in subsection 3.2.3, except that only one MHP diffusion group was proposed as the representation of users' activities. The latter was embedded in a mitigation objective function (i.e., loss function). Moreover, the limitation of employing only one diffusion group was due to the available representations in the available datasets, where only activities for misinformation and factual information existed.

We selected three considerable performance baselines for the experiments. The three baselines represented three different measures we sought to outperform, these were the state of misinformation before mitigation, misinformation after uniform, and random distribution of the incentivization budget. The state of misinformation was calculated according to evaluating a proposed loss function as in Equation 3.10 and Equation 3.11, where the latter introduced the bound C, which represented a knapsack capacity where the incentives must stay within, i.e., truth campaign incentivization budget.

$$F(\mathcal{Z}, \mathcal{Z}^H, \Phi) = \left[\frac{\sum_{i=1}^{n} \mathcal{F}(Z_i, Z_i^H)}{n}\right] - \left[\frac{\sum_{i=1}^{n} \mathcal{T}(Z_i, Z_i^H, \phi_i)}{n}\right], \tag{3.10}$$

$$\text{subject to } \sum_{i=1}^{n} \phi_i \leq C, \tag{3.11}$$

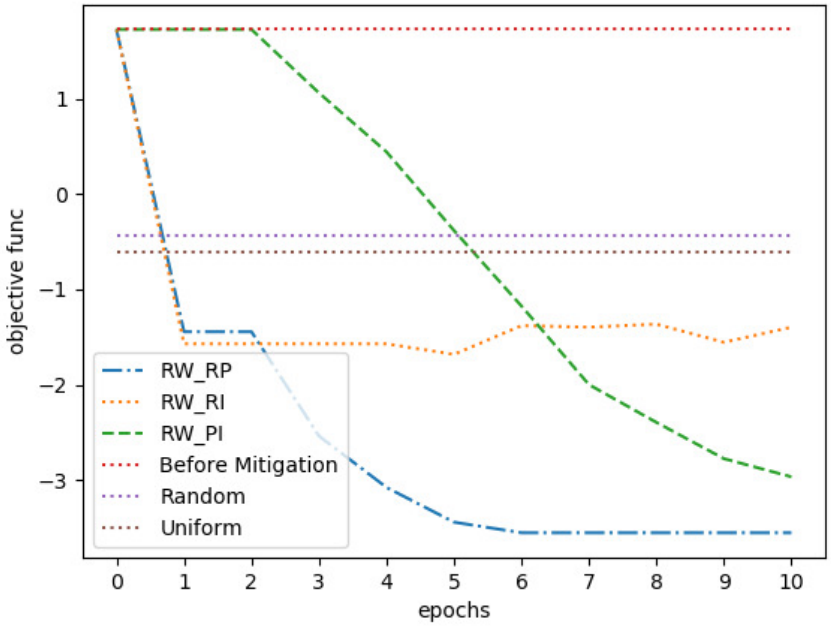


Figure 3.6: Optimization evaluation of different LA schemes on Twitter-15 misinformation dataset. The evaluated learning schemes were reward-penalty (RP), reward-Inaction (RI), and Inaction-penalty (PI). The optimization utilized the evaluation of boosted factual information over all social network users, on average

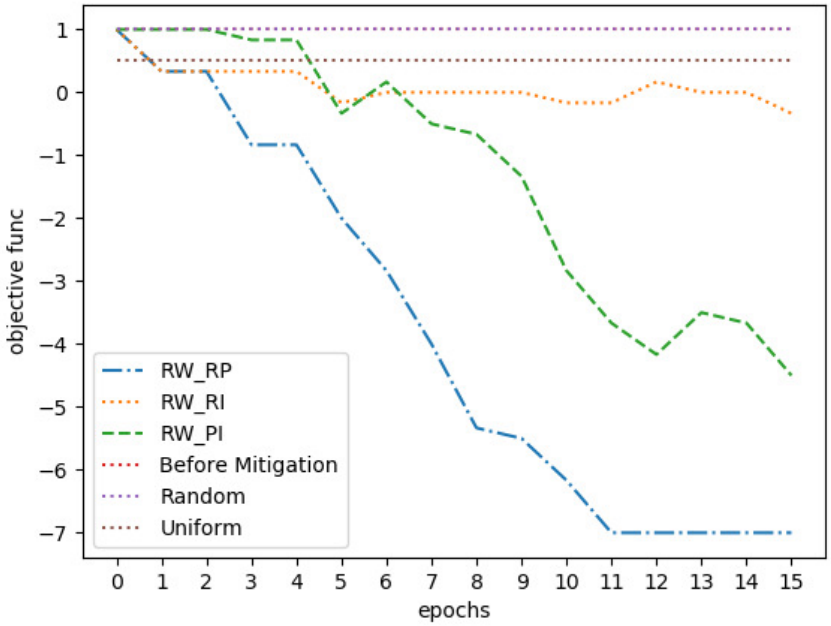


Figure 3.7: Optimization evaluation of different LA schemes on Twitter-Covid19 misinformation dataset. The evaluated learning schemes were reward-penalty (RP), reward-Inaction (RI), and Inaction-penalty (PI). The optimization utilized the evaluation of boosted factual information over all social network users, on average

where $\mathcal{Z}$ and $\mathcal{Z}^H$ had only two activities per each network user, and that represented the exposure counts of propagated misinformation activity $\mathcal{F}$ and factual information activity $\mathcal{T}$, where these counts were calculated from both historical activities $\mathcal{Z}^H$ and predicted activities $\mathcal{Z}$. Exposure counts meant how much of a particular content activity a user i was exposed to. The latter was calculated by summing all counts of that activity among the adjacent users of i, including i as adjacent to itself. In the case of calculating the exposure to factual information activities, the value ϕ_i was considered in the calculation of $\mathcal{T}$, since ϕ_i incentivizes the factual information activity counts in the simulated network.

The calculation in Equation 3.10 was as follows. On a simulated social network with n users, for each user i, $Z_i = \{\lambda_1, \lambda_2\}$ was the set of predicted network activity counts for both misinformation λ_1 and factual information activities λ_2, where the misinformation activity count was calculated from a MHP as illustrated earlier in Equation 3.4, and the factual information activity count was calculated from another MHP as in Equation 3.1. Similarly, $Z_i^H = \{H_1, H_2\}$ defined the set of historical network activity counts of both misinformation H_1 and factual information H_2. Hence, $\mathcal{F}(Z_i, Z_i^H)$ was a counting function that accumulated both historical counts and predicted counts of misinformation activities for each user i, and $\mathcal{T}(Z_i, Z_i^H, \phi_i)$ was a counting function that accumulated both historical counts and predicted counts of factual information after applying the incentivization value from a being-learned LA state value ϕ_i, for each user i. Finally, the loss function was defined as the difference between the generated counts from both misinformation and factual information, until a given time window, i.e., history + predicted time window counts. The latter was calculated over all network users and an average was taken for each activity before calculating the difference. In this manner, as the difference decreased, the criterion for evaluating the effectiveness of misinformation mitigation indicated improvement.

According to Figure 3.6 and Figure 3.7, our LA-based intervention algorithms outperformed the three baselines with the three proposed LA schemes on two real datasets: Twitter-15 political misinformation [32] and Twitter-Covid19 misinformation [163]. On Twitter-15 experiment, we noticed that the PI and RP LA schemes exhibited a nearly identical optimized loss after multiple rounds, referred to as epochs. An epoch represented a single round where all LAs experienced one state transition each. However, the RP scheme was the one with a remarkable early convergence. On Twitter-Covid19, we even observed how the RP scheme outperformed the other methods by far. Hence, it became obvious that the RP scheme was the most reliable learning scheme for our optimization since it converged earlier with better results in all experiments. For the complete details of the work done in **Paper B** and the proposed Average Difference Loss Function, see Appendix B.

3.3.1.2 Fairness Loss Function

In subsubsection 3.3.1.1, we have seen the proposed loss function from **Paper B**, where the definition of the loss was used to evaluate the performance of different

LA schemes. To this end, we concluded that the reward-penalty scheme was the best to learn truth campaign incentives for the task of intervention-based misinformation mitigation on SM. However, the proposed loss in **Paper B** was based on calculating an average difference between misinformation and factual information activity counts, which is not robust for situations when there is a sparse distribution of misinformation among users. In the latter scenario, certain users will require more attention than others when it comes to the incentives they receive. Therefore, **Paper D** (see Appendix D) introduced a new approach that combined the reward-penalty LA scheme with a unique fairness loss function. This function ensured that incentives were allocated fairly, taking into account the specific needs of each user. To evaluate the robustness of the fairness loss function over multiple social networks' scenarios, we introduced a mitigation efficiency metric which was calculated as follows.

$$1 - \frac{a}{b} : b \neq 0, \tag{3.12}$$

where a and b were the overall average misinformation percentages on the network after and before mitigation, respectively. Therefore, a higher value of that metric indicated a better misinformation mitigation performance. According to Figure 3.8, the proposed **Fair-LA** outperformed both the **AVG-LA** introduced in **Paper B** and uniform allocation of incentives in most of the scenarios. The proposed **Fair-LA** had the exact reward-penalty LA learning scheme as in **AVG-LA** but the LAs conducted their state transitions based on a fairness criterion from a novel loss function as in Equation 3.13.

$$F(\mathcal{Z}, \mathcal{Z}^H, \Phi) := \sum_{i=1}^{n} F(Z_i, Z_i^H, \phi_i) \,:\, F(Z_i, Z_i^H, \phi_i) := \sum_{j=1}^{n'} (1 - R_j^{\phi_i})^2, \tag{3.13}$$

$$\text{subject to } \sum_{i=1}^{n} \phi_i \leq C, \tag{3.14}$$

where n represented the number of network users and n' was the number of adjacent users connected to user i, where user i was also considered adjacent to itself. Therefore, j was the index that represented i and all its adjacent users over the summation. $R_j^{\phi_i}$ represented the updated ratio between factual and misinformation exposure counts after applying the recent incentivization value ϕ_i to the factual information MHP diffusion activity associated with user i, since the recent value ϕ_i influences the exposure counts of users adjacent to i. As noticed in Equation 3.13, we squared the subtraction $1 - R_j^{\phi_i}$ to maintain positive values in the interval $[0, \infty)$, while the task was to minimize the loss as much closer to 0 as possible. Thus, the expression $(1 - R_j^{\phi_i})^2$ means that the more the ratio $R_j^{\phi_i}$ approaches a value of 1, the more the individual Fairness Loss Function for user i approaches 0. The latter also means that whenever user i adjacents' exposure to factual information is increased to at least as much as the exposure to misinformation, the assigned incentive for user i becomes fair and hence the remaining incentivization budget will be consumed on

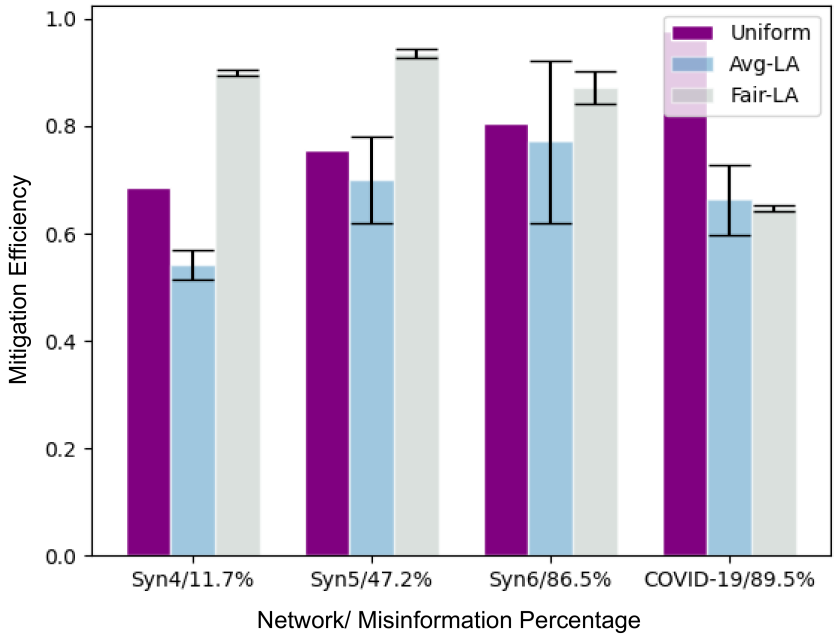


Figure 3.8: Evaluation of the previously successful reward-penalty LA scheme on different social networks with a varied distribution of misinformation. An average difference-based LA loss was compared against fairness-based LA loss

other users. In all brevity, we kept incentivizing user i until all its influenced (i.e., adjacents) users became exposed to enough facts.

It is important to highlight that the total loss was calculated through the achieved individual loss of each user during the allocation of incentives (i.e., the associated LA_i and its current state ϕ_i). That means the total loss in Equation 3.13 ensured optimal or sub-optimal assigned incentivization values over Φ as the set of all LAs current states. Finally, the consumption of all incentivization values (LAs states) must not exceed the bound C, which represented a knapsack capacity during the state transitions and optimization.

According to Figure 3.8, the proposed **Fair-LA** was evaluated in different scenarios where users' exposure to misinformation varied. For example, a synthetic network can have only half or a quarter of its users exposed to misinformation. To provide these scenarios, we created synthetic datasets to replicate different statistics of misinformation on the network. Finally, we observed how the fairness loss resulted in a more robust mitigation efficiency when applied to different scenarios. The only exception was on a version of the Twitter-Covid19 real dataset where 89.5% of network users were almost equally exposed to misinformation. In the latter, it was normal to observe how the uniform allocation of incentives outperformed both loss functions since the more the users are equally exposed to misinformation, the more a uniform allocation will be optimal and outperform. As noticed in Equation 3.13, the **Fair-LA** also conducted its state transitions based on a single MHP diffusion group criterion, where only activities of misinformation and factual information were rep-

resented. For the complete details of the work done in **Paper D** and the proposed Fairness Loss Function, see Appendix D.

3.3.1.3 Societal Acceptance Loss Function

Paper G (see Appendix G) introduced PEGYPT, an innovative dataset that focuses on misinformation in the political domain. This dataset includes temporal labels that indicate whether the information is misinformation or factual, as well as bias directions and dynamics related to societal circle formulation. These temporal labels provide insights into the activities of online users on today's SM platforms. Therefore, it became possible to model three MHP diffusion groups based on the collected dataset (see Figure 3.1). To this end, **Paper G** proposed a unique optimization loss function. This function encompassed the possible values in sets for information veracity, temporal bias, and information from societal circles, which are crucial aspects of the problem domain [44]. The loss function guided the state transitions in the reward-penalty LA scheme, enhancing its effectiveness in combating misinformation.

The work done in **Paper G** extended the Fairness Loss Function introduced in **Paper D**. In the latter, the allocation of incentives was accomplished according to how many incentives each user needs to reduce the exposure of users' adjacents to misinformation. In this manner, we kept maintaining the concept of fair incentivization in **Paper G**. However, in addition to representing only temporal activities of misinformation and factual information, we proposed additional information on the temporal societal circles and temporal bias, to model the occurrence of societal engagement and its nature, respectively (see subsection 3.1.1). That means we predicted the propagation of factual information (e.g., non-propaganda), misinformation (e.g., harmful political propaganda), engagement of users with societal circles, and the bias directions of users, all until a specific time window. Such a combination gave more close-to-reality dynamics of information diffusion and characterized the societal acceptance concept that governs social networks [44]. Equation 3.15, Equation 3.16, and Equation 3.17 demonstrate the Societal Acceptance Loss Function with its maintained fairness.

$$F(\mathcal{Z}, \mathcal{Z}^H, \Phi) := \sum_{i=1}^{n} \left[\Lambda_i'(Z_i, Z_i^H, \phi_i) + F(Z_j, Z_j^H, \phi_i) \right], \tag{3.15}$$

$$F(Z_j, Z_j^H, \phi_i) := \sum_{j=1}^{n'} \left[2 - R_j^{\phi_i} - \Lambda_j(Z_j, Z_j^H) \right]^2, \tag{3.16}$$

$$\text{subject to } \sum_{i=1}^{n} \phi_i \leq C, \tag{3.17}$$

where $\mathcal{Z}$ and $\mathcal{Z}^H$ represented different types of activity counts in multiple MHP diffusion groups, unlike what was proposed in the Average Difference Loss and the Fairness Loss. Further, to keep the fairness concept (see subsubsection 3.3.1.2), R^{ϕ_i}

defined the exposure counts ratios between non-propaganda and harmful propaganda activities for $n^{'}$ adjacent (i.e., influenced) users with user i, including i as adjacent to itself, and until the recent value of state ϕ_i, since the latter influences the exposure counts of users adjacent to i. The function $\Lambda_j(Z_j, Z_j^H)$ was calculated according to Equation 3.18 and represented a joint probability of two events. First, c_j, which was the probability that an adjacent user j with user i would engage with the societal circle that the truth campaign aimed to attract individuals to. This is because a campaign inherently possesses a certain perspective, even if it is a neutral one. Second, the probability that j had the same bias of the truth campaign and denoted as $bias_j$. Through the calculation of this joint probability, we were able to determine the likelihood of an adjacent user j, actively engaging with and accepting the factual information propagated by the incentivized user i. This consideration represented a novel approach to evaluating truth campaign incentives, as it ensured that user i would not consume resources allocated for incentivization if the adjacent user j would ultimately ignore the factual information being shared by i. It is essential to highlight that such probabilities were calculated as per the latest state transition from LA_i, where the latter suggested a recent value for the incentive ϕ_i. That is because an intervention with ϕ_i to the associated activities MHPs (i.e., societal bias and societal engagement diffusion groups), would change the generated counts for bias and societal engagement MHPs activities at each epoch since different state transitions would be performed by LA_i at each epoch. Hence, Λ traced the LAs interventions' consequences on the probability of societal engagement with the circle we sought acceptance of its concept and the agreement with that circle during such an engagement. See section 3.1 for the complete details of intervening with multiple MHP diffusion groups.

$$\Lambda_j(Z_j, Z_j^H) := P(c_j) \cdot P(bias_j). \tag{3.18}$$

While LAs interventions caused different incentives and accordingly different diffusion counts, given an increased value of $\Lambda_j(Z_j, Z_j^H)$ will decrease the individual loss in Equation 3.16, then, the associated LA_i will be rewarded. However, the individual loss for user i could be increased by the function $\Lambda_i^{'}(Z_i, Z_i^H, \phi_i)$, which represented the probability of user i not being in the same bias direction of the truth campaign. That means no matter how the adjacent users of i would engage and agree with the circle we seek, the loss will always be high if user i's bias disagrees with the truth campaign. Equation 3.19 shows how the function $\Lambda_i^{'}(Z_i, Z_i^H, \phi_i)$ was calculated.

$$\Lambda_i^{'}(Z_i, Z_i^H, \phi_i) := 1 - P(bias_i). \tag{3.19}$$

The mechanism in Equation 3.18 and Equation 3.19 means that users will be assigned incentives wisely and according to their probability of accepting the incentives and being influential in their online surroundings. Therefore, this approach aims to optimize the allocation of incentives based on users' acceptance and influence, rather than relying on a simplistic or naive assignment method.

Table 3.1 shows empirical results on the PEGYPT dataset and how the proposed societal acceptance representation in **Paper G** outperformed the traditional representation used in the fairness-only loss from **Paper D**. We can also observe that the percentage of harmful propaganda exposure mitigation was significantly higher than the percentages in both polarization mitigation and societal acceptance boost. We believe that was due to the traditional less strict definition of propaganda content exposure and its mitigation metric [20]. The latter usually considers the counts of activities a user is assumed to access through an adjacency relationship on the network such as "following" [20]. However, we believe that would be a naive assumption since "following relationships" do not guarantee actual exposure in the future. Therefore, it was essential to adopt more strict metrics from our proposed representation, such as the actual dynamics of societal acceptance and polarization, which estimated how likely an engagement would occur and to what degree it would be an agreeing engagement. Moreover, despite how we were still able to calculate the polarization mitigation and societal acceptance boost performances from the Fairness Loss, that would not be feasible without the multiple MHP diffusion groups as a novel representation of the task. Therefore, our proposed novel representation in **Paper G** improved the justification and transparency in intervention-based misinformation mitigation methods, despite the utilized learning algorithm.

Table 3.1: LA obtained performance on PEGYPT dataset when utilizing different optimization loss functions

Metric	**Fairness**	**Societal Acceptance + Fairness**
Propaganda Mitigation	**0.89** ± 0.05	0.88 ± 0.05
Polarization Mitigation	0.23 ± 0.10	**0.26** ± 0.09
Societal Acceptance Boost	0.16 ± 0.03	**0.19** ± 0.05

Figure 3.9 demonstrates the difference in behavior between the Fairness Loss Function and Societal Acceptance Loss Function on PEGYPT dataset, in terms of how the incentivization budget was consumed on which types of users. For instance, Figure 3.9 on the left side, explains how the top 200 (with higher incentive values) users engaged with what societal circles. We observed that the Fairness Loss-based LA intervention consumed most of the incentivization budget on users who contributed to around 60% of the engagements in circles F and C, while circle F was the most harmful circle and circle C also had a different bias than the truth campaign. On the contrary, the Societal Acceptance Loss-based LA intervention consumed most of the budget on less than 30% of these circles' engaged users. The latter behavior indicates how the temporal bias and societal circles' information matured the truth campaign more and incentivized users based on the probabilities of accepting the campaign incentive and being accepted by others. Moreover, on the right side of Figure 3.9, it shows how the Societal Acceptance Loss-based LA intervention highly prioritized the smaller amounts of incentives (LAs with significant small ϕ values) for users that engaged with the circles F, C.

Additionally, Figure 3.10 illustrates our approach to influencing users to join

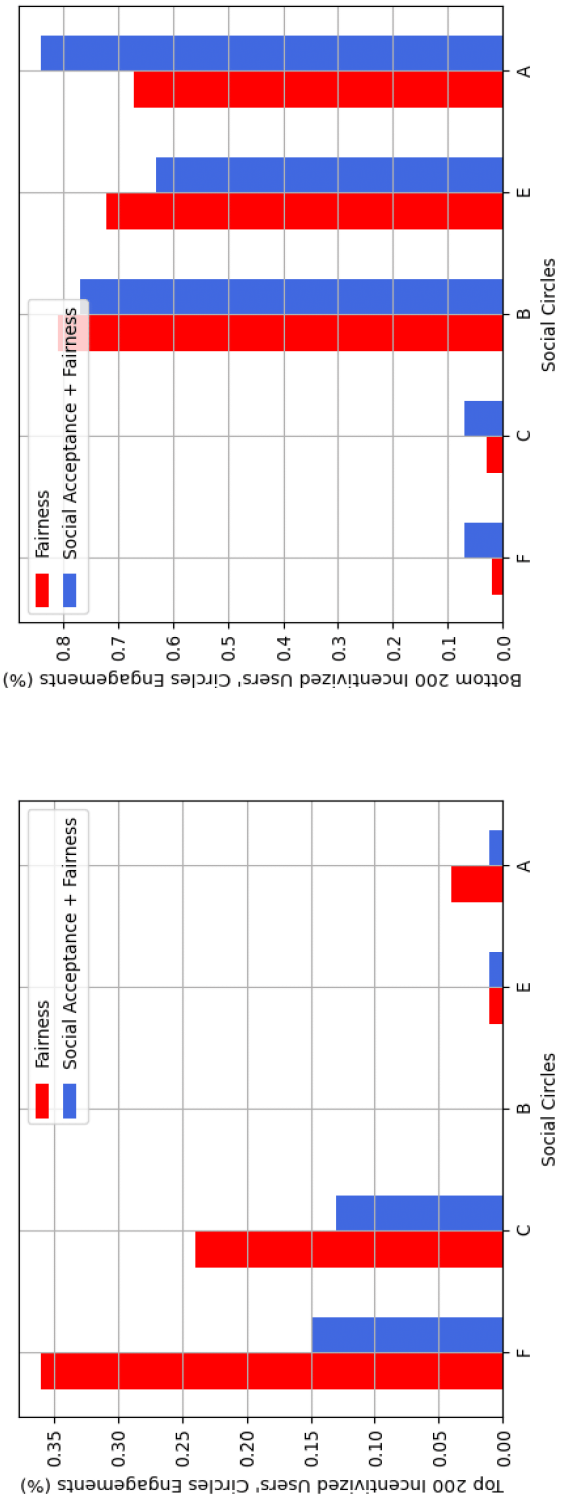


Figure 3.9: Optimization evaluation of the previously successful Fairness Loss Function and its extended capabilities in Societal Acceptance Loss, on how truth campaign incentives were assigned to the social network users on PEGYPT dataset

circle B by actively participating in and accepting its ideology, rather than engaging through disagreement. Consequently, we observed a rise in both the probability of being biased against (being in opposition to an idea where circle B opposes the same) and the probability of engaging with circle B. The latter represented modeling the engagement occurrence, while the former modeled the acceptance of that engagement since circle B represented a bias-against concept (same as truth campaign bias in **Paper G** experiments). This analysis also demonstrated how incorporating temporal bias and engagements within societal circles enabled us to track and analyze user activities associated with different diffusion groups in the context of misinformation propagation.

Finally, on the left side of Figure 3.11, we show how the LAs environment, characterized by our societal acceptance representation, was more strict and gave relatively fewer rewards. We attribute this outcome to the incorporation of a greater number of interdependent variables within the Societal Acceptance Loss Function. Nonetheless, this rigidity proved beneficial in mitigating polarization, fostering societal acceptance, and providing slightly higher confidence in the learned incentives. On the right side of Figure 3.11, to measure the level of uncertainty in the solution, we evaluate the Shannon entropy [164] of the probability distribution for individual incentives, which was obtained through Monte Carlo sampling. Notably, we observed a larger population of users with significantly lower entropies when the Societal Acceptance Loss Function was employed. These lower entropies indicate lower uncertainty in the learned incentives. To ensure clear differentiation in the results, the entropies shown in Figure 3.11 only pertain to users with discernibly different entropy values between the two loss functions.

For the complete details of the work done in **Paper G** and the proposed Societal Acceptance Loss Function and users' activities representation, see Appendix G.

3.3.1.4 Learning Automaton vs Policy Iteration

Table 3.2 demonstrates by how far our LA-based method outperformed random and uniform allocation of incentives for an intervention-based truth campaign to mitigate misinformation on SM. Further, we made an analogy of the results with previously proposed Policy Iteration-based methods for the same task. In particular, we compared with the following previously proposed methods:

- **EXP**: an exposure-based Policy Iteration algorithm [20] which depends on computing an exposure-based closeness centrality to model diffusion dynamics in terms of finding influential nodes on the social network.
- **V-MHP**: a vanilla MHP Policy Iteration algorithm [20], where a reward criterion was received from a single MHP diffusion group for information veracity related activities. Particularly, factual and misinformation activities.
- **U-MHP**: a user bias response MHP Policy Iteration algorithm [25], where another MHP diffusion group modeled the political bias activities in addition to the information veracity MHP diffusion group. However, **U-MHP** assumed

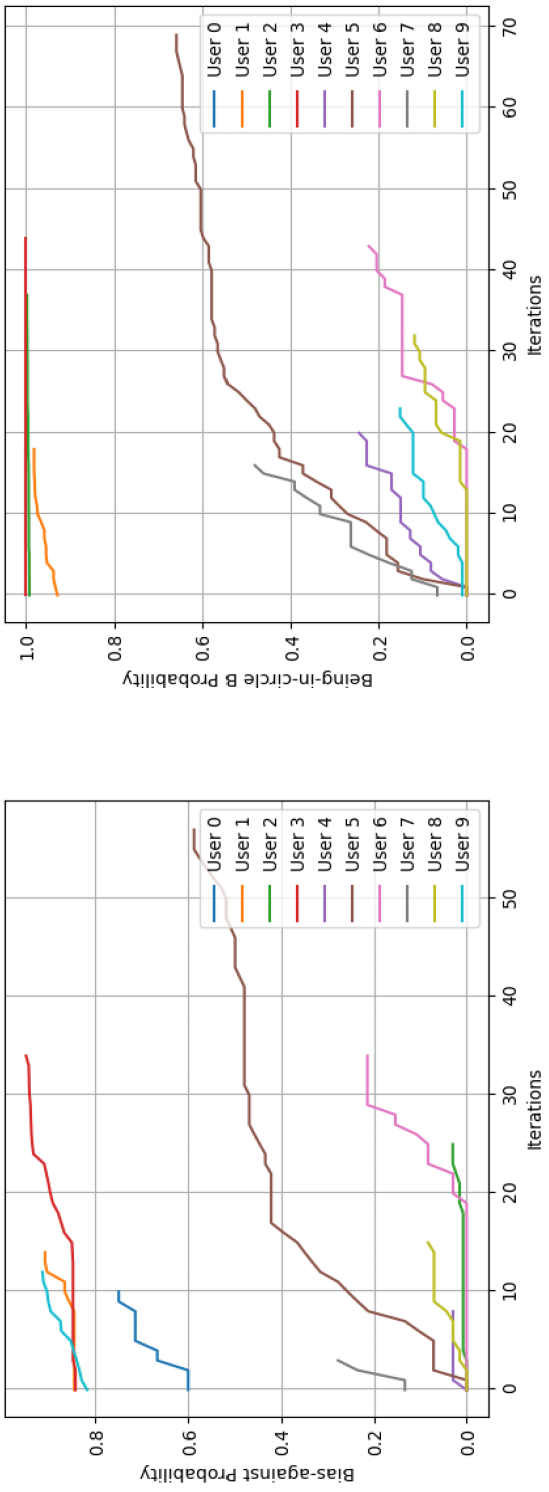


Figure 3.10: Example of breaking the harmful societal circles on PEGYPT dataset by incentivizing some users to alternatively accept and engage with the truth campaign circle B

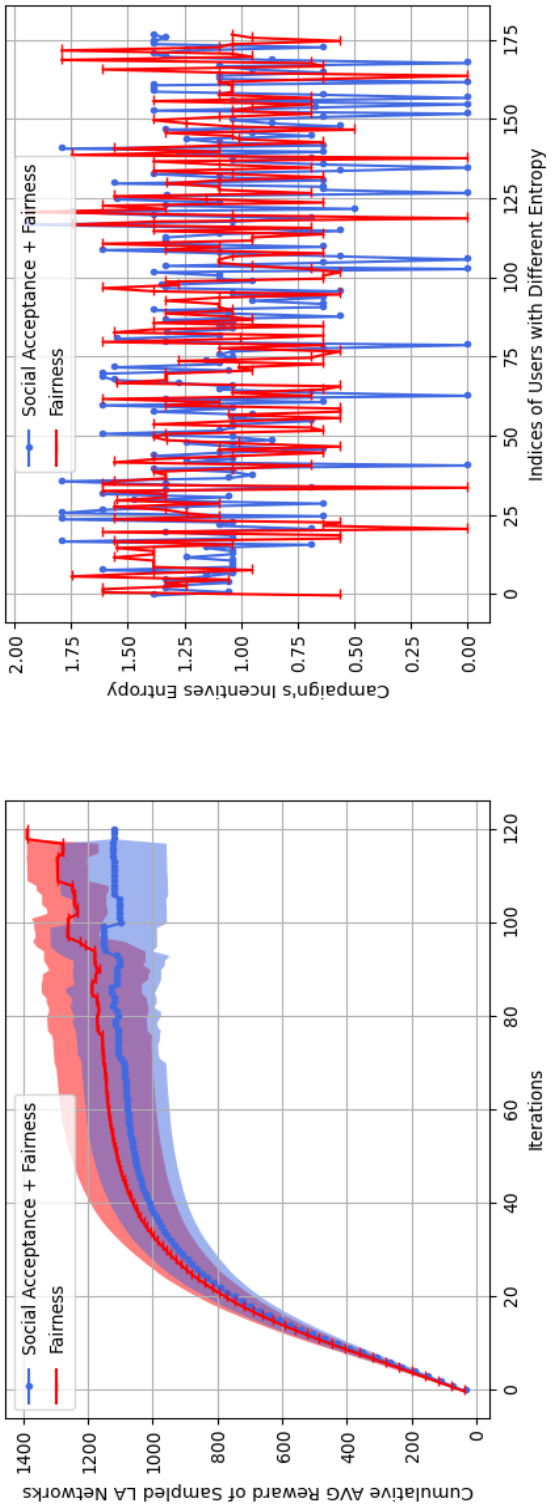


Figure 3.11: PEGYPT dataset evaluation for the previously successful Fairness Loss Function and its extended capabilities with societal acceptance awareness loss on the average cumulative rewards during incentive learning with entropy of the finally decided incentive values

static societal bias diffusion, where biases do not change once expressed. In contrast, our method incorporated temporal changes, capturing the evolving nature of biases and reflecting the dynamic influences within a social network.

Table 3.2: Relative performance of LA-based and Policy Iteration-based misinformation mitigation, against random and uniform methods

Model	***Tw15*-rnd**	***Tw15*-unif**	***Tw16*-rnd**	***Tw16*-unif**
LA-MHP	**2.37**	**2.11**	2.35	1.71
U-MHP	2.06	1.93	**3.20**	**1.80**
V-MHP	1.54	1.87	2.80	1.50
EXP	1.33	1.21	2.04	1.12

We referred to our method as LA-MHP, and the evaluation metric was a ratio between a correlation maximization Y obtained from each baseline method and a correlation maximization obtained from either random or uniform allocation of incentives when applied on two real datasets: Twitter-15 [32] and Twitter-16 [33].

The exposure counts to both factual information $\mathcal{T}$ (after applying the incentives) and misinformation $\mathcal{F}$ were considered the correlation variable and the constant, respectively. For instance, $Y = \mathcal{T} \times \mathcal{F}$, where the ratio that indicated how LA-MHP performed against the random allocation of incentives with regard to correlation maximization was calculated as $\frac{Y_{LA-MHP}}{Y_{rnd}}$, where Y was calculated twice for both LA-MHP and the random method over their incentivized and predicted MHP counts $\mathcal{T}$, and only predicted MHP counts $\mathcal{F}$. The results given in Table 3.2 proved how the proposed lightweight LA-MHP model was either outperforming or at least competing with all baselines.

3.3.2 Other Preliminary Models and Results

3.3.2.1 Trustworthiness Causal Graph

Paper A (see Appendix A) proposed a probabilistic graphical model as a theoretical view on the problem of normal users' credibility on SM during a political crisis, where polarization and deception were key properties. The paper introduced a hypothetical Causal Bayesian Network (BN) [165], inspired by the potential main entities that would be part of the online discussions dynamics. The causality-based approach provided a practical road map for some sub-problems in real-world scenarios such as individual polarization prediction, misinformation detection, and sensitivity analysis for the latter two tasks. Moreover, it facilitated intervention simulations with other causal entities if added to the causal graph, such as the learned incentives of a truth campaign. These intervention simulations could be used to verify the LAs' learned incentives from the algorithms proposed in **Papers B, D, E,** and **G**.

The proposed causal graph demonstrated a causal representation of our assumptions about the task of evaluating online users' credibility. It demonstrated the dependencies between the task's hypothetical entities (i.e., graph nodes). In a causal

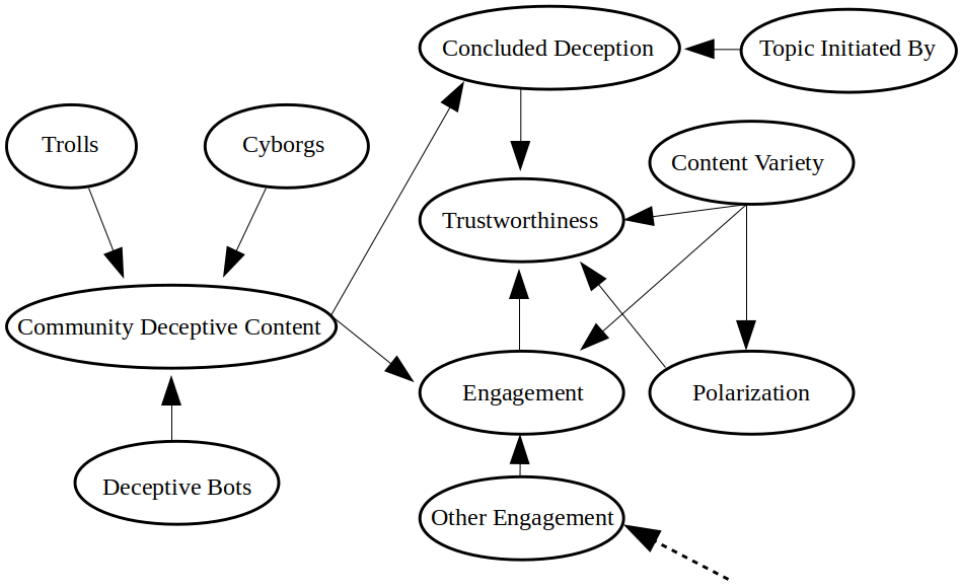


Figure 3.12: Hypothetical SM discussion causal graph

graph, edges from parent nodes to child nodes mean a causal relationship. That means a child node variable is considered as an effect of its parent node variable. In this manner, Figure 3.12 shows a causal graph of a social network from information veracity and users' trustworthiness perspectives.

As demonstrated in Figure 3.12, there were three main potential deceptive causes of manipulation in an online community (i.e., a societal circle): Trolls, Cyborgs, and Deceptive Social Bots. Because online users would agree on what was deceptively influencing them during their engagements with societal circles, users' trustworthiness degrees of their potentially propagated deception should be measured. Moreover, we hypothesized that polarization levels and the level of diversity to which content users are exposed will affect the trustworthiness degrees of users [12]. Further, we adopted the concept that users' levels of exposure to diversity can influence both polarization levels and societal engagements [166]. We also considered other societal engagements that might influence one's engagement like when a user replies to others while approving or denying their opinions. Finally, a topic initiative and a concluded deception opinion were proposed to collect more evidence about the trustworthiness degrees.

To quantify the dynamics inside the causal graph in Figure 3.12, we built a BN that can be utilized to learn the posterior joint probability distribution over the hypothetical network topology. For the complete details of the proposed BN and **Paper A**, see Appendix A.

3.3.2.2 Neural Emotion Hawkes Process

According to the definition of *emotional inertia* [167], conversation partners tend to stick to a particular emotional state, unless some external motivation excites them to change that state. Usually, the excitation comes from the other conversation partner. Therefore, by considering one partner as a chatbot agent [168], we conducted a

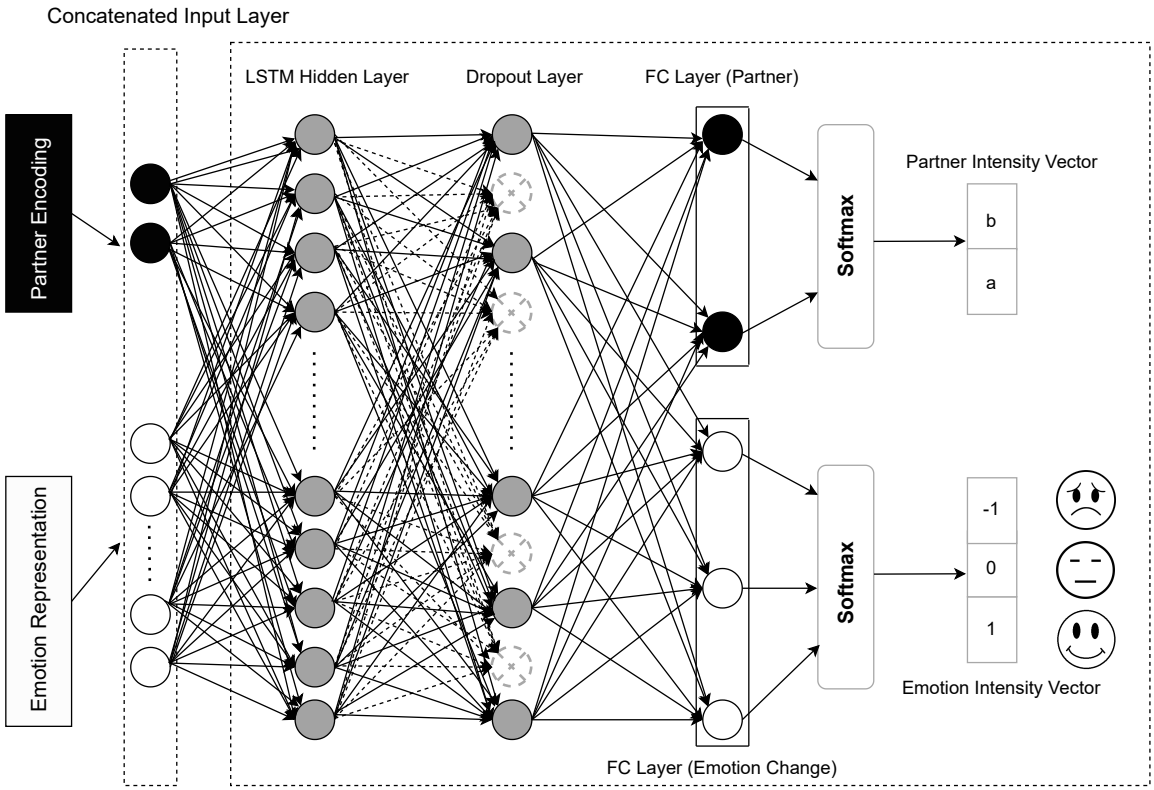


Figure 3.13: LSTM-based MHP for dyadic conversation emotion change prediction

preliminary study in **Paper C** (see Appendix C) to investigate how to learn hidden excitation patterns in emotional conversations, so the agent can control the outcome from other partner's emotions by re-planning its own (chatbot) expressed emotions.

Figure 3.13 shows a proposed End-to-End architecture of the Neural Emotion Hawkes Process (NEHP), where the latter utilized Long-Short Term Memory (LSTM) to predict emotion semantics in future conversation turns. We believed that the utilization of the LSTM would be sufficient as a preliminary study, where the conversation input sequence length was relatively short and the conversation expressed emotions were completely related, due to the complementary nature of emotions in dialogues [169]. However, self-attention mechanisms [170] are recommended for the analysis of long and multi-context conversations.

The NEHP was fed by two categories of inputs as follows. First, the two partners' one-hot-encoding vector over n steps, where n was the number of conversation turns, where multiple sequenced turns for one partner were considered as a single turn. Therefore, only one partner was represented in each turn whereas two sequenced turns must have two partners. For instance, the two partners' one-hot-encoding vector can be in the form: $(0, 1, 0, 1, 0, 1, ..., i_n)$, where each i^{th} turn represented a unique bit value for a partner. The second input category was the vector of emotion representation associated with each conversation turn. In the latter, each vector entry represented a particular emotion expressed by a conversation partner in a turn and quantified by a scalar value based on an emotion taxonomy representation technique. To learn the hidden excitation between the expressed emotions

Table 3.3: Prediction accuracy of other conversation partner over two turns in future, and P-values obtained from excitation trials

Dataset	Emotion Representation	Acc (1)	Acc (2)	Avg Acc	Excit
Topical	partner difference	0.82	0.51	0.67	**0.16**
Movies	partner difference	0.80	0.35	0.57	0.40
Topical	incremental partner difference	0.89	**0.62**	**0.76**	0.05
Movies	incremental partner difference	0.82	0.20	0.51	**0.45**
Topical	partners difference	0.85	0.54	0.64	0.06
Movies	partners difference	0.81	0.37	0.59	0.24
Topical	incremental partners difference	**0.94**	0.57	**0.76**	0.06
Movies	incremental partners difference	**0.92**	**0.76**	**0.84**	0.16

between two conversation partners, the NEHP utilized a non-parametric MHP over the hidden state vector of the LSTM. In this manner, the discrete-time intervals in the MHP setting were defined as the turns indices of the dyadic conversation where even indices were associated with one partner while the odd ones represented the other partner.

The work done in **Paper C** investigated how emotions in the NEHP model input should be represented to allow for high prediction accuracy and learning of optimal excitation strategies from intervention models. See subsubsection 2.2.3.1 for more details about excitations in HPs. To facilitate interventions with the NEHP, a RL chatbot agent can interact with the former as the RL emotion dynamics environment, and then, learn optimal emotion expression strategies. To evaluate a potential successful integration with chatbot agents in further studies, we manually re-planned the expressed emotions of one partner in the NEHP emotion sequence input. We defined the re-planning of one partner's emotions as the replacement of a particular emotion in its associated conversation utterances sequence. Then, we evaluated if such a plan would succeed to excite the other partner for a more positive emotional outcome. Hence, we conducted a T-test [171] and accepted only significant emotion change outcomes with $P\text{-}value \leq 0.05$. Then, we calculated the percentage of these successful excitation trials that influenced the emotional flow of the conversation. Two public text-based conversational datasets were studied for the preliminary experiments, one was from imagined conversations in movies [172], and the other was sentiment-annotated human-to-human conversations [173].

Table 3.3 shows the preliminary results for our proposed NEHP where four emotion representation techniques were evaluated. These results indicated promising prediction accuracy and the possibility of learning emotion excitation patterns during a dyadic conversation to help mitigate extreme polarization in online discussions. The learned excitations were not a result of any supervised learning as practiced by Poria et al. (2021) [51], since the only given ground-truths were partner ids and their associated emotion semantic representation. Such capability of the NEHP highlighted how the utilization of MHP and LSTM together was beneficial for the task.

The emotion representation methods varied from each other in terms of how the emotion taxonomies scalar values sequence was calculated. For example, "partner difference" meant that we only focused on the sequence of turns associated with each partner alone and calculated the differences between every two sequent turns of the same partner. So, we obtained a transformed sequence that represented an emotion change pattern for each partner independently from the other partner, where at each associated turn, we captured how emotions changed from the previous turn of the same partner. Additionally, the "partners difference" representation meant that we calculated the differences from every two adjacent turns' emotion scalar values. Since every two adjacent turns in our setting were associated with two different partners, the latter technique allowed for creating dependencies between partners in the representation. Finally, the incremental technique was applied to the two mentioned representation approaches but instead of calculating the differences between every two turns, differences were calculated accumulatively, hence, it could capture longer dependencies over conversation turns instead of only focusing on a previous turn. For the complete details of the proposed NEHP, see Appendix C.

3.3.2.3 Self-supervised Learning

As introduced earlier in chapter 1, fact-checking methods on SM are judgmental because they provide a definite classification of users' contents, and any error in their outcomes will violate the freedom of speech. The latter is considered an ethical concern about machine learning classifiers [19]. Moreover, as introduced in chapter 2, MHPs are highly dependent on annotated temporal data to learn the diffusion prediction function. Thus, we were motivated to study novel methods to provide both transparency and automatic data annotation for intervention-based misinformation mitigation pipelines. To this end, **Paper F** proposed a novel architecture inspired by the established Tsetlin Machine (TM) [174] to self-learn data categories. We referred to our method as the Label-Critic TM where the standard TM capacity was extended to perform classification or clustering tasks without the guidance of ground truths. To achieve that, as in the standard TM where Tsetlin Automata (TAs) are attached on a feature level, we further attached TAs on a sample level. Hence, each data sample was assigned a single TA to learn its correct label. The samples' assigned TAs shared the same structure and each had one of two possible decisions to make: either the sample belongs to class A or B. In this study, we interchangeably used the terms class and cluster to refer to the same concept. Our approach was similar to top-down hierarchal clustering [175], where all data samples were considered as one large cluster before splitting them into smaller ones until a cluster converges without further splitting.

Since a standard TM performs binary classification, we recursively split the data into two groups: A and B, where each had the group-associated positive sub-pattern(s) learned by a standard TM. Then, another standard TM learned discriminative sub-pattern(s) for A and B, separately. For each investigated data group, If the standard TM found no discriminative sub-patterns, the current group converged

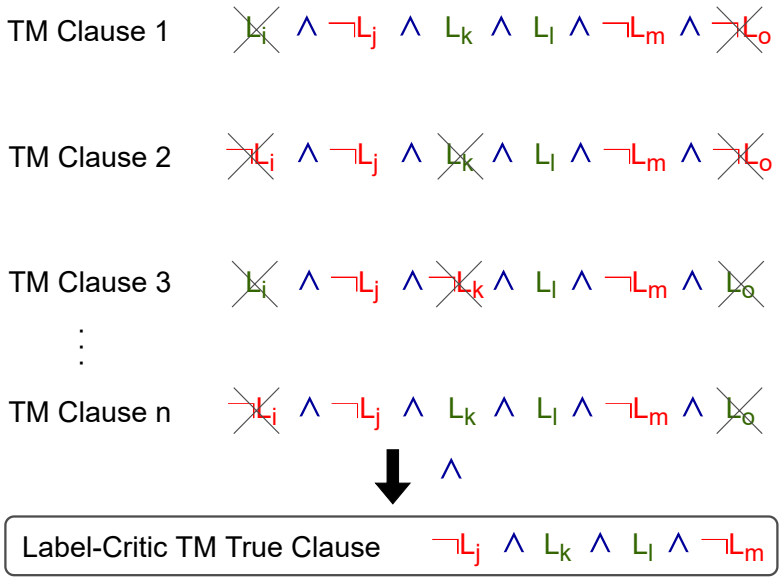


Figure 3.14: Novel Label-Critic TM invokes the True Clause generation by removing contradicted literals. The figure shows the same class clauses contradiction removal procedure

as one cluster, otherwise, more training loops were performed to further split. Finally, data samples in the learned clusters can be labeled according to the converged cluster. For instance, for four learned clusters, their underlying sample labels could be $A^{'}, A^{''}, B^{'}, B^{''}$, where $A^{'}$ and $A^{''}$ are child clusters from their original parent cluster node A. Similarly for $B^{'}$ and $B^{''}$, as child clusters from their original parent cluster node B.

In a standard TM, the training data feature space is represented by binary vectors, where a value of 1 indicates the occurrence of a feature and a value of 0 indicates the absence of the feature for a given data sample. Further, a standard TM conducts binary classification through the so-called "learned clauses". A clause is a conjunction of literals, where a literal is a propositional feature input $(L : 1)$ or its negation $(\neg L : 0)$. Hence, the final sub-pattern recognition task can be performed according to a voting scheme from all learned clauses against a given data sample to be classified. For the proposed Label-Critic TM, when learning a sample label in a data group, a Label-TA was a single TA assigned to the sample with a randomly determined initial label. The Label-TA then sent its decision to the standard TM team, where the latter learned the TM clauses that represent the sub-patterns. In parallel, a Critic-TA task was to clone the current decided label by the Label-TA and send the label to another standard TM team to perform separate clause learning. The duality of Label-TM/Critic-TM ensured validation of the final learned clauses. Both Label-TA and Critic-TA validated their decisions according to a final True Clause generated by their associated TMs after some post-processing as demonstrated in Figure 3.14.

A True Clause was the final learned literals that suggested a final pattern of a cluster, so when data samples shared a particular True Clause pattern, they

were considered as a cluster. The latter was an example of self-corrected learning facilitated by the standard TM propositional logic approach since noisy literal-contradicted clauses can be corrected and concluded into one logically True Clause. As seen in Figure 3.14, the Label-Critic TM invoked a class True Clause generation by removing contradicted clauses' literals caused by noisy training data in the standard TM because of the randomly initialized data labels. However, the contradictions would exist in literals between the two classes' clauses as well, in that case, a final True Clause per each class was concluded according to the discriminative literals. For instance, if class A had the positive literal L_i, then the class B True Clause cannot have the same literal unless it was negated.

As indicated in Figure 3.14, the generated clauses by the standard TMs would have contradicted literals due to the randomly assigned labels. However, propositional logic-based learning equipped the Label-Critic TM with self-correction capabilities. For instance for n learned clauses by a standard TM team for a particular class, if the k^{th} feature had the propositions $L_k, L_k, \neg L_k, L_k$, it was concluded to L_k, L_k since L_k was canceled only once by its contradiction $\neg L_k$. At each training epoch of the Label-Critic TM, the acquired True Clause was accumulated to previous ones, and the same literal can have multiple instances in that case, making the True Clause more informative and tuned by adding weighted literals. The generated True Clause literals were then used for a similar voting scheme, just like a standard TM but on a literal level. The latter voting mechanism determined whether a data sample should belong to which of the two clusters based on the two True Clauses. Each Label-TA evaluated its label selection and was rewarded if and only if: its decision matched the cluster which got the higher votes, and its twin Critic-TA decided the same label for the associated data sample, otherwise penalized. Critic-TAs rewards were received the same way as in Label-TAs except that they did not require evaluation of their twin Label-TAs. Eventually, it became possible to obtain transparency of the clustering results through the propositional logical statements embedded in the True Clauses associated with each converged cluster.

Figure 3.15 demonstrates the fundamental components and recursive cycles of the proposed Label-Critic TM where a recursive binary classification-based splitting was performed over the unlabeled data samples. For further details on True Clause generation and the cluster voting scheme, see Appendix F.

Figure 3.16 shows how the Label-Critic TM along with the DBSCAN method outperformed other clustering methods, including hierarchal-based clustering such as the Agglomerative method. Two synthetic data groups were evaluated for such experiments. The first data group represented small balanced datasets with 400 features and 300 samples per cluster (class). The second data group represented relatively larger and unbalanced datasets with several features up to $2,000$, and the number of samples varied among clusters. For instance, one cluster had $1,000$ samples while another had only 300 only. Further, the learning loops of the Label-Critic TM can be observed in Figure 3.17 where the learning was achieved through optimizing the total number of penalties from the Label-TAs. The number of original clusters in data dictated how many standard TM training loops were required to

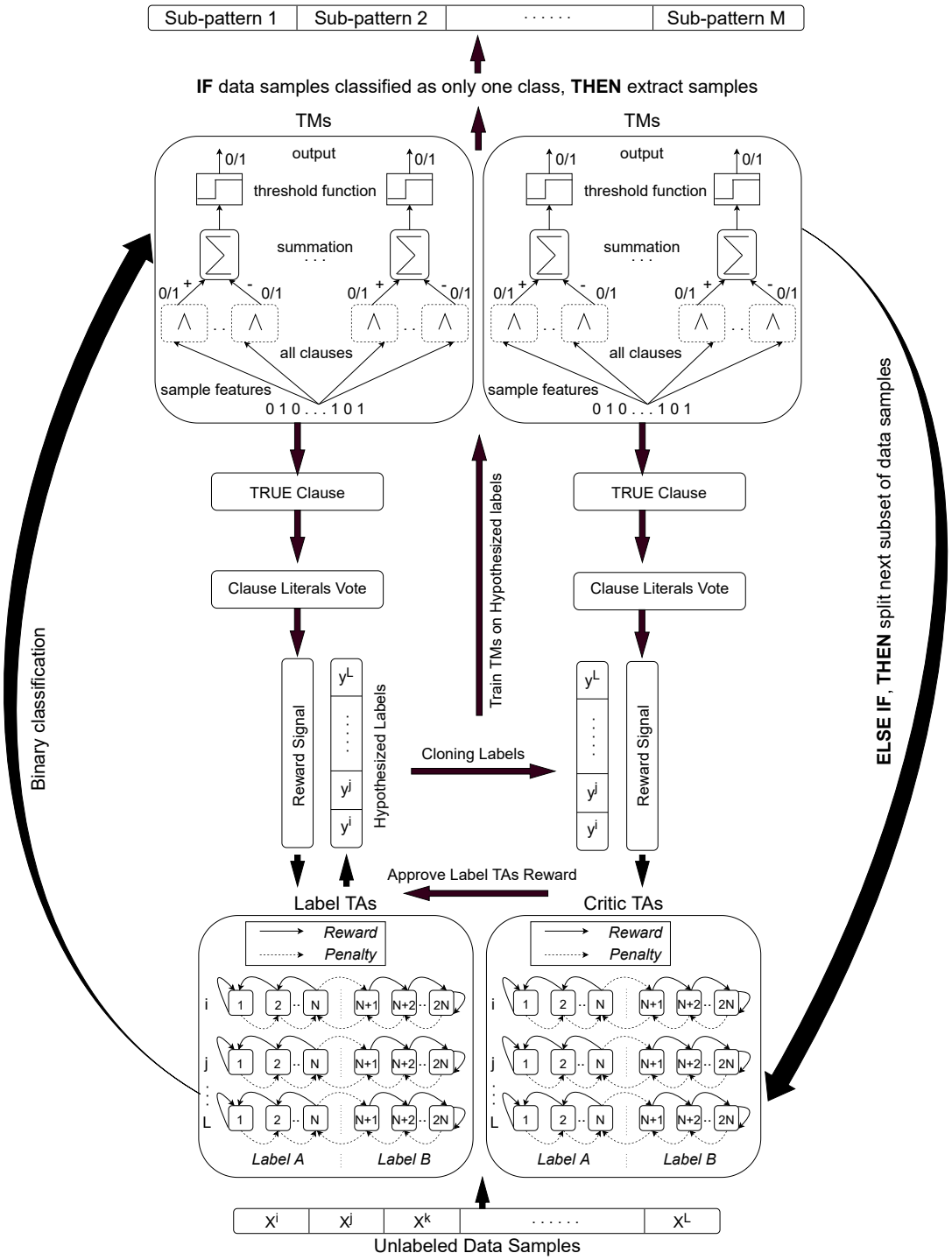


Figure 3.15: Novel recursive self-supervised Label-Critic TM architecture

keep splitting the data into smaller clusters. Loops would eventually converge after some epochs. In some cases, due to the randomness of label initialization, a loop would be repeated if it reached a particular epoch (threshold) without convergence.

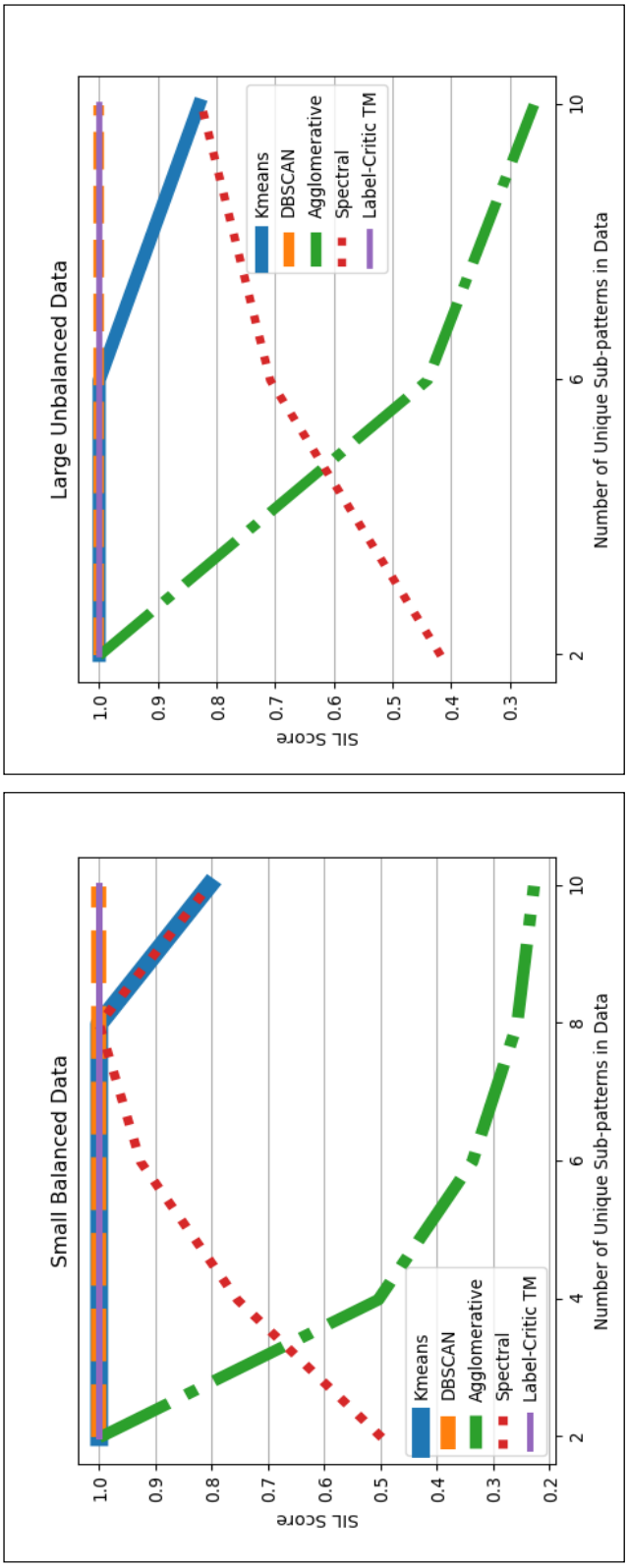


Figure 3.16: Label-Critic TM performance comparison with benchmark clustering algorithms

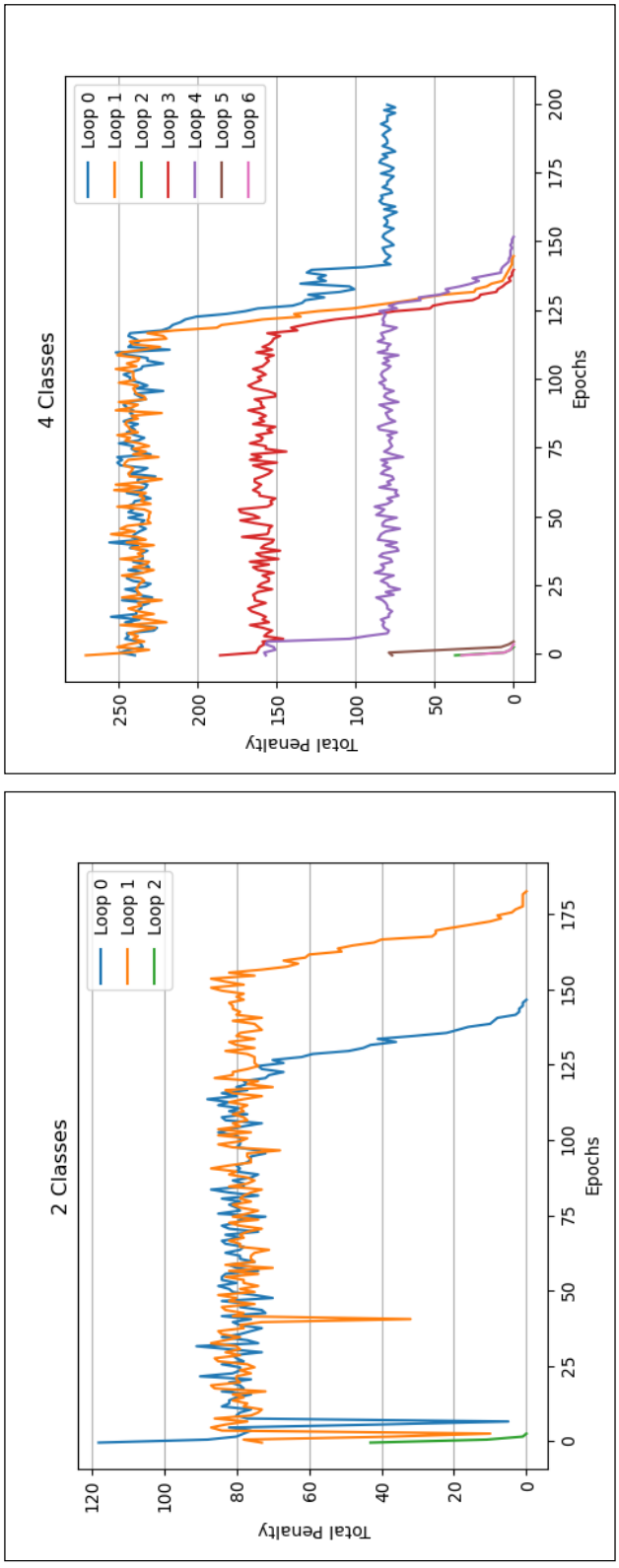


Figure 3.17: Recursive self-supervised Label-Critic TM learning curve of conducted loops over a different number of sub-patterns, with one sub-pattern per cluster

Chapter 4

Conclusion

In the past decade, intervention-based misinformation mitigation was proposed as an approach to combat misinformation on SM. In this manner, misinformation mitigation is achieved by boosting the exposure to factual information. To boost factual information on the social network, users are incentivized to propagate facts, with different incentivization levels determined based on users' needs and vulnerability to misinformation. Traditionally, intervention-based learning of incentives was performed through Reinforcement Learning (RL), where the majority of the proposed methods were Policy Iteration algorithms that learned an optimal mitigation policy for a so-called truth campaign. The relatively narrow focus on Policy Iteration algorithms led us to investigate alternative methods, leveraging recent insights on the problem of misinformation propagation on SM. For instance, two crucial questions are what incentivization strategies are most effective and which factors should be optimized for effective mitigation, apart from the individual user's needs and their vulnerability to misinformation. To address these questions, this dissertation proposed novel Learning Automaton (LA) algorithms to learn user-directed truth campaign incentives on SM, introducing novel criterion functions based on a recent understanding of SM dynamics. To learn the optimal incentivization for each user with limited incentivization resources overall, we defined the problem as a knapsack constraint optimization procedure. Then, the incentivization efforts were limited to a predefined knapsack capacity. Further, we integrated our knapsack-based model with novel modeling of temporal users' activity and optimization variables, developing dedicated parametric Multivariate Hawkes Process (MHP). The latter was our adopted information diffusion technique to simulate and predict social network activities. Then, the hypothesized incentives can be evaluated for their effectiveness over the simulated social network.

In this chapter, we present the key findings of our study and demonstrate the possible impacts of these on the current challenges of Artificial Intelligence (AI)-based misinformation mitigation on SM. Finally, we highlight the limitations of our research and suggest future research directions.

4.1 Key Findings in the Study

According to our investigation, the utilization of LA algorithms as an easily implementable and lightweight optimization framework was beneficial to boost the factual information on simulated social networks. This was particularly the case for reward-penalty LA schemes with a variable structure. In most cases, the latter showed a significant performance advantage compared to other LA schemes or traditional RL algorithms such as Policy Iteration. However, the proposed LA algorithms did not give faster convergence to the optimal incentives unless they followed a greedy probabilistic state transition strategy. The probabilistic greedy state transition strategy meant that a transition with higher probability was always performed, without exploring other possible moves over the discrete state space. Further, the proposed LA learning schemes had a dynamic rate of state transition probability updating that pursued the transition with the higher reward. We found that the combination of fully greedy exploitation, variable structure, and maximal reward pursuit-based state transition provided significantly faster convergence. The latter behavior was obtained in a stochastic and non-stationary MHP-based environment.

Further, traditionally, a loss function drives learning in RL and LA, translating the environment signal into simple rewards/penalties. The loss function thus helps the intelligent agents find proper state transitions. Accordingly, it solves its task by minimizing the loss function. Therefore, the loss function must provide a sufficiently comprehensive and rich representation of the environment to build a trustworthy and traceable solution. The loss functions should further be interpretable through the transparency of their dynamics if monitored during the optimization procedure. We found that SM misinformation activities and their mitigation should be modeled with the interdependent patterns that govern the propagation of misinformation. That is, interconnecting the temporal variables for societal bias, societal engagement, and information veracity was vital to provide a convenient criterion and traceable interventions for the incentive-learning agents.

Finally, because of the nature of the truth campaign technique and how it aims to deliver factual information to SM users, the achievement of societal fairness in the proposed algorithms is crucial and could be obtained through dedicated expressions that facilitate fairness or other societal objectives. The latter was practiced in the proposed fairness and societal acceptance-aware optimization loss functions.

4.2 Impact

The work done in this research study contributed to the following areas. First, the study contributed to the scientific knowledge of LA by providing novel techniques to variable structure LAs. Our novel LA technique demonstrated how a combination of the estimator and variable structure schemes was beneficial to the task of knapsack constraint stochastic optimization, with a multiple-dimensional random walk. The latter also showed how more complicated knapsack settings will require new approaches in the design of the proposed LA, unlike traditional interaction between

LAs and knapsack-based environments where the latter needed only a deterministic LA in terms of its state transition selection.

Second, the LA proposed algorithms showed robustness after being evaluated on multiple misinformation datasets with statistical variations regarding the percentage of misinformation exposure. Such robustness is crucial for generalizing the proposed misinformation mitigation optimization functions to unseen social network statistics.

Third, unlike only conducting empirical experiments for the proposed AI models on existing datasets, the study collected more descriptive and representative data to establish a more realistic learning environment and objective for the proposed learning agent. In this manner, the study bridged AI and Social Science by investigating relevant theoretical studies on the problem of misinformation propagation on SM and the interdependent dynamics that govern such propagation. All the above led to a novel computation model that encapsulated Social Science theories and allowed for their empirical applications and a novel misinformation dataset.

Fourth, we contributed with the Multiplex-Controlled Multivariate Hawkes Processes (MCMHP), a mathematical framework to model problems where LAs are required to optimize a volume-based diffusion model to boost the counts of a particular activity on a social network. In our proposed MCMHP, different components can be easily replaced by a more suitable one depending on the task. For instance, the MCMHP framework can utilize other proposed loss functions, LA schemes, and volume-based diffusion models.

4.3 Limitation and Future Work

The proposed SM temporal users' activity representation was important to extend the analytical capacity to the truth campaign optimization procedure. However, this research study did not investigate the improvement of these activity diffusion dynamics in terms of the diffusion prediction approach. The latter is a deep topic and requires an entirely separate study. For instance, future work in this regard can focus on multimodal diffusion modeling where both linguistic and temporal features can be concatenated to predict future diffusions. Moreover, instead of utilizing the classical parametric MHPs, neural network-based MHPs can be investigated where both internal and external influence parameters become dynamic, instead of representing them with fixed values.

Users' activities on SM can have multiple cascades, e.g., retweets of retweets on Twitter. The latter means that exposure to some influence can be indirect through the indirect relationships over the multiple cascades. In this manner, the investigations done in this dissertation have only focused on one level of cascades, where exposure to influence was measured according to the direct relationship in the network, e.g., when two users have a following relationship on Twitter. Hence, proposing exposure measures and optimization over multiple cascaded influences is an interesting future contribution to our study.

Despite the proposed sampling techniques in **Papers D, E** and **G** to facilitate

robustness on large-scale social networks, the study was constrained by the limited size of the available misinformation datasets. Investigating larger scale data, involving millions of users, will be beneficial for further improvement and understanding of the proposed misinformation mitigation framework. For instance, how will our framework behave in an increased optimization space, and will it be possible to build an efficient neural diffusion model for scalability?

Another limitation was how the work evaluated the application of truth campaigns during an assumed normal scenario when misinformation unintentionally circulates. However, there could be other scenarios when a truth campaign challenges another manipulation campaign at the same time. For example, when deceptive bots spread misinformation in the same time intervals as the truth campaign. The evaluation of the truth campaign incentivization procedure will be very interesting in such adversarial settings.

Finally, the study contributed with preliminary algorithms and structures to verify the learned truth campaign incentives, and to provide interpretable self-learning for the detection of some categorical data such as SM activities. Therefore, a future contribution can be extending these preliminary models to provide a complete analysis of the dynamics of users' activities with the learned incentives, where verifications of optimized misinformation mitigation strategies would be possible. Moreover, matured models for interpretable self-learning of SM activities such as misinformation, factual information, polarization directions, and societal engagement — will empower the misinformation mitigation pipeline with independence and transparency.

Bibliography

[1] Joshua A Tucker, Andrew Guess, Pablo Barberá, Cristian Vaccari, Alexandra Siegel, Sergey Sanovich, Denis Stukal, and Brendan Nyhan. Social media, political polarization, and political disinformation: A review of the scientific literature. *Political polarization, and political disinformation: a review of the scientific literature (March 19, 2018)*, 2018.

[2] Paolo Montemurro, Ales Porcnik, Per Hedén, and Maximilian Otte. The influence of social media and easily accessible online information on the aesthetic plastic surgery practice: literature review and our own experience. *Aesthetic plastic surgery*, 39:270–277, 2015.

[3] Raina M Merchant and Nicole Lurie. Social media and emergency preparedness in response to novel coronavirus. *Jama*, 323(20):2011–2012, 2020.

[4] Yasmim Mendes Rocha, Gabriel Acácio de Moura, Gabriel Alves Desidério, Carlos Henrique de Oliveira, Francisco Dantas Lourenço, and Larissa Deadame de Figueiredo Nicolete. The impact of fake news on social media and its influence on health during the covid-19 pandemic: A systematic review. *Journal of Public Health*, pages 1–10, 2021.

[5] Irina Khaldarova and Mervi Pantti. Fake news: The narrative battle over the ukrainian conflict. In *The Future of Journalism: Risks, Threats and Opportunities*, pages 228–238. Routledge, 2020.

[6] Sora Park, Caroline Fisher, Terry Flew, and Uwe Dulleck. Global mistrust in news: The impact of social media on trust. *International Journal on Media Management*, 22(2):83–96, 2020.

[7] Clayton Wukich and Ines Mergel. Reusing social media information in government. *Government Information Quarterly*, 33(2):305–312, 2016.

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

[9] Matteo Cinelli, Gabriele Etta, Michele Avalle, Alessandro Quattrociocchi, Niccolò Di Marco, Carlo Valensise, Alessandro Galeazzi, and Walter Quattrociocchi. Conspiracy theories and social media platforms. *Current Opinion in Psychology*, page 101407, 2022.

[10] Gregory Eady, Jonathan Nagler, Andy Guess, Jan Zilinsky, and Joshua A Tucker. How many people live in political bubbles on social media? evidence from linked survey and twitter data. *Sage Open*, 9(1):2158244019832705, 2019.

[11] Michela Del Vicario, Antonio Scala, Guido Caldarelli, H Eugene Stanley, and Walter Quattrociocchi. Modeling confirmation bias and polarization. *Scientific reports*, 7(1):1–9, 2017.

[12] Patricia Moravec, Randall Minas, and Alan R Dennis. Fake news on social media: People believe what they want to believe when it makes no sense at all. *Kelley School of Business research paper*, (18-87), 2018.

[13] Patricia Rossini. Beyond incivility: Understanding patterns of uncivil and intolerant discourse in online political talk. *Communication Research*, 49(3):399–425, 2022.

[14] Laurenz Ennser-Jedenastik, Christina Gahn, Anita Bodlos, and Martin Haselmayer. Does social media enhance party responsiveness? how user engagement shapes parties' issue attention on facebook. *Party Politics*, 28(3):468–481, 2022.

[15] James A Piazza. Fake news: the effects of social media disinformation on domestic terrorism. *Dynamics of Asymmetric Conflict*, 15(1):55–77, 2022.

[16] Kai Shu, Suhang Wang, Dongwon Lee, and Huan Liu. *Disinformation, misinformation, and fake news in social media*. Springer, 2020.

[17] Sonal Garg and Dilip Kumar Sharma. New politifact: a dataset for counterfeit news. In *2020 9th International Conference System Modeling and Advancement in Research Trends (SMART)*, pages 17–22. IEEE, 2020.

[18] D Graves. Understanding the promise and limits of automated fact-checking. 2018.

[19] Derek Leben. Normative principles for evaluating fairness in machine learning. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, pages 86–92, 2020.

[20] Mehrdad Farajtabar, Jiachen Yang, Xiaojing Ye, Huan Xu, Rakshit Trivedi, Elias Khalil, Shuang Li, Le Song, and Hongyuan Zha. Fake news mitigation via point process based intervention. In *International conference on machine learning*, pages 1097–1106. PMLR, 2017.

[21] Ole-Christoffer Granmo, B John Oommen, Svein Arild Myrer, and Morten Goodwin Olsen. Learning automata-based solutions to the nonlinear fractional knapsack problem with applications to optimal resource allocation. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37(1):166–175, 2007.

[22] Sakshini Hangloo and Bhavna Arora. Content-based fake news detection using deep learning techniques: Analysis, challenges and possible solutions. In *2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT)*, pages 411–417. IEEE, 2022.

[23] Mohammad Hammas Saeed, Shiza Ali, Jeremy Blackburn, Emiliano De Cristofaro, Savvas Zannettou, and Gianluca Stringhini. Trollmagnifier: Detecting state-sponsored troll accounts on reddit. In *2022 IEEE Symposium on Security and Privacy (SP)*, pages 2161–2175. IEEE, 2022.

[24] Yik Chan Chin, Ahran Park, and Ke Li. A comparative study on false information governance in chinese and american social media platforms. *Policy & Internet*, 2022.

[25] Mahak Goindani and Jennifer Neville. Social reinforcement learning to combat fake news spread. In *Uncertainty in Artificial Intelligence*, pages 1006–1016. PMLR, 2020.

[26] Moti Michaeli and Daniel Spiro. From peer pressure to biased norms. *American Economic Journal: Microeconomics*, 9(1):152–216, 2017.

[27] Ayesha S Dina and D Manivannan. Intrusion detection based on machine learning techniques in computer networks. *Internet of Things*, 16:100462, 2021.

[28] Mei Li, Xiang Wang, Kai Gao, and Shanshan Zhang. A survey on information diffusion in online social networks: Models and methods. *Information*, 8(4):118, 2017.

[29] Davinder Kaur, Suleyman Uslu, Kaley J Rittichier, and Arjan Durresi. Trustworthy artificial intelligence: a review. *ACM Computing Surveys (CSUR)*, 55(2):1–38, 2022.

[30] Xiaofei Xu, Ke Deng, and Xiuzhen Zhang. Identifying cost-effective debunkers for multi-stage fake news mitigation campaigns. In *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, pages 1206–1214, 2022.

[31] Yalcin Akcay, Haijun Li, Susan H Xu, et al. Greedy algorithm for the general multidimensional knapsack problem. *Annals of Operations Research*, 150(1):17–29, 2007.

[32] Xiaomo Liu, Armineh Nourbakhsh, Quanzhi Li, Rui Fang, and Sameena Shah. Real-time rumor debunking on twitter. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, pages 1867–1870, 2015.

[33] Jing Ma, Wei Gao, Prasenjit Mitra, Sejeong Kwon, Bernard J Jansen, Kam-Fai Wong, and Meeyoung Cha. Detecting rumors from microblogs with recurrent neural networks. 2016.

[34] William Yang Wang. " liar, liar pants on fire": A new benchmark dataset for fake news detection. *arXiv preprint arXiv:1705.00648*, 2017.

[35] Fatima K Abu Salem, Roaa Al Feel, Shady Elbassuoni, Mohamad Jaber, and May Farah. Fa-kes: A fake news dataset around the syrian war. In *Proceedings of the international AAAI conference on web and social media*, volume 13, pages 573–582, 2019.

[36] Kai Shu, Deepak Mahudeswaran, Suhang Wang, Dongwon Lee, and Huan Liu. Fakenewsnet: A data repository with news content, social context and spatialtemporal information for studying fake news on social media. *arXiv preprint arXiv:1809.01286*, 2018.

[37] Heidi A Vuletich and B Keith Payne. Stability and change in implicit bias. *Psychological science*, 30(6):854–862, 2019.

[38] John R Anderson. *How can the human mind occur in the physical universe?* Oxford University Press, 2009.

[39] Cynthia Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5):206–215, 2019.

[40] Kumpati S Narendra and Mandayam AL Thathachar. Learning automata-a survey. *IEEE Transactions on systems, man, and cybernetics*, (4):323–334, 1974.

[41] Baichuan Yuan, Hao Li, Andrea L Bertozzi, P Jeffrey Brantingham, and Mason A Porter. Multivariate spatiotemporal hawkes processes and network reconstruction. *SIAM Journal on Mathematics of Data Science*, 1(2):356–382, 2019.

[42] A Mirzaei Rahimi, Amir Ziaeddini, and Shu Gonglee. A novel approach to efficient resource allocation in load-balanced cellular networks using hierarchical drl. *Journal of Ambient Intelligence and Humanized Computing*, 13(5):2887–2901, 2022.

[43] Huan Liu, Qiang Chen, and Richard Evans. How official social media affected the infodemic among adults during the first wave of covid-19 in china. *International Journal of Environmental Research and Public Health*, 19(11):6751, 2022.

[44] Femi Olan, Uchitha Jayawickrama, Emmanuel Ogiemwonyi Arakpogun, Jana Suklan, and Shaofeng Liu. Fake news on social media: the impact on society. *Information Systems Frontiers*, pages 1–16, 2022.

[45] Geir Thore Berge, Ole-Christoffer Granmo, Tor Oddbjørn Tveit, Morten Goodwin, Lei Jiao, and Bernt Viggo Matheussen. Using the tsetlin machine to learn human-interpretable rules for high-accuracy text categorization with medical applications. *IEEE Access*, 7:115134–115146, 2019.

[46] Mauricio Reyes, Raphael Meier, Sérgio Pereira, Carlos A Silva, Fried-Michael Dahlweid, Hendrik von Tengg-Kobligk, Ronald M Summers, and Roland Wiest. On the interpretability of artificial intelligence in radiology: challenges and opportunities. *Radiology: artificial intelligence*, 2(3):e190043, 2020.

[47] Zexian Zeng, Xia Jiang, and Richard Neapolitan. Discovering causal interactions using bayesian network scoring and information gain. *BMC bioinformatics*, 17(1):1–14, 2016.

[48] Dragan S Pamučar, Darko Božanić, and Aca Ranđelović. Multi-criteria decision making: An example of sensitivity analysis. *Serbian journal of management*, 12(1):1–27, 2017.

[49] Hongyuan Mei and Jason M Eisner. The neural hawkes process: A neurally self-modulating multivariate point process. *Advances in neural information processing systems*, 30, 2017.

[50] Anbang Xu, Zhe Liu, Yufan Guo, Vibha Sinha, and Rama Akkiraju. A new chatbot for customer service on social media. In *Proceedings of the 2017 CHI conference on human factors in computing systems*, pages 3506–3510, 2017.

[51] Soujanya Poria, Navonil Majumder, Devamanyu Hazarika, Deepanway Ghosal, Rishabh Bhardwaj, Samson Yu Bai Jian, Pengfei Hong, Romila Ghosh, Abhinaba Roy, Niyati Chhaya, et al. Recognizing emotion cause in conversations. *Cognitive Computation*, pages 1–16, 2021.

[52] Firoj Alam, Fahim Dalvi, Shaden Shaar, Nadir Durrani, Hamdy Mubarak, Alex Nikolov, Giovanni Da San Martino, Ahmed Abdelali, Hassan Sajjad, Kareem Darwish, et al. Fighting the covid-19 infodemic in social media: A holistic perspective and a call to arms. In *ICWSM*, pages 913–922, 2021.

[53] Tian Lan, Gang Mao, Fei Xia, Shengqi Yu, Rishad Shafik, and Alex Yakovlev. An asynchronous tsetlin automaton architecture with integrated non-volatile memory. In *2022 International Symposium on the Tsetlin Machine (ISTM)*, pages 37–40. IEEE, 2022.

[54] Sabine Matook, Alan R Dennis, and Yazhu Maggie Wang. User comments in social media firestorms: A mixed-method study of purpose, tone, and motivation. *Journal of Management Information Systems*, 39(3):673–705, 2022.

[55] Hunt Allcott and Matthew Gentzkow. Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2):211–36, 2017.

[56] Kai Shu, Suhang Wang, Dongwon Lee, and Huan Liu. Mining disinformation and fake news: Concepts, methods, and recent advancements. In *Disinformation, misinformation, and fake news in social media*, pages 1–19. Springer, 2020.

[57] Fantahun Bogale Gereme and William Zhu. Fighting fake news using deep learning: Pre-trained word embeddings and the embedding layer investigated. In *2020 The 3rd International Conference on Computational Intelligence and Intelligent Systems*, pages 24–29, 2020.

[58] Yaqing Wang, Fenglong Ma, Zhiwei Jin, Ye Yuan, Guangxu Xun, Kishlay Jha, Lu Su, and Jing Gao. Eann: Event adversarial neural networks for multi-modal fake news detection. In *Proceedings of the 24th acm sigkdd international conference on knowledge discovery & data mining*, pages 849–857, 2018.

[59] Eslam Amer, Kyung-Sup Kwak, and Shaker El-Sappagh. Context-based fake news detection model relying on deep learning models. *Electronics*, 11(8):1255, 2022.

[60] Jeffrey Pennington, Richard Socher, and Christopher D Manning. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543, 2014.

[61] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*, 19(1):22–36, 2017.

[62] Lingwei Wei, Dou Hu, Yantong Lai, Wei Zhou, and Songlin Hu. A unified propagation forest-based framework for fake news detection. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 2769–2779, 2022.

[63] Ke Wu, Song Yang, and Kenny Q Zhu. False rumors detection on sina weibo by propagation structures. In *2015 IEEE 31st international conference on data engineering*, pages 651–662. IEEE, 2015.

[64] Qiao Zhang, Shuiyuan Zhang, Jian Dong, Jinhua Xiong, and Xueqi Cheng. Automatic detection of rumor on social network. In *Natural Language Processing and Chinese Computing*, pages 113–122. Springer, 2015.

[65] Savvas Zannettou, Tristan Caulfield, William Setzer, Michael Sirivianos, Gianluca Stringhini, and Jeremy Blackburn. Who let the trolls out? towards understanding state-sponsored trolls. In *Proceedings of the 10th acm conference on web science*, pages 353–362, 2019.

[66] Chengcheng Shao, Giovanni Luca Ciampaglia, Onur Varol, Kai-Cheng Yang, Alessandro Flammini, and Filippo Menczer. The spread of low-credibility content by social bots. *Nature communications*, 9(1):1–9, 2018.

[67] Zhiwei Jin, Juan Cao, Han Guo, Yongdong Zhang, and Jiebo Luo. Multimodal fusion with recurrent neural networks for rumor detection on microblogs. In *Proceedings of the 25th ACM international conference on Multimedia*, pages 795–816, 2017.

[68] Aseel Addawood, Adam Badawy, Kristina Lerman, and Emilio Ferrara. Linguistic cues to deception: Identifying political trolls on social media. In *Proceedings of the international AAAI conference on web and social media*, volume 13, pages 15–25, 2019.

[69] Nafiza Rahman, Maisha Maimuna, Afroja Begum, Md Ahmed, Mohammed Shamsul Arefin, et al. A survey of data mining techniques in the field of cyborg mining. In *Soft Computing for Security Applications*, pages 781–797. Springer, 2022.

[70] Alessandro Bessi and Emilio Ferrara. Social bots distort the 2016 us presidential election online discussion. *First Monday*, 21(11-7), 2016.

[71] Isa Inuwa-Dutse, Mark Liptrott, and Ioannis Korkontzelos. Detection of spam-posting accounts on twitter. *Neurocomputing*, 315:496–511, 2018.

[72] Samantha Bradshaw and Philip Howard. Troops, trolls and troublemakers: A global inventory of organized social media manipulation. 2017.

[73] Jon Roozenbeek, Sander Van Der Linden, Beth Goldberg, Steve Rathje, and Stephan Lewandowsky. Psychological inoculation improves resilience against misinformation on social media. *Science advances*, 8(34):eabo6254, 2022.

[74] Anupam Biswas, Ripon Patgiri, and Bhaskar Biswas. *Principles of Social Networking: The New Horizon and Emerging Challenges*. Springer, 2022.

[75] Dung V Pham, Giang L Nguyen, Tu N Nguyen, Canh V Pham, and Anh V Nguyen. Multi-topic misinformation blocking with budget constraint on online social networks. *IEEE Access*, 8:78879–78889, 2020.

[76] Shoujin Wang, Xiaofei Xu, Xiuzhen Zhang, Yan Wang, and Wenzhuo Song. Veracity-aware and event-driven personalized news recommendation for fake news mitigation. In *Proceedings of the ACM Web Conference 2022*, pages 3673–3684, 2022.

[77] Ryota Kobayashi and Renaud Lambiotte. Tideh: Time-dependent hawkes process for predicting retweet dynamics. In *Tenth International AAAI Conference on Web and Social Media*, 2016.

[78] Melissa A Schilling and Corey C Phelps. Interfirm collaboration networks: The impact of large-scale network structure on firm innovation. *Management science*, 53(7):1113–1126, 2007.

[79] Justin Cheng, Lada Adamic, P Alex Dow, Jon Michael Kleinberg, and Jure Leskovec. Can cascades be predicted? In *Proceedings of the 23rd international conference on World wide web*, pages 925–936, 2014.

[80] KZ Nanjo, B Enescu, Robert Shcherbakov, Donald L Turcotte, Takaki Iwata, and Yosihiko Ogata. Decay of aftershock activity for japanese earthquakes. *Journal of Geophysical Research: Solid Earth*, 112(B8), 2007.

[81] Jonathan M Read, Ken TD Eames, and W John Edmunds. Dynamic social networks and the implications for the spread of infectious disease. *Journal of The Royal Society Interface*, 5(26):1001–1007, 2008.

[82] Huacheng Li, Chunhe Xia, Tianbo Wang, Sheng Wen, Chao Chen, and Yang Xiang. Capturing dynamics of information diffusion in sns: A survey of methodology and techniques. *ACM Computing Surveys (CSUR)*, 55(1):1–51, 2021.

[83] Susan C Herring and Sanja Kapidzic. Teens, gender, and self-presentation in social media. *International encyclopedia of social and behavioral sciences*, 2:1–16, 2015.

[84] Adrien Guille, Hakim Hacid, Cecile Favre, and Djamel A Zighed. Information diffusion in online social networks: A survey. *ACM Sigmod Record*, 42(2):17–28, 2013.

[85] M Trupthi, Suresh Pabboju, and G Narasimha. Sentiment analysis on twitter using streaming api. In *2017 IEEE 7th International Advance Computing Conference (IACC)*, pages 915–919. IEEE, 2017.

[86] Hao Peng, Azadeh Nematzadeh, Daniel M Romero, and Emilio Ferrara. Network modularity controls the speed of information diffusion. *Physical Review E*, 102(5):052316, 2020.

[87] Xiaofeng Gao, Zhenhao Cao, Sha Li, Bin Yao, Guihai Chen, and Shaojie Tang. Taxonomy and evaluation for microblog popularity prediction. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 13(2):1–40, 2019.

[88] Dong Li, Yongchao Zhang, Zhiming Xu, Dianhui Chu, and Sheng Li. Exploiting information diffusion feature for link prediction in sina weibo. *Scientific reports*, 6(1):1–8, 2016.

[89] Bongwon Suh, Lichan Hong, Peter Pirolli, and Ed H Chi. Want to be retweeted? large scale analytics on factors impacting retweet in twitter network. In *2010 IEEE second international conference on social computing*, pages 177–184. IEEE, 2010.

[90] Nasir Naveed, Thomas Gottron, Jérôme Kunegis, and Arifah Che Alhadi. Bad news travel fast: A content-based analysis of interestingness on twitter. In *Proceedings of the 3rd international web science conference*, pages 1–7, 2011.

[91] Emilio Ferrara and Zeyao Yang. Quantifying the effect of sentiment on information diffusion in social media. *PeerJ Computer Science*, 1:e26, 2015.

[92] René Pfitzner, Antonios Garas, and Frank Schweitzer. Emotional divergence influences information spreading in twitter. In *Sixth international AAAI conference on weblogs and social media*, 2012.

[93] Liangjie Hong, Ovidiu Dan, and Brian D Davison. Predicting popular messages in twitter. In *Proceedings of the 20th international conference companion on World wide web*, pages 57–58, 2011.

[94] Zongyang Ma, Aixin Sun, and Gao Cong. Will this# hashtag be popular tomorrow? In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*, pages 1173–1174, 2012.

[95] Jiang Yang and Scott Counts. Predicting the speed, scale, and range of information diffusion in twitter. In *fourth international AAAI conference on weblogs and social media*, 2010.

[96] Cédric Lagnier, Ludovic Denoyer, Eric Gaussier, and Patrick Gallinari. Predicting information diffusion in social networks using content and user's profiles. In *European conference on information retrieval*, pages 74–85. Springer, 2013.

[97] Flavio Figueiredo, Jussara M Almeida, Marcos A Gonçalves, and Fabricio Benevenuto. Trendlearner: Early prediction of popularity trends of user generated content. *Information Sciences*, 349:172–187, 2016.

[98] Liangda Li and Hongyuan Zha. Learning parametric models for social infectivity in multi-dimensional hawkes processes. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 28, 2014.

[99] Cheng Yang, Maosong Sun, Haoran Liu, Shiyi Han, Zhiyuan Liu, and Huanbo Luan. Neural diffusion model for microscopic cascade study. *IEEE Transactions on Knowledge and Data Engineering*, 33(3):1128–1139, 2019.

[100] Janani Kalyanam, Mauricio Quezada, Barbara Poblete, and Gert Lanckriet. Prediction and characterization of high-activity events in social media triggered by real-world news. *PloS one*, 11(12):e0166694, 2016.

[101] Dong Li, Zhiming Xu, Yishu Luo, Sheng Li, Anika Gupta, Katia Sycara, Shengmei Luo, Lei Hu, and Hong Chen. Modeling information diffusion over social networks for temporal dynamic prediction. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management*, pages 1477–1480, 2013.

[102] Manuel Gomez-Rodriguez, Jure Leskovec, and Andreas Krause. Inferring networks of diffusion and influence. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 5(4):1–37, 2012.

[103] Minkyoung Kim, Dean Paini, and Raja Jurdak. Real-world diffusion dynamics based on point process approaches: A review. *Artificial Intelligence Review*, 53(1):321–350, 2020.

[104] RA Blythe and AJ Bray. Survival probability of a diffusing particle in the presence of poisson-distributed mobile traps. *Physical Review E*, 67(4):041101, 2003.

[105] Joseph Davies, Huy Truong-Ba, Michael E Cholette, and Geoffrey Will. Optimal inspections and maintenance planning for anti-corrosion coating failure on ships using non-homogeneous poisson processes. *Ocean Engineering*, 238:109695, 2021.

[106] Alan G Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971.

[107] Marcello Rambaldi, Emmanuel Bacry, and Fabrizio Lillo. The role of volume in order book dynamics: a multivariate hawkes process analysis. *Quantitative Finance*, 17(7):999–1020, 2017.

[108] Kyungsub Lee and Byoung Ki Seo. Marked hawkes process modeling of price dynamics and volatility estimation. *Journal of Empirical Finance*, 40:174–200, 2017.

[109] Fei Xu, Yumin Shi, Zhihua Feng, and Ming Li. Self-excited vibration in production, economy and society. *Procedia Manufacturing*, 39:1709–1714, 2019.

[110] Mateusz Bocian, John HG Macdonald, and Jeremy F Burn. Biomechanically inspired modeling of pedestrian-induced vertical self-excited forces. *Journal of Bridge Engineering*, 18(12):1336–1346, 2013.

[111] Shuai Gao, Jun Ma, and Zhumin Chen. Modeling and predicting retweeting dynamics on microblogging platforms. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, pages 107–116, 2015.

[112] Jonathan L Zelner, Benjamin A Lopman, Aron J Hall, Sebastien Ballesteros, and Bryan T Grenfell. Linking time-varying symptomatology and intensity of infectiousness to patterns of norovirus transmission. *PloS one*, 8(7):e68413, 2013.

[113] Minkyoung Kim, Dean Paini, and Raja Jurdak. Causal inference in disease spread across a heterogeneous social system. *arXiv preprint arXiv:1801.08133*, 2018.

[114] Pietro Della Briotta Parolo, Raj Kumar Pan, Rumi Ghosh, Bernardo A Huberman, Kimmo Kaski, and Santo Fortunato. Attention decay in science. *arXiv e-prints*, pages arXiv–1503, 2015.

[115] Minkyoung Kim, Daniel A McFarland, and Jure Leskovec. Modeling affinity based popularity dynamics. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pages 477–486, 2017.

[116] Aleksandr Simma. *Modeling events in time using cascades of Poisson processes*. University of California, Berkeley, 2010.

[117] Yosihiko Ogata. Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical association*, 83(401):9–27, 1988.

[118] Kamilya Smagulova and Alex Pappachen James. A survey on lstm memristive neural network architectures and applications. *The European Physical Journal Special Topics*, 228(10):2313–2324, 2019.

[119] Yosihiko Ogata. On lewis' simulation method for point processes. *IEEE transactions on information theory*, 27(1):23–31, 1981.

[120] Michele Starnini, Andrea Baronchelli, and Romualdo Pastor-Satorras. Temporal correlations in social multiplex networks. *arXiv preprint arXiv:1606.06626*, 2016.

[121] Peiyuan Suny, Jianxin Li, Yongyi Mao, Richong Zhang, and Lihong Wang. Inferring multiplex diffusion network via multivariate marked hawkes process. *arXiv preprint arXiv:1809.07688*, 2018.

[122] King-Sun Fu. Learning control systems–review and outlook. *IEEE transactions on Automatic Control*, 15(2):210–221, 1970.

[123] Amanpreet Singh, Narina Thakur, and Aakanksha Sharma. A review of supervised machine learning algorithms. In *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, pages 1310–1315. Ieee, 2016.

[124] Nagdev Amruthnath and Tarun Gupta. A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance. In *2018 5th international conference on industrial engineering and applications (ICIEA)*, pages 355–361. IEEE, 2018.

[125] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[126] Alireza Rezvanian, Behnaz Moradabadi, Mina Ghavipour, Mohammad Mehdi Daliri Khomami, and Mohammad Reza Meybodi. *Learning automata approach for social networks*, volume 820. Springer, 2019.

[127] Pengfei Wang, Yu Fan, Long Xia, Wayne Xin Zhao, ShaoZhang Niu, and Jimmy Huang. Kerl: A knowledge-guided reinforcement learning model for sequential recommendation. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval*, pages 209–218, 2020.

[128] KS Narendra and MALT Thathachar. Learning automata: An introduction prentice-hall. *New Jersey*, 1989.

[129] Beakcheol Jang, Myeonghwi Kim, Gaspard Harerimana, and Jong Wook Kim. Q-learning algorithms: A comprehensive classification and applications. *IEEE access*, 7:133653–133667, 2019.

[130] Georgios I Papadimitriou, Andreas S Pomportsis, S Kiritsi, and E Talahoupi. Absorbing stochastic estimator learning automata for s-model stationary environments. *Information Sciences*, 147(1-4):193–199, 2002.

[131] Ying Guo, Chong Di, and Shenghong Li. A novel reduced parameter s-model of estimator learning automata in the switching non-stationary environment. *Neural Computing and Applications*, 34(9):6811–6824, 2022.

[132] Kumpati S. Narendra and M. A. L. Thathachar. Learning automata - a survey. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-4(4):323–334, 1974.

[133] ML Tsetlin. On the behavior of finite automata in random media," automat. i. *Remote Control*, pages 1210–1219, 1961.

[134] Mina Ghavipour and Mohammad Reza Meybodi. A streaming sampling algorithm for social activity networks using fixed structure learning automata. *Applied Intelligence*, 48(4):1054–1081, 2018.

[135] S Roohollahi, A Khatibi Bardsiri, and F Keynia. Using an evaluator fixed structure learning automata in sampling of social networks. *Journal of AI and Data Mining*, 8(1):127–148, 2020.

[136] Geir Horn and B John Oommen. A fixed-structure learning automaton solution to the stochastic static mapping problem. In *19th IEEE International Parallel and Distributed Processing Symposium*, pages 8–pp. IEEE, 2005.

[137] Chuang Liu and Zi-Ke Zhang. Information spreading on dynamic social networks. *Communications in Nonlinear Science and Numerical Simulation*, 19(4):896–904, 2014.

[138] B. John Oommen and Daniel C. Y. Ma. Deterministic learning automata solutions to the equipartitioning problem. *IEEE Transactions on Computers*, 37(1):2–13, 1988.

[139] Ahsan Saleem, Muhammad Khalil Afzal, Muhammad Ateeq, Sung Won Kim, and Yousaf Bin Zikria. Intelligent learning automata-based objective function in rpl for iot. *Sustainable Cities and Society*, 59:102234, 2020.

[140] Kaddour Najim and Alexander S Poznyak. *Learning automata: theory and applications*. Elsevier, 2014.

[141] M Frank Norman. Markovian learning processes. *SIAM Review*, 16(2):143–162, 1974.

[142] S Lakshmivarahan and THATHACHAR MAL. Bounds on the convergence probabilities of learning automata. 1976.

[143] Mandayam AL Thathachar and P Shanti Sastry. Varieties of learning automata: an overview. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 32(6):711–722, 2002.

[144] B John Oommen and Mariana Agache. Continuous and discretized pursuit learning schemes: Various algorithms and their comparison. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 31(3):277–287, 2001.

[145] B John Oommen and Joseph K Lanctôt. Discretized pursuit learning automata. *IEEE Transactions on systems, man, and cybernetics*, 20(4):931–938, 1990.

[146] MAL Thathachar and P Shanti Sastry. A new approach to the design of reinforcement schemes for learning automata. *IEEE transactions on systems, man, and cybernetics*, (1):168–175, 1985.

[147] Mandayam AL Thathachar. A class of rapidly converging algorithms for learning automata. In *IEEE International Conference on Cybernetics and Society*, pages 602–606, 1984.

[148] Karl Pearson. The problem of the random walk. *Nature*, 72(1865):294–294, 1905.

[149] Sabir Umarov and Rudolf Gorenflo. On multi-dimensional random walk models approximating symmetric space-fractional diffusion processes. *Fractional Calculus and Applied Analysis*, 8(1):73–88, 2005.

[150] Ya Li, Yichao He, Xuejing Liu, Xiaohu Guo, and Zewen Li. A novel discrete whale optimization algorithm for solving knapsack problems. *Applied Intelligence*, 50(10):3350–3366, 2020.

[151] Jianjun Liu, Changzhi Wu, Jiang Cao, Xiangyu Wang, and Kok Lay Teo. A binary differential search algorithm for the 0–1 multidimensional knapsack problem. *Applied Mathematical Modelling*, 40(23-24):9788–9805, 2016.

[152] Mohammad Mehdi Daliri Khomami, Alireza Rezvanian, and Mohammad Reza Meybodi. Distributed learning automata-based algorithm for community detection in complex networks. *International Journal of Modern Physics B*, 30(8):1650042, 2016.

[153] Behnaz Moradabadi and Mohammad Reza Meybodi. A novel time series link prediction method: Learning automata approach. *Physica A: Statistical Mechanics and its Applications*, 482:422–432, 2017.

[154] Mina Ghavipour and Mohammad Reza Meybodi. A dynamic algorithm for stochastic trust propagation in online social networks: Learning automata approach. *Computer Communications*, 123:11–23, 2018.

[155] Hao Ge, Jinchao Huang, Chong Di, Jianhua Li, and Shenghong Li. Learning automata based approach for influence maximization problem on social networks. In *2017 IEEE Second International Conference on Data Science in Cyberspace (DSC)*, pages 108–117. IEEE, 2017.

[156] Hao Ge, Wen Jiang, Shenghong Li, Jianhua Li, Yifan Wang, and Yuchun Jing. A novel estimator based learning automata algorithm. *Applied Intelligence*, 42(2):262–275, 2015.

[157] David E Allen and Michael McAleer. Trump's covid-19 tweets and dr. fauci's emails. *Scientometrics*, 127(3):1643–1655, 2022.

[158] Chuang Liu, Xiu-Xiu Zhan, Zi-Ke Zhang, Gui-Quan Sun, and Pak Ming Hui. How events determine spreading patterns: information transmission via internal and external influences on social networks. *New Journal of Physics*, 17(11):113045, 2015.

[159] H.C. Kelman. Compliance, identification, and internalization: Three processes of attitude change. *Journal of Conflict Resolution*, 2(1):51–60, 1958.

[160] Angelia Nedic, Asuman Ozdaglar, and Pablo A Parrilo. Constrained consensus and optimization in multi-agent networks. *IEEE Transactions on Automatic Control*, 55(4):922–938, 2010.

[161] Luca F Bertuccelli and Jonathan P How. Estimation of non-stationary markov chain transition models. In *2008 47th IEEE Conference on Decision and Control*, pages 55–60. IEEE, 2008.

[162] Xiaotian Hao, Zhaoqing Peng, Yi Ma, Guan Wang, Junqi Jin, Jianye Hao, Shan Chen, Rongquan Bai, Mingzhou Xie, Miao Xu, et al. Dynamic knapsack optimization towards efficient multi-channel sequential advertising. In *International Conference on Machine Learning*, pages 4060–4070. PMLR, 2020.

[163] Ahmed Abouzeid, Ole-Christoffer Granmo, Christian Webersik, and Morten Goodwin. Learning automata-based misinformation mitigation via hawkes processes. *Information Systems Frontiers*, pages 1–20, 2021.

[164] Saurabh Mishra and Bilal M Ayyub. Shannon entropy for quantifying uncertainty and risk in economic disparity. *Risk Analysis*, 39(10):2160–2181, 2019.

[165] Parvaneh Sarshar, Ole-Christoffer Granmo, Jaziar Radianti, and Jose J Gonzalez. A bayesian network model for evacuation time analysis during a ship fire. In *2013 IEEE Symposium on Computational Intelligence in Dynamic and Uncertain Environments (CIDUE)*, pages 100–107. IEEE, 2013.

[166] Cigdem Aslay, Antonis Matakos, Esther Galbrun, and Aristides Gionis. Maximizing the diversity of exposure in a social network. In *2018 IEEE international conference on data mining (ICDM)*, pages 863–868. IEEE, 2018.

[167] Peter Kuppens, Nicholas B Allen, and Lisa B Sheeber. Emotional inertia and psychological maladjustment. *Psychological science*, 21(7):984–991, 2010.

[168] Runnan Li, Zhiyong Wu, Jia Jia, Jingbei Li, Wei Chen, and Helen Meng. Inferring user emotive state changes in realistic human-computer conversational dialogs. In *Proceedings of the 26th ACM international conference on Multimedia*, pages 136–144, 2018.

[169] Jonathan H Turner and Jan E Stets. Sociological theories of human emotions. *Annu. Rev. Sociol.*, 32:25–52, 2006.

[170] Che Liu, Junfeng Jiang, Chao Xiong, Yi Yang, and Jieping Ye. Towards building an intelligent chatbot for customer service: Learning to respond at the appropriate time. In *Proceedings of the 26th ACM SIGKDD international conference on Knowledge Discovery & Data Mining*, pages 3377–3385, 2020.

[171] Tae Kyun Kim. T test as a parametric statistic. *Korean journal of anesthesiology*, 68(6):540–546, 2015.

[172] Cristian Danescu-Niculescu-Mizil and Lillian Lee. Chameleons in imagined conversations: A new approach to understanding coordination of linguistic style in dialogs. *arXiv preprint arXiv:1106.3077*, 2011.

[173] Karthik Gopalakrishnan, Behnam Hedayatnia, Qinglang Chen, Anna Gottardi, Sanjeev Kwatra, Anu Venkatesh, Raefer Gabriel, Dilek Hakkani-Tür, and Amazon Alexa AI. Topical-chat: Towards knowledge-grounded open-domain conversations. In *INTERSPEECH*, pages 1891–1895, 2019.

[174] Rupsa Saha and Sander Jyhne. Interpretable text classification in legal contract documents using tsetlin machines. In *2022 International Symposium on the Tsetlin Machine (ISTM)*, pages 7–12. IEEE, 2022.

[175] Maryna Lukach, David Dufton, Jonathan Crosier, Joshua M Hampton, Lindsay Bennett, and Ryan R Neely III. Hydrometeor classification of quasi-vertical profiles of polarimetric radar measurements using a top-down iterative hierarchical clustering method. *Atmospheric Measurement Techniques*, 14(2):1075–1098, 2021.

Part II

Publications

Appendix A

Paper A

Title:	Causality-based Social Media Analysis for Normal Users Credibility Assessment in a Political Crisis
Authors:	Ahmed Abouzeid, Ole-Christoffer Granmo, Christian Webersik, Morten Goodwin
Affiliation:	University of Agder, Faculty of Engineering and Science, P. O. Box 509, NO-4898 Grimstad, Norway
Conference:	*2019 25th Conference of Open Innovations Association (FRUCT)*, Helsinki, Finland.
DOI:	10.23919/FRUCT48121.2019.8981500.

Causality-based Social Media Analysis for Normal Users Credibility Assessment in a Political Crisis

Ahmed Abouzeid, Ole-Christoffer Granmo, Christian Webersik, Morten Goodwin
Department of Information and Communication Technology
Faculty of Engineering and Science, University of Agder
P.O. Box 509, NO-4898 Grimstad, Norway
E-mails: {ahmed.abouzeid, ole.granmo, christian.webersik, morten.goodwin}@uia.no

Abstract **— Information trustworthiness assessment on political Social Media (SM) discussions is crucial to maintain the order of society, especially during emergent situations. The polarity nature of political topics and the echo chamber effect by SM platforms allow for a deceptive and a dividing environment. During a political crisis, a vast amount of information is being propagated on SM, that leads up to a high level of polarization and deception by the beneficial parties. The traditional approaches to tackling misinformation on SM usually lack a comprehensive problem definition due to its complication. This paper proposes a probabilistic graphical model as a theoretical view on the problem of normal users credibility on SM during a political crisis, where polarization and deception are keys properties. Such noisy signals dramatically influence any attempts for misinformation detection. Hence, we introduce a causal Bayesian Network (BN), inspired by the potential main entities that would be part of the process dynamics. Our methodology examines the problem solution in a causal manner which considers the task of misinformation detection as a question of cause and effect rather than just a classification task. Our causality-based approach provides a practical road map for some sub-problems in real-world scenarios such as individual polarization level, misinformation detection, and sensitivity analysis of the problem. Moreover, it facilitates intervention simulations which would unveil both positive and negative effects on the deception level over the network.**

A

A.1 Introduction

Nowadays, SM are an essential part of humans everyday life. The notable increase of the number of users, the ease and cheap cost of information sharing, and the consecutive technological enhancements of different SM platforms, have indeed boosted SM to be a tough competitor to traditional news outlets. The great benefit of SM does not lie only in news circulation, but further contexts of propagating it have been introduced. These modern paradigms of fast and cross-distance social interaction have allowed additional perspectives on how people are responding to the news. For example, emotions, questions, and disapproval from eyewitnesses have became part of the process.

Emergencies are not an exception of how people would depend on SM. During critical scenarios like political rebellions, terrorist attacks, and disasters caused by natural hazards, a significant amount of information is being propagated. In such circumstances, authorities and citizens construct various usage patterns of information through SM [1]. For instance, citizens would like to get updated by following authorities verified Facebook pages or Twitter accounts. Also, authorities might rely on information disseminated by citizens to feed up their emergency management systems in order to support decisions, since people could act as eyewitnesses. Moreover, citizens would interact together to enable more information diffusion, to express emotions, or to offer help. It has been observed from previous studies that there is a major challenge in all these patterns, which is the information manipulation and the lack of trust between citizens and authorities or between citizens themselves [2], [3], [4].

Information veracity assessment on SM is a critical topic because of how misleading news could affect the social order and the recovery from an emergency. Besides, the lack of trust between SM consumers threatens SM to serve as a reliable source of information, and waste all the technological efforts that have been previously accomplished. Moreover, how people are more likely to disseminate information regardless its correctness has its roots in psychological and social literature [5], [6]. Fortunately, with the growing number of consumers and how they depend on SM, users are playing a fundamental role in questioning and verifying received information, which can be viewed as a self-defense mechanism. Despite the contradiction of how both individuals and societies could positively or negatively shape information credibility, this self-defense mechanism of SM reveals the feasibility of overcoming such difficulty. Therefore, the study of information trustworthiness on SM has brought more focus in recent years.

Information accuracy dilemma on SM can be broken into multiple sub-problems: *rumor detection*, *cyborg/ trolls/ social bots detection*, and *fake news detection*. In rumor detection, a rumor could be either correct or incorrect. Commonly, a rumor is created during an emergency and due to the absence of a reinforced report from official entities. Some literature defines a rumor as a possibility to be either true or false [7]. Ref. [8] defined a rumor as an item of information which is deemed to be false. A potential cause of incorrect or inconsiderable information is SM fake

accounts. Hence, trolls, cyborg, and social bots detection have been studied in some literature in the preceding years [9], [10], [11]. In the context of SM, trolls are deceptive accounts ran by a human whom purpose is to motivate the others to an emotional reaction. On the other hand, cyborgs are a semi-automated accounts which objectively try to spread fake information. Social bots are usually ran by a computer program and used in many cases like advertising and fake news circulation. Fake news detection is the process of discovering false news. Either it was misinformation or disinformation. Misinformation refers to the unintentionally spreading of false information. On the contrary, disinformation is purposely circulating of fake news and usually adopted in political propaganda or in financial manipulations [12], [13]. In the rest of this paper, we will use the term Misinformation to refer to any fake news, regardless of the intentions.

Our theoretical study focuses on the circulation of misinformation in political emergencies like revolutions and uprisings, where corrupted regimes and citizens might confuse and mislead the public by disseminating deceptive content. One of the recently revealed methods on SM misinformation is the propagation-based method which considers that more information trustworthiness evidence to be retrieved from a majority of eyewitnesses or verified accounts [14]. In propagation-based methods, credibility networks are built to employ optimization techniques over different pieces of news giving the underlying point of view. Mining different viewpoints and reactions to news is referred to as Stance Detection [15], [16]. However, in a political crisis, everyone is biased with their opinions and reactions to other opinions or shared news [17]. Hence, such context is challenging the assumption that we can unveil information credibility by investigating different opinions and find out what the majority of people are believing in. On the other hand, it would be less complicated during disasters caused by natural hazards because people are less biased and the available opinions would be easily trusted.

A.1.1 Contribution and Paper Organization

This paper focuses on the problem of normal users content credibility assessment from the perspective of cause and effect as an interpretation of some evidence during the investigation. The paper studies a potential novel approach to the problem by engaging a theoretical foundation from Bayesian analysis and causal inference [18], [19]. The study challenges the assessment of different opinions trustworthiness about a specific claim. For that, we propose a probabilistic graphical model based on a causal BN to reason about possible causes and effects within the dynamics of information propagation on SM platforms. Our proposed method tries to solve the challenge of the unreliable opinion-based solutions in polarized scenarios by calculating a posterior marginal probability of the trustworthiness degree of opinions after obtaining some evidence. Our research contributions are summarized as follows.

- explain the capabilities of both predictive and diagnostic analysis on BNs to infer about the trustworthiness of an opinion and estimating polarization level

and other unknown information, given some evidence and observed causes and effects;

- illustrate how causal-based SM analysis opens the road to a potential novel approach for misinformation detection with the three layers of causal inference;

Although the different opinions and reactions to the news are taken into consideration along with the detected biased communities and deceptive accounts in a social network. However, it is important to highlight that the stance knowledge extraction task and community deceptive accounts detection techniques are not the focus of this research paper. The remainder of the paper is organized as follows. section A.2 gives a summary of the related work. section A.3 demonstrates the problem statement and notations. Our causality-based approach is explained in more details in section A.4. Our proposed methodology provided by a toy example is explained in section A.5. Finally, we conclude the whole paper and discuss our future work in section A.6.

A.2 Related Work

Research efforts on SM misinformation have studied the construction of prediction models for a misinformation classification task [2]. These models can be categorized into two classes: *content-based models*, and *context-based models*. Content-based models are divided into two main approaches: *knowledge-based*, and *style-based*. Knowledge-based methods propose examining external sources to fact-check the suspected information [20]. Various approaches could be applied for fact-checking as it could be automated or managed by human experts or crowd-sourcing. Computational fact-checking methods usually use either open web or a structured knowledge graph [21]. A knowledge graph is a structured network topology which could be constructed from the open web such as DBpedia and Google Relation Extraction Corpus. A fact-checking procedure is adopting a knowledge graph in order to infer about facts on its graph to verify information by exploring evidence from the external information source [21].

Style-based methods try to capture information manipulators by their writing style. Style-based methods can be categorized into two main classes: *deception-oriented*, and *objectivity-oriented* [2]. Earlier studies from forensic psychology investigated the credibility and manipulation of statements [22]. Such studies motivated the deceptive-oriented methods to detect misinformation. Deep neural network models, such as Convolutional Neural Networks (CNN), have been applied to classify deceptive contents according to their deceptive attitudes [23]. Objectivity-oriented refers to the manipulation of news by decreasing or hiding a key piece of information. Such scenarios are likely to happen in political emergencies and political manipulation campaigns. Linguistic-based features were used to detect objectively manipulated news articles [24].

Context-based models have two main approaches: *stance-based*, and *propagation-based*. Stance-based studies users reactions on the news. Some work proposed a

bipartite network of users and Facebook posts using the "like" stance [16]. This network was used for a semi-supervised probabilistic model to predict how likely a Facebook post is a hoax. Propagation-based approach focuses on people opinions on SM, it relies on the assumption that information credibility is highly related to the sincerity of SM opinions in relevant contents. Propagation-based approach attempts to infer if there is any conflict in the shared information by exploring other circulated details associated with a particular topic. Two types of propagation networks could be built: *homogeneous* and *heterogeneous* credibility networks [2]. Homogeneous credibility networks consist of a single kind of entities, such as posts or events [15]. Heterogeneous credibility networks connect different types of entities, such as posts, and sub-events [25]. These networks performs an optimization task on their graphs to conclude the veracity of the information.

It is acknowledged that the problem of misinformation in political situations cannot be solved by only applying any state-of-the-art technology in similar domains. For instance, stance detection and text-based solutions can just act as a first phase for a complicated pipeline. That is because in polarized political scenarios, the definition of fake news is relative, due to the different perspectives each sub-group would have. Therefore, what is really misinformation differs from the perceived false content. Recent efforts in studying the relationship between fake information and political polarization have revealed a correlation between polarization and what people consider as fake news on Twitter [26]. That claims an obstacle in the combat of misinformation detection on SM since concepts like biased opinions, actual fake information, relatively fake contents can be easily confused because of such correlation.

The polarization caused by both SM platforms and human nature threatens the reliability of opinion-based misinformation detection methods. In general, many social network community detection algorithms have been adopted [27]. Previous studies tried to tackle such problem by assuming that if we enforced more information diversity to each social bubble, it would reduce the polarization since the latter is an effect of the lack of information diversity itself [28]. Other previous work aimed to detect these communities and identify them as sub-networks or similar connected nodes in the social graph by analyzing the network cohesion [29]. One of the common real-world networks in community detection is Zachary's karate club which is a real example of a social network of 34 members (nodes) in a karate club and usually used as a benchmark dataset to evaluate community detection algorithms as well [30]. One of the recent contributions was an incremental method to detect communities in dynamic evolving social networks which was motivated by how previous community detection methods were static [31].

A similar concept to misinformation on SM is disease diagnosis and detection, both issues are putting people's lives on danger and they have symptoms and causes. Both also can spread among societies and their sub-communities. One of the most advanced techniques in modern medical diagnosis is the BNs [32]. BNs are probabilistic graphical models used to represent conditional relationships between random variables (graph nodes). These random variables could represent both evidence and

quires which we aim to reason and infer about. The relations in BNs can be modeled as causal relations which are more suitable for problems when the causes and effects are the core of the situation dynamics.

The problem of information veracity assessment on SM is intersected with many other tasks in the domain of SM analysis during disasters and other related contexts. Hence, it is important to highlight that in general, research community aims to extract knowledge from SM during crisis but there are differences between each sub-task. Knowledge extraction can be applied for sentiment analysis to explain the social behaviour of citizens during different stages of a crisis [33]. Opinion extraction used for news credibility tasks or political analysis [34]. Geo-location extraction tasks are being approached during disasters caused by natural hazards [33]. Hate speech towards certain groups of people which commonly increases during refugee crisis or extremely polarized political crisis [35].

A.3 Problem

A.3.1 Misinformation Definition

We aim to define the problem of normal users content credibility assessment on SM during a political crisis as a cause and effect problem instead of a classification task. The reason behind such definition is that an ordinary classification approach would not provide a complete control of such critical issue in our societies. On the other hand, an intervention view could unveil the root causes or suggest strategies to prevent misinformation. For that, we define misinformation propagation in terms of both predictive and diagnostic analysis tasks where causal inference approach is strongly followed. Misinformation could be viewed as a disease and the task is to understand when that disease occurs, and why it happens, and how to stop such issue in advance.

In the process of misinformation spreading, individuals approval to deceptive contents, and information shared by extremely polarized persons, could be considered symptoms of the deception phenomenon. Figure A.1 shows misinformation analysis causal-inspired solution framework. The framework declares how the stance detection, polarization measures, variety of social content a user is exposed to, deception information, and causal relations are considered as evidence to be collected in order to compute the trustworthiness of a user opinion. Moreover, an advanced causal analysis would be applied, such as intervention and sensitivity analysis to provide more confidence and insights about the inference or to hopefully suggest defensive strategies [19].

The definition of deception is critical to our proposed causal approach. We consider a political crisis as an environment where trolls, cyborgs, and deceptive social bots are trying to manipulate the public and motivate them to a specific reaction (stance). We differentiate between the collected deception information (deceptive accounts) and the unknown credibility of normal users. The latter is our focus in this study as we believe normal users are the threatening carriers of a deception

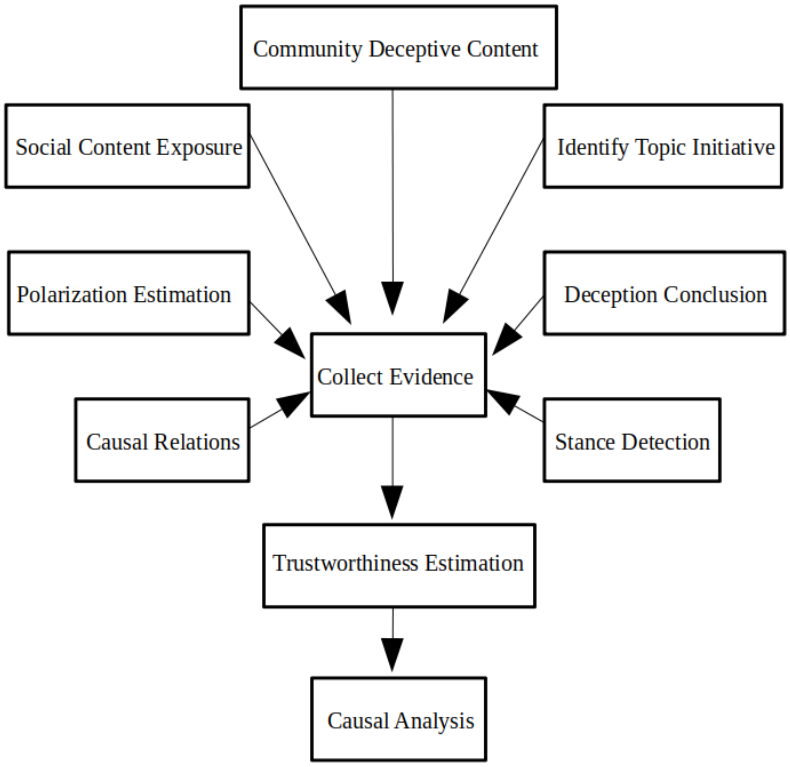


Figure A.1: Misinformation analysis causal-inspired solution framework

disease in political discussions. Hence, a strategy to only detect trolls, cyborgs, and deceptive social bots is not sufficient in our opinion.

As Figure A.1 indicates, we define a community deceptive content and a deception conclusion as two different things. Community deceptive content are all deceptive accounts in all biased communities over the network. For instance, right-wing trolls would manipulate the right-biased users to agree on a certain topic, on the other hand, left-wing cyborgs would defend that by propagating a refusal stance, however, both left/right-wings might share the same stance in some cases. We consider users as less trusted if they agreed on a common deceptive stance which was disseminated by all detected community deceptive content (left/right-wings). Although, in most cases, these community manipulating accounts would disagree with each others, therefore, a conclusion of the deception should be defined. Such conclusion means which opinion is considered less trustworthy and which could be dealt with as a defensive mechanism. To set a conclusion and draw the boundary lines between the differences in deceptive content stances, one more causal entity should be introduced, that is the topic initiative.

Topic initiative is defined as which biased party initially circulated the stance about the claim. For example, initially sharing something with an agreement or disagreement on it. The topic initiative would help to conclude the actual deceptive stance when left/right-wings share different opinions which is the most probable scenario. For instance, if the topic was started by a right-wing party, and right-wing users agreed to it including their community deceptive trolls, disagreements stances would be considered high trustworthy. If right-wing users disagreed on a

topic initiated by their biased sphere and circulated by their trolls while the latter agreed on the claim, disagreements stances credibility would even become higher, regardless of how left-trolls responded to it.

A.3.2 Social Media as an Environment

Given Twitter as an example, Figure A.2 illustrates the social engagement of main tweets X_i, X_j, X_k and their relationships (1=*agree*/ -1=*disagree*) with other reactions such as other main tweets, re-tweets, replies, and pressing a love button. Since SM have a lot of uncertainty and noise, we should differentiate between two scenarios. The first case is a certain environment where stances are certain guidelines for the misinformation detection task. The second scenario is the uncertainty about such engagements, since they might be biased instead of being subjective. Also they might be manipulated by other deceptive factors such as deceptive accounts.

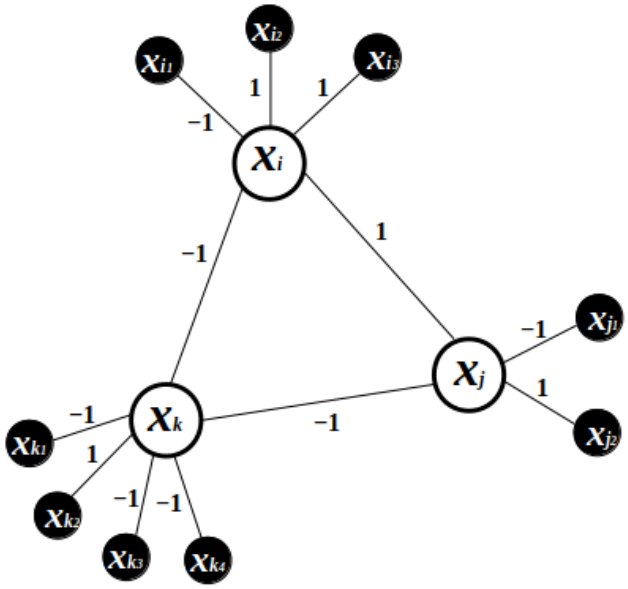


Figure A.2: Main tweets and their social engagement

Stance-based methods assume that the majority of opinions would be trusted. That means the more common opinion a single main tweet X is sharing, the more likely it is not false. However, it is crucial to define that majority since a polarized political discussion has an extreme deception possibility, even for its majority of opinions. In our proposed approach, we define a more likely credible user opinion according to its unbiased measures along with other main factors such as evidence indicating a less manipulation by deceptive content and a more variety of social content the user is exposed to. Hence, all biased and immature opinions should be more likely low trustworthy.

Figure A.3 indicates the complexity of the problem as a noisy transformation from certainty to uncertainty. The latter occurs because the stance detection task is a probabilistic solution to opinion mining problems. Moreover, detecting a social engagement E with a probability $Pr(E)$ close to unity from a stance detection

model could still be misleading, since it could be a biased opinion or influenced and manipulated by other false information or driven by psychological reasons.

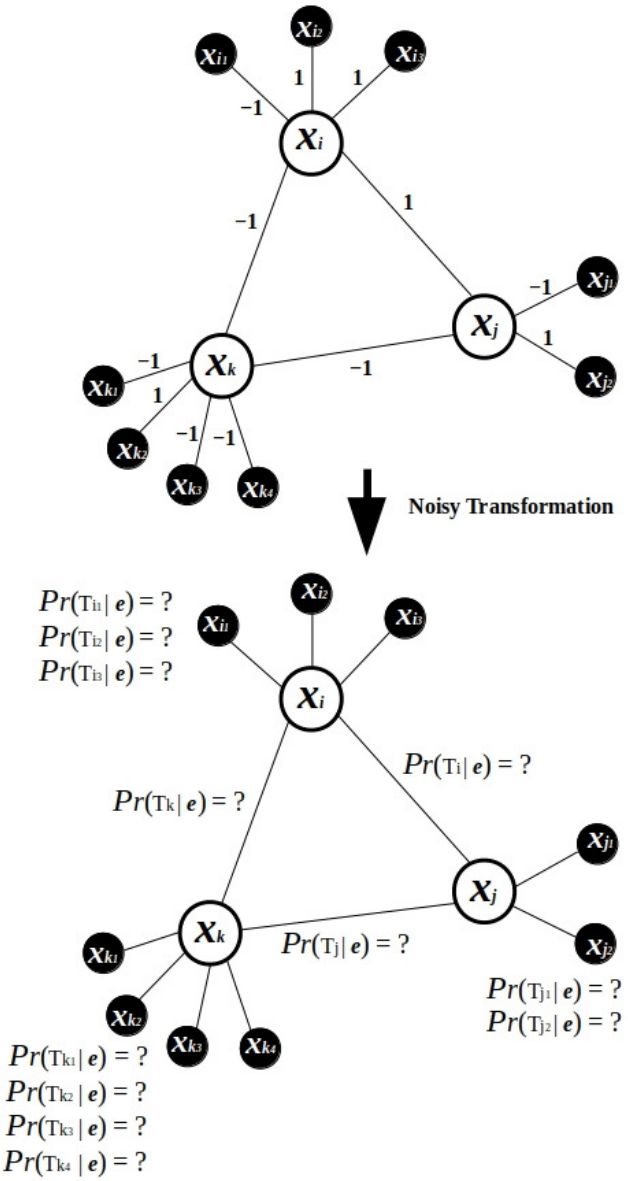


Figure A.3: Noisy transformation from certainty to uncertainty

Stance detection models only infer about the semantics of an opinion and do not consider how honest that opinion was. Hence, it is more convenient to define the uncertainty of social engagement trustworthiness as a conditional probability where the veracity of the stance is depending on other factors such as polarization and other causal relations in the social network. These factors could be referred to as evidence e. Therefore, we define the social engagement trustworthiness as $Pr(T|e)$. It is useful to formalize and analyze that transformation process using a causality and a probabilistic model to represent the uncertainty following the three layers of causality: *observing*, *intervention*, and *sensitivity analysis* [19].

A.3.3 Polarization Definition

The problem of misinformation detection is highly correlated with polarization in a political crisis. Therefore, a definition of polarization is critical to the problem analysis and solution. We define an honest opinion with regard to its root causes in the network. One of these causes is the polarized community an opinion is driven by. We consider polarized opinions in a political crisis to be categorized into the following main categories which represent the possible communities in the social network: *Far left*, *left*, *neutral*, *right*, and *far right*.

Our proposed categorization should differentiate between misleading opinions and biased ones, since a false stance is always misinformation but any of the five bias levels of opinions could be either misinformation or not. Moreover, we define a polarized opinion as a relative value where one SM stance can be considered less biased with regard to another content (agree with another community claim), while the same first stance could disagree with its own community claim. In our opinion, such relative definition per each case is useful for credibility assessment, for instance, if two tweets disagreed while they belong to the same community, such stance is very important since it indicates a certain level of subjectivity and a lower polarization level.

Some literature considered the less diversity of social content as a major cause of polarization, that means the more diversity of content a user is exposed to, the less polarized the user could be in most cases [28]. Furthermore, polarization is not only influencing normal users trustworthiness, it also dictates the objectively active deceptive accounts on the network. For instance, different community deceptive accounts would try to influence their communities such as right-wing and left-wing trolls, each would try to motivate its community in a certain direction with regard to a certain topic. Typically, these directions are opposite. Hence, agreeing with a deceptive content from the right-wing would mean disagreement with another from the left-wing. Therefore, there should be some measurements for which of these biased deceptive stances are less trustworthy and which ones are ironically higher in their trustworthiness.

A.3.4 Notations

Table A.1 describes the problem notations and their descriptions.

Definition of BN: Let $BN = (G, \theta)$ be the BN as the pair of Directed Asyclic Graph (DAG) G and θ as Conditional Probability Tables (CPTs) set. Let $Z = \{T, E, P, V, D, I, L, Y, B\}$ be the set of discrete random variables (nodes) of G, where the edges are causal relations over Z.

Definition of Trustworthiness Degree: Let's denote $T = [1, 10]$ as a discrete random variable where its value ranges between 1 and 10, indicating lower to higher degree of an opinion trustworthiness, respectively. Hence, the trustworthiness degree of the ith user stance E_i can be denoted as $Pr(T_i|e)$. Where e are all the occurred evidence calculated through (G, θ) when E_i had a certain value.

Table A.1: Problem notations and descriptions

Notations	Descriptions
(G, θ)	BN
G	DAG
θ	CPTs set
Z	Set of random variables (network structure nodes)
e	Some evidence over the network
$pa(Z)$	Parent node(s) of Z
$Y(Z)$	Child node(s) of Z
T	Trustworthiness
E	Social engagement (opinion)
P	Polarization level
V	Social content variety exposure
D	Concluded deception
L	Troll
Y	Cyborg
B	Deceptive bot
I	Topic initiative

Definition of Stance: Let's denote E as a discrete random variable for the network social engagements (stances) where $E = \{-1, 0, 1\}$ (*disagree*$=-1/$ *neutral*$=0/$ *agree*$=1$).

Definition of Polarization: Let $P = [1, 10]$ be the discrete random variable for the user polarization degree. P value ranges between 1 and 10, indicating lower to higher degree of polarization, respectively. $Pr(P_i|e)$ is the probability of the ith user polarization degree given evidence e.

Definition of Content Variety Exposure: Let $V = [1, 10]$ be the discrete random variable for the user social content variety exposure degree. V value ranges between 1 and 10, indicating lower to higher degree of content exposure, respectively. $Pr(V_i|e)$ is the probability of the ith user content exposure degree given evidence e.

Definition of Topic Initiative: Let I be the discrete random variable for the topic initiating polarized party, where $I = \{-2, -1, 0, 1, 2\}$ indicating *far left*, *left*, *neutral*, *right*, and *far right*, respectively.

Definition of Deception: Let's denote D as a discrete random variable for the concluded deceptive content opinion, where $D = \{-1, 1\}$ (*disagree*$=-1/$ *agree*$=1$). D can be observed as an evidence through its root causes, for instance, $Pr(D|D_L, D_R, I)$, where D_L, D_R, I are left/right-wings communities deceptive content stances, and the community which initiated the topic, respectively.

A.4 Causal Modelling

In this Section, we explain our causality-based approach to clarify the dynamics and relationships that shape the spreading of misinformation on SM during a political crisis. The conducted causality analysis in this paper should explicitly demonstrate our hypothetical assumptions on the problem as discussed in the previous sections. We aim to ask a question of which interventions are highly linked to information veracity rather than asking a prediction question only. Therefore, our main task is to model the cause and effect of the major variables on a social network that might influence or be affected by misinformation.

A causal graph is a visual representation of our assumptions about the problem and its data generating process. It should demonstrate the dynamics and relationships of the problem main entities (nodes) and the dependencies which are results of causal relations. In a causal graph, edges from parent nodes to child nodes mean a causal relationship. A child node variable is considered as an effect of its parent node variable. Figure A.4 shows a causal graph of a social network from information veracity perspective. There could be different hypothesized causal graphs for the same problem, hence, different probabilistic graphical models could be constructed as well. Evaluating different causal models is recommended in that case. In this theoretical paper, we provide one assumption of the problem and the given causal graph shows the details of this assumption.

A.4.1 Causal Graph

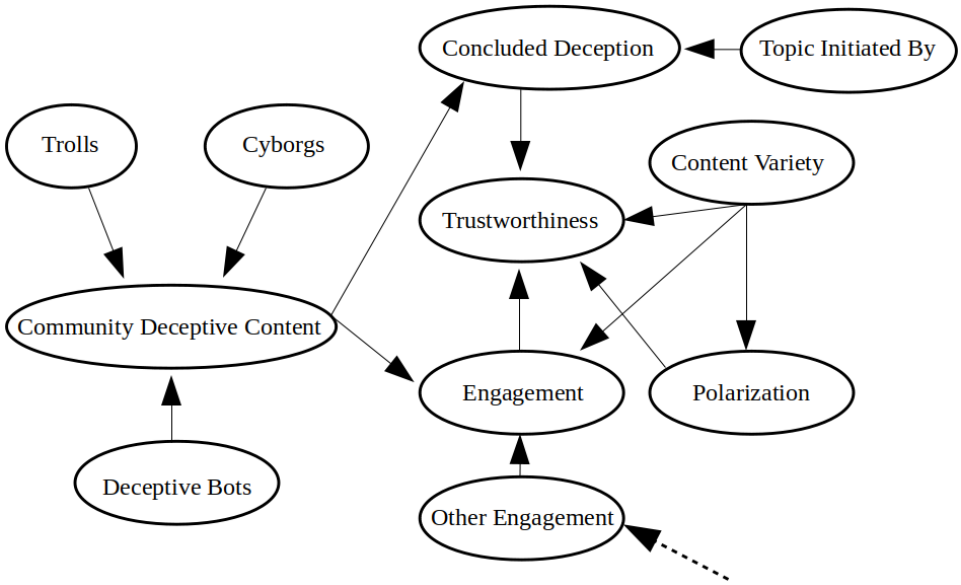


Figure A.4: Polarized political SM discussion causal graph

In Figure A.4, we consider a community deceptive content to be a common cause of users social engagement on that community. The assumption is that all users are distributed across different communities on SM, representing their mindset and preferences, each community will be influenced somehow by being exposed

to a deceptive content targeting that community. The idea of small communities dedicated deceptive content is crucial to the challenge of biased and manipulated opinions, since by investigating such causal relations, we would be able to weight different opinions according to their causes.

In general, we consider the three main potential deceptive accounts to be the cause of a community manipulation: *trolls*, *cyborgs*, and *deceptive social bots*. In a highly polarized discussion, people would be easily manipulated and would agree on what is deceptively influencing them in their social bubbles. Hence, normal users trustworthiness degrees of each reaction to deceptive information should be measured to evaluate their credibility. Moreover, the trustworthiness degree is affected by the measures of polarization levels and social content variety exposure, the latter is directly influencing both polarization levels and social engagements as well [28]. Eventually, we consider other social engagement that might influence one's engagement like when a user is replying to others and approving or denying their opinions.

As mentioned in the problem definition, a topic initiative and a concluded deception stance from different community deceptive content should be defined in order to collect more evidence about the stances trustworthiness degrees. In our causal graph, we consider the concluded deceptive stance as a result of measuring its hypothesized causes. These causes are the community which has initiated topic, the stance on the initiated topic, and stances from other community deceptive content.

A.4.2 Graph Semantics

There are three main structures a causal graph could have and each one describes a unique concept of how the joint probability distribution function will be factorized. These causal semantics guide the creation of the CPTs. These CPTs are crucial since they are the model parameters. Figure A.5 demonstrates the three different causality graph structures. In the chain structure, a cause Z is influencing an effect X, the latter will trigger another effect Y. That indicates how a directly connected child node is dependent to its parent node. Moreover, that pattern holds one important property and it is crucial to the computation, which is that Y is conditionally independent from Z given that the intermediate node X occurred. By given X, we can infer about Y even if we do not know anything about Z. That conditional independence is denoted as $(Y \perp Z|X)$. That means if X occurred, $Pr(Y|Z,X) = Pr(Y|X)$ and that simplifies the calculation. In the common cause structure, Y, Z, and K are also conditionally independent if X occurred. In such causal pattern, X is called a confounder of Y, Z, and K as it is considered a common cause and they are dependent on it.

As the opposite to the previous described casual structures, the collider path or the common effect structure is different when it comes to the definition of its conditional independence, so by given that X occurred, Y, Z, and K are conditionally dependent on each other which is denoted as $(Y \not\perp Z, K|X)$. A special case for the collider path is when X is a child node for another parent node $pa(X)$, if $pa(X)$

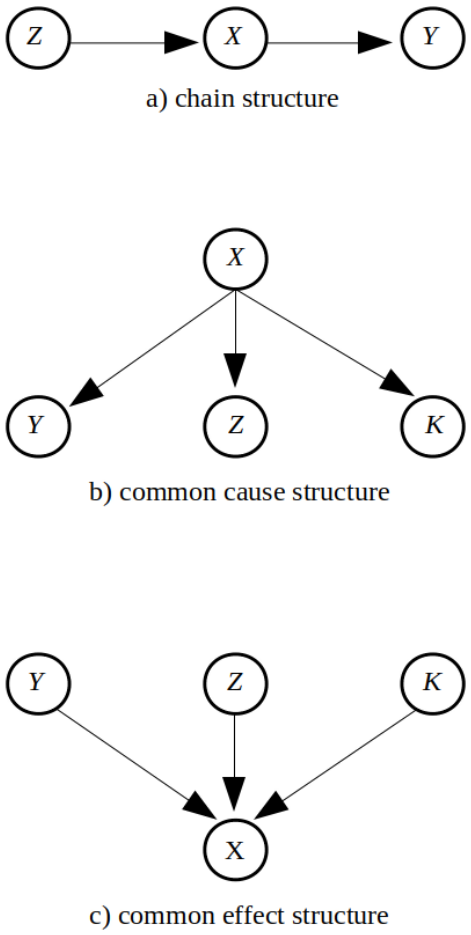


Figure A.5: Possible causal structures

occurred, then Y, Z, and K are also conditionally dependent to each other, even if we do not know about X.

A.5 Methodology

A.5.1 Causal Bayesian Networks

BNs are fundamental methods in the field of Artificial Intelligence (AI). They provide efficient ways to calculate large and complex probabilistic inference tasks under uncertainty [32], [36], [37]. The relations in the network (directed edges) can be causal relations and the network is constructed as a DAG where no loops inside any part of the graph can be found. The DAG property is also important for how the reasoning would be performed, since variables independence in DAG is compatible with how we can calculate the joint probability distribution of all the random variables. Figure A.6 shows an abstract BN, modeled according to our hypothesized causal graph (see Figure A.4) with the defined domain variables (see Table A.1).

Our proposed BN is a connected graph and its complexity is bounded by the number of stances E it will investigate. Abstractly, Figure A.6 has three social

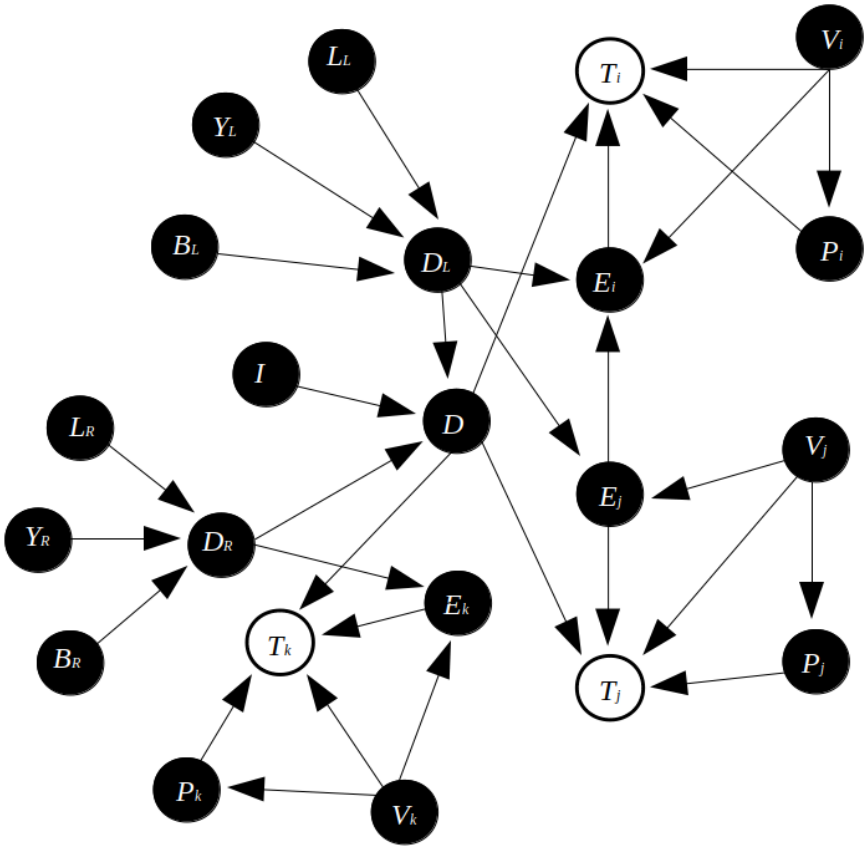


Figure A.6: The derived BN from the assumed causal graph

engagements where E_i and E_j belong to a left-wing community, and E_k is considered a part of a right-wing community. Although E_i and E_j are social engagements from the same social bubble, they could have different polarization levels P_i, P_j if they have different social content exposure measures V_i, V_j, for instance, if user i is more exposed to other community related content.

Each stance on the given BN represents a possible scenario of interaction such as the case when E_i is influenced by its community deceptive content D_L, its exposed social content variety estimation V_i, and another stance E_j that motivated user i to reply. On the other hand, E_k is only influenced by its community deceptive content D_R and its exposed content variety estimation variable V_k. In the real world scenarios, the number of nodes on such BN could be extremely bigger and the connectivity degree of the graph will be remarkably higher.

In BN, we can perform two possible types of reasoning: *predictive* and *diagnostic*, where each one dictates the direction of reasoning on the graph, either from the child node to a parent node (bottom-up) or the other way (top-down). The task of any BN is to calculate a marginal posterior probability of an unknown variable given some prior probabilities and likelihoods for other known variables. The process of calculating a marginal posterior probability is called belief update or probabilistic inference.

To build up a simulation model based on BN, we should first obtain some prior and conditional probabilities. Prior and conditional probabilities can be ob-

tained from observations and conditional frequencies on data samples. Equation A.1 demonstrates how the joint probability distribution of all discrete random variables on BN is calculated. Furthermore, the factors of the joint probability distribution function are interpreted as CPTs for child nodes and prior probabilities for root nodes. These probabilities are considered the network parameters for calculating the targeted unknown variable. Figure A.6, declares these unknown variables with a white circle, while other black circles are representing observed evidence (assignments of variables). For example, evidence that are collected by applying stance detection, polarization estimation, and exposed content variety estimation. Moreover, deceptive accounts detection tools should be applied to collect evidence about the concluded deceptive content stance D on the network. What remains after collecting these evidence, is to calculate the marginal posterior probability of the discreet random variable T which represents the trustworthiness degree of the user social engagement.

$$Pr(z_1, ..., z_n) = \prod_{i=1}^{n} pr(z_i|pa(z_i)). \tag{A.1}$$

A.5.2 Belief Update

The task of the BN belief update algorithm is to learn the posterior joint probability distribution along with the network topology. There are different update belief algorithms [38]. In this Section, we will give a brief statement on the EPIS-BN algorithm, which is an evidence pre-propagation importance sampling algorithm for BNs [39]. In general, importance sampling algorithms seem to be more successful with extremely unlikely evidence, which would be the case for SM remarkable randomness. It has been stated that exact inference in BNs is NP-hard [40]. Moreover, with thousands of variables in the network, it becomes infeasible to obtain an exact inference. Sometimes, the only way to obtain results is the approximate inference. Approximate inference is also NP-hard [41]. In general, the complexity of the computation increases if the number of parents increases for a child node, that is because the computational cost of the many entries and calculations in the CPTs.

Importance sampling-based algorithms are inherited from the family of stochastic sampling algorithms [38]. The former seem to provide a more robust performance, giving the research efforts to obtain a better importance function which is crucial to the precision of the inference. Theoretically, the convergence rate of the importance sampling-based algorithms is in the order of $\frac{1}{\sqrt{n}}$, where n is the number of samples.

In general, an update belief algorithm works by determining the number of samples and initializing the prior and CPTs of the network. According to our proposed BN in Figure A.6, for an unknown variable T_i (user i stance trustworthiness), to collect evidence e and update the beliefs for T_i, two subsets (e^+, e^-) should be defined. These subsets declare the ancestors and descendants of T_i, respectively. Then, the algorithm constructs two types of messages calculated and accumulated through e^+ and e^-: *parent to child messages* and *child to parent messages*, respectively. Fig-

ure A.7 indicates how these messages are being propagated when updating the belief of any targeted variable Z over the BN, where Z beliefs are updated through all its incoming messages. We have used the notations $pa(Z)$ and $Y(Z)$ to refer to parents and children of Z, respectively.

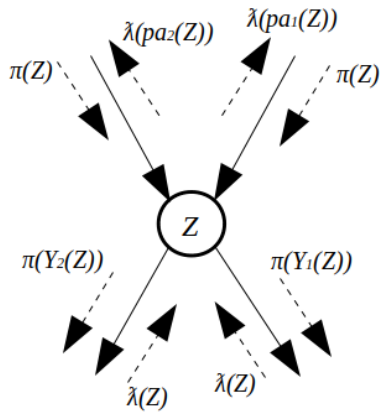


Figure A.7: Information propagation over BN

In a more compact form, Equation A.2 and Equation A.3 demonstrate how to calculate the incoming messages to the ith user trustworthiness degree variable T_i over the BN.

$$\pi(T_i) = \prod_{e^+} Pr(T_i|e^+_{T_i}), \tag{A.2}$$

$$\lambda(T_i) = \prod_{e^-} Pr(e^-_{T_i}|T_i), \tag{A.3}$$

where $\pi(T_i)$ and $\lambda(T_i)$ are representing messages sent to the stance trustworthiness variable T_i from its direct causes (E_i, P_i, V_i, D), and messages sent to T_i from its effects (no effects for T_i), respectively.

Obviously, our proposed causal BN declares that the trustworthiness node T_i has no effect on any descendants, hence, $\lambda(T_i) = 1$ in that case. On the other hand, $\pi(T_i)$ can be rewritten with Equation A.4.

$$\pi(T_i) = Pr(T_i|P_i, E_i, D) \cdot \pi(P_i) \cdot \pi(E_i) \cdot \pi(D) \cdot Pr(V_i). \tag{A.4}$$

As noticed, we did not explicitly include the content variety exposure V_i variable for θ_{T_i} since it will be calculated from the messages coming from the trustworthiness node to its ancestor content variety node $\lambda(V_i)$ for the ith user. The same dropping goes for any discrete random variable that would be duplicated in the equations. Also, we have added $Pr(V_i)$ instead of $\pi(V_i)$ since content variety has no parents to receive messages from. In general, and by using Equation A.2 and Equation A.3, the belief update algorithm calculates the beliefs of a variable according to Equation A.5.

$$\Pr(Z|e) = \alpha \cdot \pi(Z) \cdot \lambda(Z), \tag{A.5}$$

where $\alpha = \frac{1}{e}$ as the normalization constant and the multiplication of both π and λ is a pairwise multiplication since they both are considered as probability distribution vectors over their investigated variables possible values. The result of this equation should be also a marginal posterior probability distribution for Z over the evidence e.

The Importance Conditional Probability Tables (ICPTs) are the new introduced concept to the previous general demonstrated calculations in belief update. In ICPTs, a table of a node T_i is a posterior probabilities table where $Pr(T_i|pa(T_i), e)$. The probabilities are conditional on the evidence as well, instead of conditioning only on the ancestors of T_i.

A.5.3 Toy Example

Our hypothesis about the problem of normal users credibility on polarized SM discussions is slightly tested in this Section. Our toy example provides three scenarios to evaluate the proposed causal structure. We have used the EPIS-BN algorithm from GeNIe software academic version to simulate these scenarios [42]. First, we evaluate how the algorithm will perform when not all evidence are observed and the trustworthiness variable T is unknown. Second, we test the performance further by making a fully observed evidence. Third, we try to mislead the network in the second scenario by intervene and change some values for some evidence to check if there would be any contradiction in the results.

As discussed in Figure A.4 and Figure A.6, the community deceptive content is caused by objectively deceptive accounts like trolls, cyborgs, and deceptive social bots. For simplicity reasons, we have omitted the variables for these three causes and instead, we will consider only the community deceptive content variable, regardless of its causes. The main setup in the three scenarios is as follows:

- two biased communities (left-wing, right-wing) and five users are part of a political discussion: **Bob**, **Alice**, **Charlotte**, **Daisy**, and **Eric**;
- **Bob**, **Alice**, and **Eric** are part of the right-wing community, on the other hand, **Charlotte** and **Daisy** are considered members of the left-wing society;
- the community deceptive content of the left side is disagreeing on a claim, while the right side deceptive content is agreeing on it. Moreover, the topic is initiated by the right-wing community;
- the social engagement of **Charlotte** is also influenced by a social engagement from **Daisy**, and **Eric** social engagement is also influencing **Alice** opinion;

In order to initialize our proposed causal BN, CPTs should be constructed. Figure A.8, indicates an example of CPTs for the network. These values were defined as dummy data, nevertheless, they give a logical conditional frequency of how likely people would agree or disagree. In case of real data, the values could be constructed from conditional frequencies in the data itself, for example, given a time series data,

how many times a user tended to agree to its own community deceptive account when the user content variety exposure was low.

Bob Social Exposure	Low		Medium		High	
Deceptive Content	Agree	Disagree	Agree	Disagree	Agree	Disagree
Agree	0.67	0.14	0.47	0.42	0.01	0.98
Neutral	0.02	0.11	0.01	0.03	0.01	0.01
Disagree	0.31	0.75	0.52	0.55	0.98	0.01

Figure A.8: Bob social engagement from CPTs

Figure A.9, shows a simple community discussion over SM. Users: **Alice**, **Bob**, **Charlotte**, **Daisy**, and **Eric** were communicating with their different social background and experience. In this scenario, we have considered that not all evidence were observed and the task is to update the belief of the five users trustworthiness degree T, given the collected evidence for all causes of T except the stances E.

In Figure A.9 scenario, the discrete random variable I was indicating that the topic was initiated by the right-wing community (either normal users or deceptive accounts in that social bubble). Then, the right side deceptive accounts stances agreed on the claim of the topic, then, the left side disagreed. Since the topic was initiated by the right side and the right-wing deceptive accounts reacted with agreements, the BN updated the belief of the discrete random variable D and considered the agreement stance as the deceptive stance for the topic claim with probability $Pr(D = 1|e) = 83\%$. Since normal users stances E were not given as part of the evidence, the BN calculated their beliefs according to the incoming messages for all the corresponding nodes E_A, E_B, E_C, E_D, E_E for users **Alice**, **Bob**, **Charlotte**, **Daisy**, and **Eric**, respectively.

Notably, both **Charlotte** and **Daisy** were already a left-side community members and they both were highly polarized, hence, they both contradicted with the right-wing initiative and disagreed on it. Moreover, the more the user will disagree on the claim, the higher the trustworthiness degree will be. For instance, **Charlotte** would disagree with a belief $Pr(E_C = -1|e) = 91\%$. On the other hand, **Daisy** would disagree with a belief $Pr(E_C = -1|e) = 75\%$. Consequently, the beliefs for T_C and T_D were 61% and 55%, respectively. Furthermore, **Bob** has a higher belief of disagreement and trustworthiness $Pr(E_B = -1|e) = 98\%$, $Pr(T_B = 10|e) = 82\%$, even if he was a right-wing, that might be because of the evidence which indicated his less polarization and high exposure to diversity of content. In addition, it was noticed how **Alice** was considered less trusted since her stance belief was almost to agree and to share the same deceptive stance $Pr(E_A = 1|e) = 78\%$, $Pr(T_A = 0|e) = 74\%$.

In the second scenario, Figure A.10 explains what has happened when we replaced the beliefs probabilities of E_A, E_B, E_C, E_D, E_E with certain evidence to increase the probabilities in the first scenario to be certain values with a probability equal to unity. For instance, from $Pr(E_A = 1|e) = 78\%$ to just $E_A = 1$, and from $Pr(E_D = -1|e) = 75\%$ to just $E_D = -1$. Then, the updated beliefs of the trustworthiness degree of users became closer to 1. For instance, **Daisy** high trustworthiness degree belief changed from 55% to 65%, after giving more evidence and information. Same occurred to **Eric**, since his high trustworthiness degree in

the first scenario was 51%, giving that the beliefs of his stance were distributed as $Pr(E_E = 1|e) = 47\%, Pr(E_E = 0|e) = 1\%, Pr(E_E = 1|e) = 52\%$. However, in the second scenario with a given evidence of how he has reacted exactly, his high trustworthiness degree belief became 67%.

Figure A.11 indicates an intervention in the experiment as the third case scenario. Given the first two scenarios, it was always a majority stance which was considered as a high trustworthy. For example, first two cases considered the agreement as a less trustworthy social engagement while only one out of five users had such an opinion. Our third situation tried to evaluate the challenge of a biased majority of opinions that might mislead any stance or propagation-based misinformation detection solution. Hence, we intervened to make the agreement stance as the major opinion in the discussion with even more confusing evidence, for instance, we made **Daisy** agrees but also we made her a less polarized person. Nevertheless, the results in Figure A.11 shows how the disagreement stance still considered as a high trustworthy despite of being a minority.

A.6 Conclusion

In this paper, we have introduced a theoretical study for the problem of normal users credibility on social media in a political crisis. Our proposed methodology could be a novel solution to the problem of misinformation. We have modeled the problem of misinformation in social media as a cause and effect process, where causes and effects are evidence to be collected before calculating the marginal posterior probability of the trustworthiness degree of the user opinion about a claim. On the other hand, recent approaches on misinformation lack the definition of polarization and biased opinions along with a full adoption to the causality approach. For instance, how traditional misinformation stance and propagation-based methods would be less efficient in polarized situations. Hence, it is crucial to define the uncertainty that occurs in a polarized political discussion over social media. Such uncertainty could not be only the extreme biased opinions as anomalies in the data, therefore, it would be more efficient to define the cause and effect between all key variables including the polarization causes, effects, and the effects of the effects. Our proposed causal BN considered these key variables as the social engagement (stance), the polarization level, the amount of information and its variety a user is exposed to, and the deceptive content in the discussion along with the topic initiative. Our toy example provided three scenarios representing partial observation of these variables, full observation, and an intervention scenario to evaluate any contradiction in the proposed causal structure.

Along with updating the beliefs for the normal users stances trustworthiness degrees, the given study would be suggested to trace the deceptive accounts, predict stances, and estimate polarization levels. Eventually, that would lead to the computation of each normal user credibility by employing a Dynamic Bayesian Network (DBN) to infer the trustworthiness degrees of users stances over time as a tempo-

ral feature for the credibility assessment [36]. Furthermore, the proposed approach would be applied on other domains such as fake reviews on commercial products or disasters caused by natural hazards, by modeling the problem causal relations and variables within a causal BN.

In order to adopt with the complexity of the social network and the numerous number of nodes our final BN would reach, the study of how to design the system with a proper computational cost is necessary. In addition, further work should be applying some experiments based on artificial and real world data. Moreover, a complete sensitivity analysis and intervention simulation should be studied and applied on all demonstrated variables. Finally, the study of the DBN is important since the time dimension is critical to our problem, especially for measuring the temporal trustworthiness of normal users along with polarization and content exposure correlation.

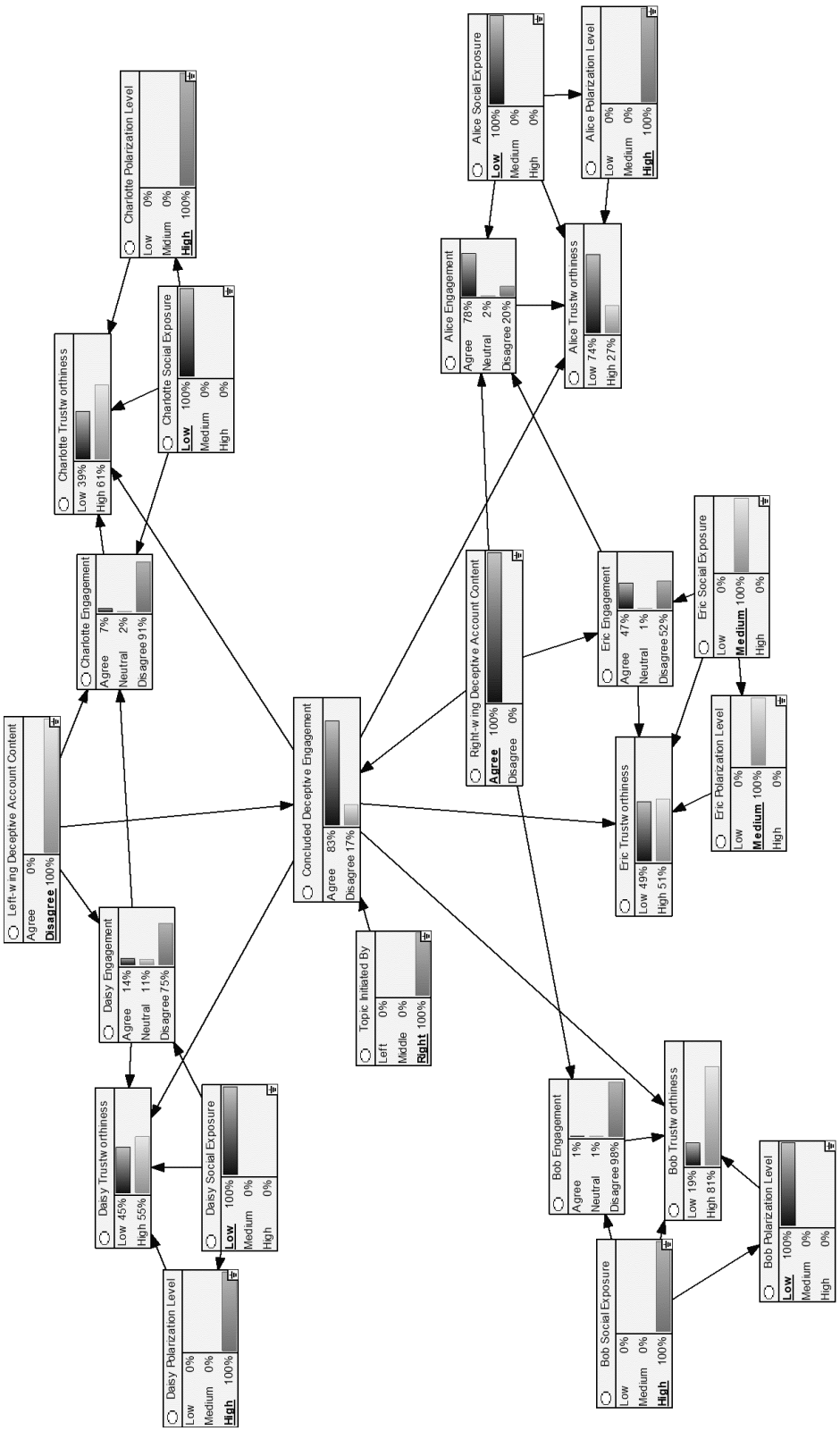


Figure A.9: Partially observed evidence scenario

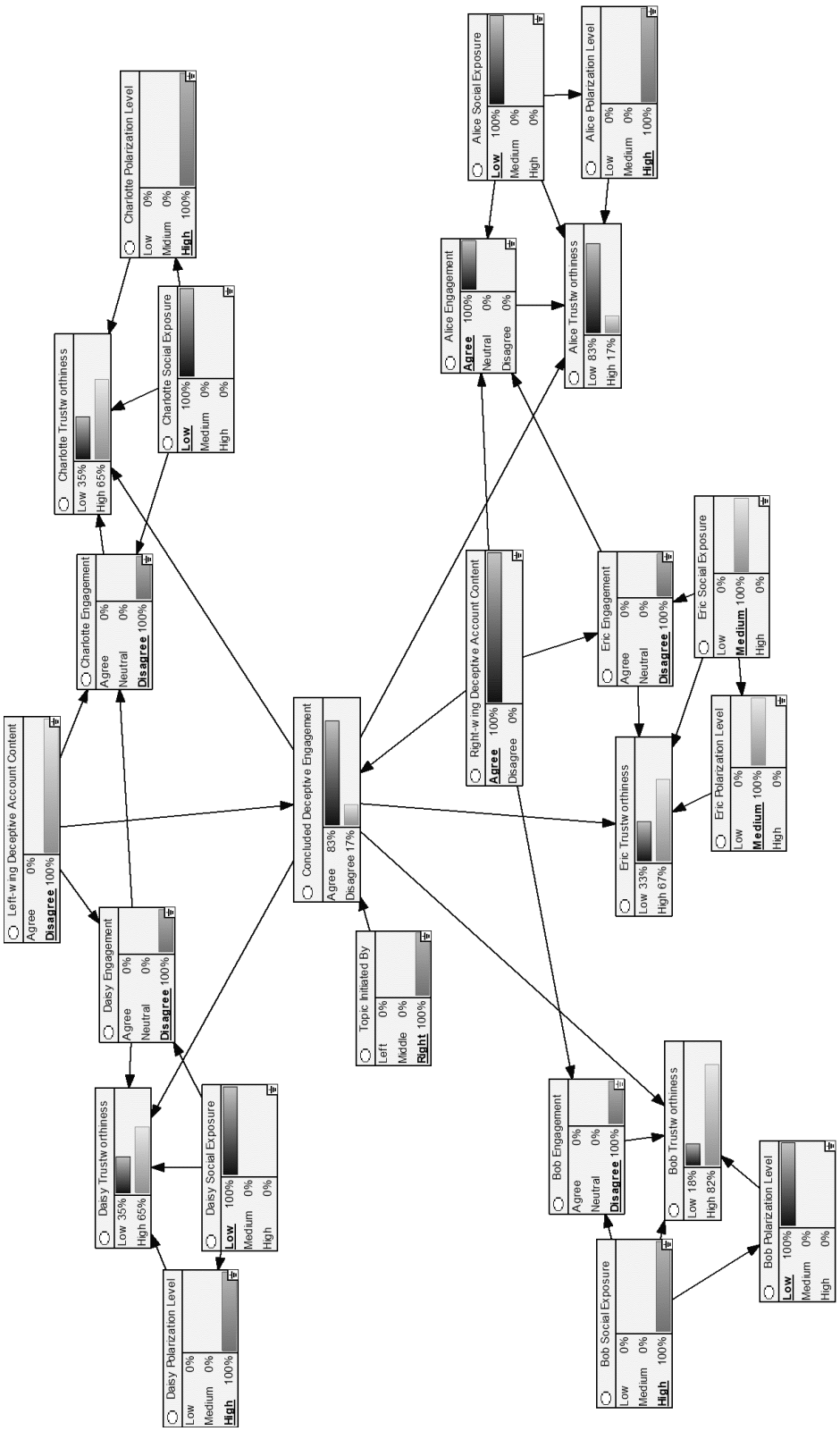


Figure A.10: Fully observed evidence scenario

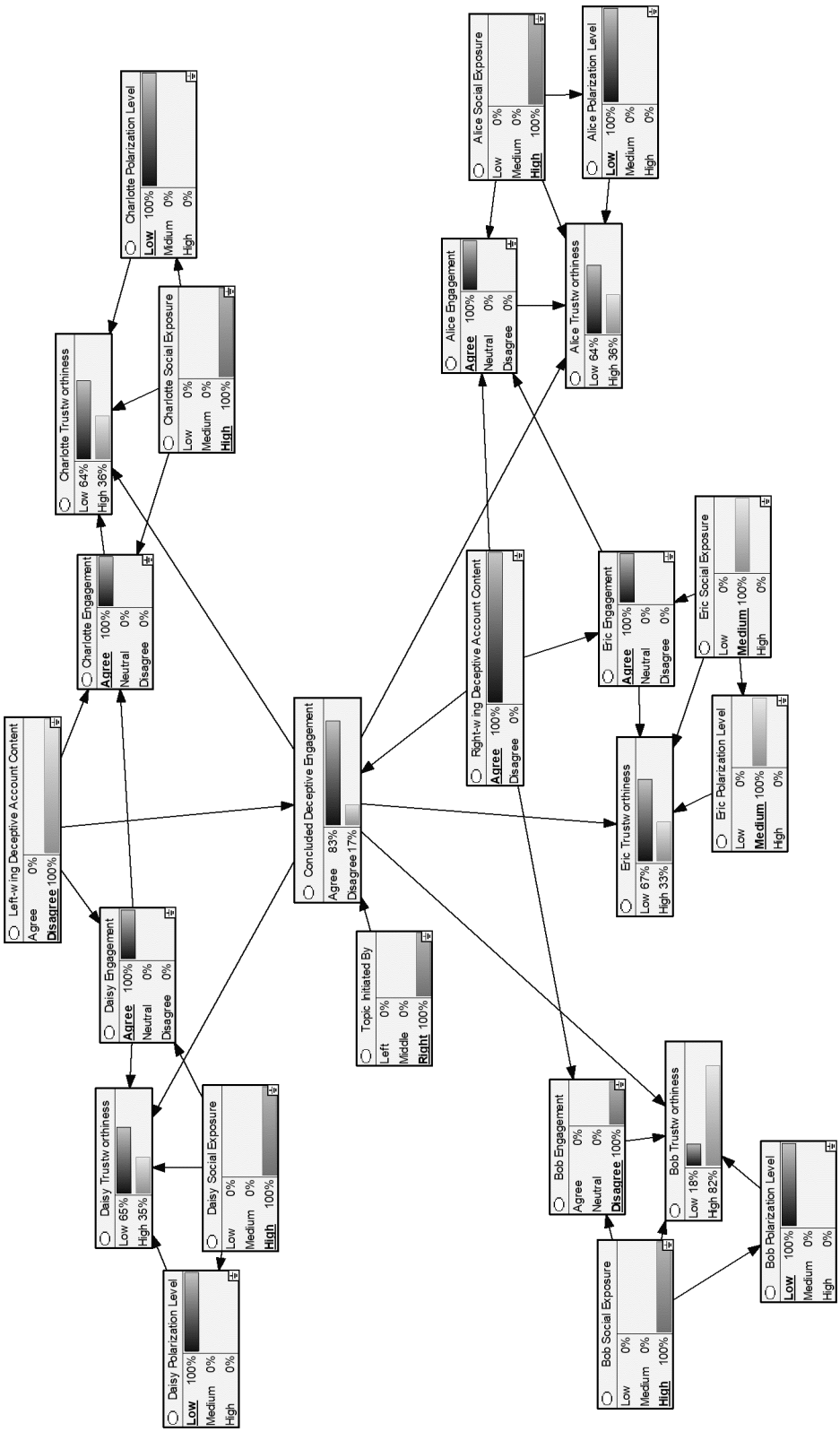


Figure A.11: Intervention scenario

Bibliography

[1] Christian Reuter and Marc-André Kaufhold. Fifteen years of social media in emergencies: a retrospective review and future directions for crisis informatics. *Journal of Contingencies and Crisis Management*, 26(1):41–57, 2018.

[2] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, 19(1):22–36, 2017.

[3] Arkaitz Zubiaga, Ahmet Aker, Kalina Bontcheva, Maria Liakata, and Rob Procter. Detection and resolution of rumours in social media: A survey. *ACM Computing Surveys (CSUR)*, 51(2):32, 2018.

[4] Aditi Gupta, Hemank Lamba, Ponnurangam Kumaraguru, and Anupam Joshi. Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy. In *Proceedings of the 22nd international conference on World Wide Web*, pages 729–736. ACM, 2013.

[5] Christopher Paul and Miriam Matthews. The russian "firehose of falsehood" propaganda model. *Rand Corporation*, pages 2–7, 2016.

[6] Walter Quattrociocchi, Antonio Scala, and Cass R Sunstein. Echo chambers on facebook. *Available at SSRN 2795110*, 2016.

[7] Nicholas DiFonzo and Prashant Bordia. Rumor, gossip and urban legends. *Diogenes*, 54(1):19–35, 2007.

[8] Gang Liang, Wenbo He, Chun Xu, Liangyin Chen, and Jinquan Zeng. Rumor identification in microblogging systems based on users' behavior. *IEEE Transactions on Computational Social Systems*, 2(3):99–108, 2015.

[9] Alessandro Bessi and Emilio Ferrara. Social bots distort the 2016 us presidential election online discussion. *First Monday*, 21(11-7), 2016.

[10] Zi Chu, Steven Gianvecchio, Haining Wang, and Sushil Jajodia. Detecting automation of twitter accounts: Are you a human, bot, or cyborg? *IEEE Transactions on Dependable and Secure Computing*, 9(6):811–824, 2012.

[11] Justin Cheng, Michael Bernstein, Cristian Danescu-Niculescu-Mizil, and Jure Leskovec. Anyone can become a troll: Causes of trolling behavior in online

discussions. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, pages 1217–1230. ACM, 2017.

[12] Gillian Bolsover and Philip Howard. Chinese computational propaganda: automation, algorithms and the manipulation of information about chinese politics on twitter and weibo. *Information, Communication & Society*, pages 1–18, 2018.

[13] Thomas Renault. Market manipulation and suspicious stock recommendations on social media. 2017.

[14] Marcelo Mendoza, Barbara Poblete, and Carlos Castillo. Twitter under crisis: Can we trust what we rt? In *Proceedings of the first workshop on social media analytics*, pages 71–79. ACM, 2010.

[15] Zhiwei Jin, Juan Cao, Yongdong Zhang, and Jiebo Luo. News verification by exploiting conflicting social viewpoints in microblogs. In *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.

[16] Eugenio Tacchini, Gabriele Ballarin, Marco L Della Vedova, Stefano Moret, and Luca de Alfaro. Some like it hoax: Automated fake news detection in social networks. *arXiv preprint arXiv:1704.07506*, 2017.

[17] Hunt Allcott and Matthew Gentzkow. Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2):211–36, 2017.

[18] Rens Van de Schoot, David Kaplan, Jaap Denissen, Jens B Asendorpf, Franz J Neyer, and Marcel AG Van Aken. A gentle introduction to bayesian analysis: Applications to developmental research. *Child development*, 85(3):842–860, 2014.

[19] Judea Pearl. Theoretical impediments to machine learning with seven sparks from the causal revolution. *arXiv preprint arXiv:1801.04016*, 2018.

[20] Andreas Vlachos and Sebastian Riedel. Fact checking: Task definition and dataset construction. In *Proceedings of the ACL 2014 Workshop on Language Technologies and Computational Social Science*, pages 18–22, 2014.

[21] Baoxu Shi and Tim Weninger. Discriminative predicate path mining for fact checking in knowledge graphs. *Knowledge-based systems*, 104:123–133, 2016.

[22] Max Steller. Recent developments in statement analysis. In *Credibility assessment*, pages 135–154. Springer, 1989.

[23] William Yang Wang. " liar, liar pants on fire": A new benchmark dataset for fake news detection. *arXiv preprint arXiv:1705.00648*, 2017.

[24] Martin Potthast, Tim Gollub, Matthias Hagen, and Benno Stein. The clickbait challenge 2017: towards a regression model for clickbait strength. *arXiv preprint arXiv:1812.10847*, 2018.

A

[25] Zhiwei Jin, Juan Cao, Yu-Gang Jiang, and Yongdong Zhang. News credibility evaluation on microblog with a hierarchical propagation model. In *2014 IEEE International Conference on Data Mining*, pages 230–239. IEEE, 2014.

[26] Manoel Horta Ribeiro, Pedro H Calais, Virgílio AF Almeida, and Wagner Meira Jr. " everything i disagree with is# fakenews": Correlating political polarization and spread of misinformation. *arXiv preprint arXiv:1706.05924*, 2017.

[27] Elanor Colleoni, Alessandro Rozza, and Adam Arvidsson. Echo chamber or public sphere? predicting political orientation and measuring political homophily in twitter using big data. *Journal of communication*, 64(2):317–332, 2014.

[28] Antonis Matakos and Aristides Gionis. Tell me something my friends do not know: Diversity maximization in social networks. In *2018 IEEE International Conference on Data Mining (ICDM)*, pages 327–336. IEEE, 2018.

[29] Jihui Han, Wei Li, and Weibing Deng. Multi-resolution community detection in massive networks. *Scientific reports*, 6:38998, 2016.

[30] Wayne W Zachary. An information flow model for conflict and fission in small groups. *Journal of anthropological research*, 33(4):452–473, 1977.

[31] Zhongying Zhao, Chao Li, Xuejian Zhang, Francisco Chiclana, and Enrique Herrera Viedma. An incremental method to detect communities in dynamic evolving social networks. *Knowledge-Based Systems*, 163:404–415, 2019.

[32] Anthony Costa Constantinou, Norman Fenton, William Marsh, and Lukasz Radlinski. From complex questionnaire and interviewing data to intelligent bayesian network models for medical decision support. *Artificial intelligence in medicine*, 67:75–93, 2016.

[33] Yuqin Jiang, Zhenlong Li, and Susan L Cutter. Social network activity space sentiment and evacuation: What can social media tell us? *Annals of the American Association of Geographers*, pages 1–16, 2019.

[34] Ophélie Fraisier, Guillaume Cabanac, Yoann Pitarch, Romaric Besançon, and Mohand Boughanem. Uncovering like-minded political communities on twitter. In *Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval*, pages 261–264. ACM, 2017.

[35] Mai ElSherief, Vivek Kulkarni, Dana Nguyen, William Yang Wang, and Elizabeth Belding. Hate lingo: A target-based linguistic analysis of hate speech in social media. In *Twelfth International AAAI Conference on Web and Social Media*, 2018.

[36] Jaziar Radianti, Ole-Christoffer Granmo, Parvaneh Sarshar, Morten Goodwin, Julie Dugdale, and Jose J Gonzalez. A spatio-temporal probabilistic model of hazard-and crowd dynamics for evacuation planning in disasters. *Applied Intelligence*, 42(1):3–23, 2015.

[37] Sondre Glimsdal and Ole-Christoffer Granmo. A bayesian network based solution scheme for the constrained stochastic on-line equi-partitioning problem. *Applied Intelligence*, 48(10):3735–3747, 2018.

[38] Haipeng Guo and William Hsu. A survey of algorithms for real-time bayesian network inference. In *Join Workshop on Real Time Decision Support and Diagnosis Systems*, 2002.

[39] Changhe Yuan and Marek J Druzdzel. Importance sampling algorithms for bayesian networks: Principles and performance. *Mathematical and Computer Modelling*, 43(9-10):1189–1207, 2006.

[40] Gregory F Cooper. The computational complexity of probabilistic inference using bayesian belief networks. *Artificial intelligence*, 42(2-3):393–405, 1990.

[41] Paul Dagum and Michael Luby. Approximating probabilistic inference in bayesian belief networks is np-hard. *Artificial intelligence*, 60(1):141–153, 1993.

[42] GeNIe software academic version, Web: https://www.bayesfusion.com/genie/.

Appendix B

Paper B

Title:	Learning Automata-based Misinformation Mitigation via Hawkes Processes
Authors:	Ahmed Abouzeid, Ole-Christoffer Granmo, Christian Webersik, Morten Goodwin
Affiliation:	University of Agder, Faculty of Engineering and Science, P. O. Box 509, NO-4898 Grimstad, Norway
Journal:	*Information Systems Frontiers*
DOI:	10.1007/s10796-020-10102-8.

Learning Automata-based Misinformation Mitigation via Hawkes Processes

Ahmed Abouzeid, Ole-Christoffer Granmo, Christian Webersik, Morten Goodwin
Department of Information and Communication Technology
Faculty of Engineering and Science, University of Agder
P.O. Box 509, NO-4898 Grimstad, Norway
E-mails: {ahmed.abouzeid, ole.granmo, christian.webersik, morten.goodwin}@uia.no

B

Abstract **— Mitigating *misinformation* on Social Media (SM) is an unresolved challenge, particularly because of the complexity of information dissemination. To this end, Multivariate Hawkes Processes (MHPs) have become a fundamental tool because they model social network dynamics, which facilitates execution and evaluation of mitigation policies. In this paper, we propose a novel light-weight intervention-based *misinformation* mitigation framework using decentralized Learning Automata (LAs) to control the MHPs. Each automaton is associated with a single user and learns to what degree that user should be involved in the mitigation strategy by interacting with a corresponding MHP, and performing a joint random walk over the state space. We use three *Twitter* datasets to evaluate our approach, one of them being a new COVID-19 dataset provided in this paper. Our approach shows fast convergence and increased valid information exposure. These results persisted independently of network structure, including networks with central nodes, where the latter could be the root of *misinformation*. Further, the LAs obtained these results in a decentralized manner, facilitating distributed deployment in real-life scenarios.**

B.1 Introduction

The spread of *misinformation* on SM can have critical consequences during a crisis. Whether the crisis is a disaster, political struggle, terrorist attack, natural hazard, or a pandemic, misleading information such as rumors and false alarm can impede or endanger a successful outcome, such as effective response to a natural hazard. According to a recent study [1], at least 50% of the world's countries suffer from organized political manipulation campaigns over SM. Other examples of the damaging effect of *misinformation* circulated over SM includes the Ebola outbreak in West Africa [2], which was believed to be three times more worse than the previous Ebola outbreaks. Nowadays, with a more connected world, the impact of *misinformation*[1] is getting more severe, even becoming a global threat. For instance, the propagated climate change denying content. There is thus an increasing interest among researchers, and society in general, in finding solutions for combating *misinformation*.

There are two main strategies for combating SM *misinformation* [3]. Some research focus on classifying fake news, rumors, or fake accounts such as social bots, cyborgs, and trolls. Usually, such an approach is referred to as fake news, *misinformation*, or rumor identification. To this end, several solutions have been proposed. For instance, opinion-based or content-based solutions [4] can be used to classify fake news based on textual content. Another approach is to mitigate actively, through proactive intervention [5], or after *misinformation* already is spreading throughout the social network.

Large-scale manipulation carried out across SM during political events is one of the greatest threats to social justice [1]. So-called cyber armies like the Russian Trolls attack on the U.S.A 2016 presidential elections is a well-known example [6]. To the best of our knowledge, most of the attempts to automate the detection of such malicious accounts are not real-time. Furthermore, these cyber armies change their behavior over time, and each context would often require a new model. That is, cyber armies acting in different societies and cultures will have their own linguistic- and behavioral patterns [7]. This diversity makes it difficult to build all-encompassing linguistic models, leading to sub-optimal performance. Too high false negative rate leads to undetected *misinformation* attempts while too high false positive rates can be ethically problematic because accounts or content may be falsely flagged as malicious.

Our work presented in this paper addresses the above challenges by mitigating *misinformation* attempts by countering *misinformation* with rectifying information. That is, we seek to reducing the harmful effects of *misinformation* through a targeted real-news campaign. We propose an approach to single out candidate users for real-news, so as to maximize the remedying effect of injecting real-news into the

[1]The term *misinformation* is sometimes replaced with *disinformation* in some literature. More conveniently, *misinformation* is the unintended spread of malicious content, while disinformation is purposely spreading malicious content. For the rest of this paper, we will use the term *misinformation* to refer to both phenomenons.

social network. A real news campaign can be viewed as a counteraction to the *misinformation* process over the network. That means selecting some users such that by sharing a suggested content through them, a maximal influence would occur on all the other network users, as the latter would became more exposed to valid information.

B.1.1 Hawkes Simulation

Since real-time intervention with SM platforms is not feasible for experimentation purposes, we simulate the process of information diffusion by employing Hawkes Processes (HPs) [8]; [9], as applied in [5]; [10]. HPs are point processes which can model the arrival or occurrence of events, indexed by time or location. There is a range of application domains that fall into such a model. For example, in finance, a HP can describe how a buy or sell transaction on the stock market (an event) influences future stock prices and transaction volume. Similarly, in geophysics, a HP can capture how an earthquake event influences the likelihood of another earthquake event happening as an after effect. For SM, we consider content such as tweets or Facebook posts as events, that have at least time-associated indices. The introduction of new content may trigger a chain of new content, for instance through retweeting, sharing, replying, and quoting.

For all of the above example processes, a HP is particularly suitable because it is a self-exciting point process where the arrival of an event is dependant on the history of all other relevant events. In this paper, we use HPs to modeling each user, so that we can simulate different user behaviors and social network dynamics, including the effect of mitigation.

HPs are random and non-linear, suitable for capturing the unpredictable and intricate nature of SM dynamics. Optimizing mitigation effects thus becomes a challenging problem, involving spatio-temporal reasoning. Furthermore, the randomness, uncertainty and incomplete information on real-life SM aggravates the difficulty of finding a global or a local minimum. We therefore Therefore, we propose a novel LA architecture in this paper, designed to operate in stochastic and unknown environments.

B.1.2 Problem Statement

Let us consider a scenario where a certain amount of *misinformation* is circulating in a social network. The *misinformation* is affecting different users to varying degree, depending on the mix of correct information and misinformation facing each user. We thus define the impact of misinformation for a single user by degree of exposure, relative to correct information [5]. Similarly, the overall influence of misleading information on the whole network could be measured by the average exposure on all users. Since SM events are typical spatio-temporal, these measures should be quantified and reconsidered over different time stages as well.

In order to mitigate the spread of false content, we can apply an intervention-

based strategy to increase the amount of valid information against malicious content, or at least to obtain a balance between the impact of false and true content on the network. The amount of either false and accurate information could be viewed as a counts. Such count represents how much each type of content was generated on the network by each user and at a specific time.

Let A be an adjacency matrix indicating an explicit influence. Let $A_{ij} = 1$ if there is a directed edge or an influence indicating that user i follows user j or quotes (with agreement) content from j, and $A_{ij} = 0$ if not. For a realization of r time steps $\{t_0, t_1, \ldots t_r\}$, let F^{t_r}, and T^{t_r} be the impacts of false- and true content exposure prior to and including time step t_r, respectively. Hence, the impact of both false and true content on user i till the time stage t_r can be calculated as per Equation B.1, and Equation B.2, respectively.

$$F_i^{t_r} = \sum_{s=0}^{t_r} \sum_{j=1}^{n} A_{ij} \cdot F_j^{t_s}, \tag{B.1}$$

$$T_i^{t_r} = \sum_{s=0}^{t_r} \sum_{j=1}^{n} A_{ij} \cdot T_j^{t_s}. \tag{B.2}$$

The outer summation $\sum_{s=0}^{t_r}$ accumulates the impact of information up to and including time step t_r. Furthermore, the impact of *misinformation* on user i, should be measured through all possible chances of being exposed to *misinformation*. That could be achieved by calculating the amount of malicious content from n adjacent users where user i is exposed to their content due to a direct following/ retweeting relationship. The overall average network impacts of both false and true content prior to and including the time stage t_r can be obtained by Equation B.3, and Equation B.4, respectively, where $|U|$ is the cardinality of the network users set.

$$F^{t_r} = \frac{1}{|U|} \sum_{i=1}^{|U|} F_i^{t_r}, \tag{B.3}$$

$$T^{t_r} = \frac{1}{|U|} \sum_{i=1}^{|U|} T_i^{t_r}. \tag{B.4}$$

To achieve actual mitigation during the spread of *misinformation*, a reasonable result would be by making $T^{t_r} \geq F^{t_r}$. That requires some intervention to change the counts which result in T^{t_r}. To apply such interference, we need to obtain the initial counts before modifying them. Therefore, a HP can be engaged to model the quantity of the generated content by each user at various simulation time stages.

B.1.2.1 Counting Generated Content

If we look at a point process on the non-negative real numbers line, where the latter is representing the time, the point process is a random process whose realizations r consist of the event times stages $\{t_0, t_1, \ldots t_r\}$ and they define the time by when an event has occurred. A point process on a specific time realization t_s can be redefined

with an equivalent counting process. A counting process N^{t_s} is a random function defined on a given time stage $t_s \geq 0$, and takes the integer values {1, 2, 3, ...} as the number of events of the point process by the time stage t_s. Hence, a random variable N^{t_s} counts the number of events up to time t_s as the one below.

$$N^{t_s} := \sum_{i \geq t_0} 1_{\{t_s \geq e_i\}}, \tag{B.5}$$

where e_i represents an event occurred by time t_i and $\mathbb{1}_{\{.\}}$ is an indicator function that takes the value 1 when the condition is true, and takes the value 0 when it is false, making it a counting function with a jump of 1 within each time stage it counts for, while starting from the initial time stage t_0 and finishing by the time stage t_s.

The most straightforward class of point processes is the Poisson process. In Poisson processes, the random variables which represent the counts have an inter-arrival time, the rate of such arrivals per a time unit (stage) is denoted as λ, which refers to the intensity of the process. The latter is describing how likely and dense these counts or events to occur in a time sequence. However, in a Poisson process, the inter-arrival times are independent, in other words, the arrival of historical events do not influence the arrival of future events.

A well known self-exciting process was introduced by Hawkes [11], the proposed model was based on a counting process where the intensity λ depends on all previously occurred events. In a HP, the arrival of an event shifts the conditional intensity function to an increase. Such a process determines its conditional intensity output based on two fundamental quantities, base intensity μ, and historical events arrival prior to a certain point in time H^{t_s}. With an analogy to *Twitter* and the problem of *misinformation*, the counts are the number of tweets, either true or dishonest ones. The base intensities can be viewed as the exogenous motivational factors which influenced a user to react, while the historical events can be viewed as the network endogenous factors, for instance, how the sub-network of user followees are acting on the network. In order to mimic *Twitter* dynamics as an environment for our mitigation method, we consider MHP by defining U-dimensional point processes $N_U^{t_s}$, where U is the network users set, which emphasizes the self-excitation between events on SM [12]. N^{t_s} can be interpreted as F^{t_s} or T^{t_s} as described in Equation B.1 and Equation B.2, while $|U|$ is the number of individual users a single HP is associated with, and t_s is the specific time realization or stage. The best way to describe a HP, is by its conditional intensity function as per Equation B.6.

$$\lambda_i(t_s|H^{t_s}) = \mu_i + \sum_{t_{s'} < t_s} g(t_s - t_{s'}), \tag{B.6}$$

where g is some kernel function over the history associated with user i from the time stage $t_{s'}$ prior to time t_s. g is concerned with the history of some influence $A_{i.}$. We used an exponential decay kernel function $g = A_{i.}e^{-wt}$ as practiced in [5], where w is the decay factor which represents the rate for how the influence is reduced over time. For all U, the base intensity vector μ, and the influence matrix A can be estimated using maximum likelihood as presented in [13]. To simulate *Twitter*

dynamics for our mitigation task, we can rewrite Equation B.6 as the one below.

$$\lambda_i(t_s|H^{t_s}) = \mu_i + \int_0^{t_s} g(t_s - t_{s'})dN(t_{s'}), \tag{B.7}$$

where $N(t_{s'})$ is the integration variable and the count of the historical generated content that influences user i and determined by the HP. In turn, the conditional intensity $\lambda_i(t_s|H^{t_s})$, tells how likely user i would act and generate content herself, by the time t_s. Since $N(t_{s'})$ is interpreted as $F(t_{s'})$ and $T(t_{s'})$ and can be calculated from Equation B.1 and Equation B.2, it is important to highlight that the influence matrix A is considered explicit influence (following/ retweeting) when measuring content impacts on users after the simulation. On the other hand, A is considered implicit or hidden temporal influence when calculating the conditional intensity function, since the latter is a result of the estimated Hawkes parameters before the simulation, and indicating the independence from the network explicit structure. That is, we estimate the simulation parameters to obtain reasonable and inferred simulated network dynamics from the hidden temporal influence, then, we measure content impacts based on the explicit relationships on the simulated network dynamics.

B.1.2.2 Limited Budget Mitigation

To observe the process of the intervention-based mitigation, we followed a social network reshaping approach as employed in previous work [5]; [10]; [14]. To achieve such a resolution on the network, we are interested in the base intensity μ, since it defines any external motivation on the users. Hence, we are interested in adjusting the value of the base intensity by increasing it. However, there are two main challenges in this method, not all users would respond to an exogenous motivation, and not all of them are capable of boosting the activity of the network. Besides, some users would be spreading *misinformation* on purpose and they will not respond to an opposite campaign. Moreover, the time spent for incentivizing users is limited due to the crisis time criticality. Therefore, the modification of μ is bounded by a small amount of incentivization that should be allocated wisely among users to reach the optimum mitigation results.

Let us denote C as the optimization constraint, which represents the limited budget of incentivization. The optimization objective is to minimize the difference between *misinformation* and valid information impacts on the network by incentivizing the true content simulation-base intensities of users with respect to C. We define the optimization problem as a stochastic knapsack problem [15], where the selection of some users at a specific time stage is aimed in order to maximize the mitigation performance. The stochastic knapsack solution is bounded by the maximum allowed amount the knapsack can afford, in our case, this is referred to as C.

The purpose of the knapsack optimization is to fill a knapsack with materials amounts $X = \{x_i, ..., x_n\}$ such that they maximize some value $\mathcal{F}(X)$ but, at the same time, staying within the limited capacity of the knapsack ($\sum_{i=1}^{n} x_i = C$).

With an analogy to our problem, we can define the below minimization objective and constraint functions.

$$\min \mathcal{F}(X) = \sum_{i=1}^{|U|} \mathcal{F}(x_i), \text{where } \mathcal{F}(x_i) = \frac{1}{|U|} \sum_{i=1}^{|U|} F_i^{t_r} - \frac{1}{|U|} \sum_{i=1}^{|U|} T_i^{t_r}, \tag{B.8}$$

$$\text{subject to } \sum_{i=1}^{|U|} x_i = C, \text{where } x_i > 0 : \forall u_i \in U, \tag{B.9}$$

where x_i is the incentivization amount for the user i and both $F_i^{t_r}$ and $T_i^{t_r}$ are random variables generated from a Hawkes count process $N(t_r)$ prior to realization r at time. Where $F_i^{t_r}$ is calculated through the simulation Equation B.7, and $T_i^{t_r}$ is calculated through the simulation Equation B.10, with replacing N by F and T in both equations, respectively. Therefore, the optimization problem is stochastic with regard to the objective function. Hence, and by finding the optimum incentivization amount x, the intervention can be applied by employing another HP for each user as the one below.

$$\lambda_i(t_s|H^{t_s}) = x_i + \mu_i + \int_0^{t_s} g(t_s - t_{s'}) dN(t_{s'}), \tag{B.10}$$

where $N(t_{s'})$ represents the count of true information events in the Hawkes model prior to the specified time stage t_s, giving that $t_{s'} < t_s$, and $t_s \leq t_r$ when r realizations (time stages) are the time intervals of the whole process.

B.1.3 Paper Contribution and Limitation

It is essential to highlight that our approach is different from traditional approaches for finding graph centrality measures or most influential users on a social network [16]; [17]. That is because our method would be under-performing if applied to a network where most influential users are spreading fake news. On the other hand, our purpose is to learn normal users who can be effective at a specific moment and independently from the graph structure and network centrality measures. Such independence is a crucial advantage of our approach, since it allows for further exploration and analysis of the temporal hidden influence structure on social networks. Besides, the timing driven feature is fundamental to crisis mitigation applications.

This paper introduces an adaptative learning method to achieve stochastic optimization over a social network. The optimization task is constrained and stochastic regarding its objective function. We applied our experiments on *Twitter* data, and evaluated our model on two publicly available real-world datasets that were used in previous work. Namely, *Twitter15* and *Twitter16* datasets [18]; [19]; [20]. Moreover, we introduce a new *Twitter* dataset for the *COVID-19* pandemic. The latter represents a different situation that would demonstrate the flexibility of our solution. Results showed that with our light-weight computation method, we were able to find at least a local minimum that serves the required mitigation aims.

In our solution, we used a LA [21] as an adaptative learning technique. A LA is a stochastic model, operating in the framework of Reinforcement Learning (RL). The LA has been found to be a robust method for solving many complex and real-world problems where randomness is a primary characteristic of the problem. We built a network of LA, while each is assigned to a user on the social network. The individual LA should learn if the associated user is a good candidate for the mitigation campaign or not. Additionally, each LA determines an amount from a limited budget. The amount reflects how much we can spend from a limited budget. The latter can be viewed as an optimization constraint and the capacity of how likely we would depend on each of the suggested candidate(s), who would be part of an intervention process.

Our LA-based method differs from previous *misinformation* mitigation with RL approaches [5]; [10]. The latter had a three dimension state per a user at a time, where user amounts of true and false events were observed with number of "like" responses received. Then, the task was to learn a mitigation policy over the constructed state space. On the other hand, we redefine the task as a natural optimization problem, and we reconsider the problem of state space from being multidimensional to a single dimension, considering an overall network objective function with one single variable instead of calculating a multidimensional function across all users. Therefore, our objective function is only calculating one single value per a user, that drastically reduced the solution state space required for convergence, since the number of users on the network becomes the size of the state space. That means linear state space increase instead of exponential in case of scaling up the solution. We present our empirical results that show how that led to a faster and more reliable resolution without a notable loss in accuracy, where differences between users exposures to fake and true news have no skewed distribution.

We propose a novel exercise of the LA in the domain of SM *misinformation* mitigation. To the best of our knowledge, LA-based approaches were not employed in that area, and this paper is the first to conduct an LA study on online *misinformation* mitigation tasks. We approach that by evaluating three primary learning schemes for the LA. Moreover, and compare to similar mitigation approaches [5]; [10]. However, as a limitation in our work, we do not consider the political bias of users, compared to what has been done in [10], in addition, we focus on non-skewed data points distribution scenarios where an average value could be the performance measure. Hence, political bias and skewed data scenarios are left for future improvements.

We evaluate our method and our implementation of the HPs-based simulation on two baselines datasets (*Twitter15*, *Twitter16*) after applying some post-processing on the original data. Furthermore, we introduce a new dataset (*Twitter-COVID19*) which was collected and annotated by us, and represents a different definition from the traditional fake news cases. The new dataset demonstrates the applicability and flexibility of our approach in different situations, such as the infodemic of *COVID-19*. In such a scenario, a mitigation task would be targeting the reduction of some propagated content effects, for instance, the irrational statements about an already-found cure or any incorrect crisis relief content that might motivate people to be

less careful. Our experiments show promising results on all the evaluated datasets.

B.1.4 Paper Organization

This paper is organized as follows. section B.2 introduces a literature review, where some of the previous adaptation of HPs are mentioned. Furthermore, other intervention-based mitigation approaches are briefly explained. The applied method and datasets are explained in section B.3, a statistical comparison between the datasets is demonstrated as well. section B.4 shows our empirical results and performance metrics. A discussion with comparable results to other mitigation methods is demonstrated in section B.5. Eventually, we conclude the paper and highlight possible future work in section B.6.

B.2 Related Work

The definition of fake news or *misinformation* has evolved through time, not only due to the increased complexity of such social problem but also because of how recent technological efforts have progressed. For a long time, the spread of fake news on SM has been considered as the intentional dissemination of false information in news articles [22]. However, other research work started to give an attention to the broader scope of the problem [3]; [61]. For instance, rumor detection [23], malicious accounts classification [6]; [24], and the causal aspects of *misinformation* [25]. However, and to the best of our knowledge, it has been an obstacle to effectively solve the problem in real-time or without data selection-bias concerns. Moreover, ethical questions are being asked [26] since fake news detection solutions are judgemental by nature. Therefore, the need for safer and online strategies that would lead to more generic and authentic resolutions are critically desirable.

As a motivation for more online resolutions to SM *misinformation*, intervention-based mitigation strategies have been practiced in the literature. Reshaping users activities by applying an interference strategy was introduced in [27]. In addition, dynamic programming was employed to optimally distribute incentivization resources among users in different time stages [28]. In such previous work, objective functions were designed using expected values of exposure counts of the user content, generated from a HP. The latter has been applied as a simulation for the SM information diffusion in many recent applications as well [29]; [30]. MHPs have proven efficiency and robustness in SM analysis and more specifically in the domain of *misinformation*.

Since recent advances in *misinformation* mitigation approaches have achieved an online and interactive (simulation-based) resolutions [5]; [10], future work would focus on improving and wider applications rather than a narrow definition of *misinformation* or limited datasets that were examined in previous work [18]; [19]; [20]. For example, the conducted intervention-based mitigation could be used for polarization, hate speech, and infodemics mitigation resolutions. The latter is one of our interests at this study.

The use of RL for *misinformation* mitigation on SM has revealed a promising future for how such a problem could be tackled. However, to model a large state space as practiced in [5] and to evaluate an optimum policy through that, is still a big concern, especially for a mitigation task that needs to be achieved timelessly. Moreover, user incentivization should be applied according to the problem context, which causes a loss of generality in some cases. For example, in a political scenario, a mitigation objective function would consider the political bias of users [10] before preference them as suitable candidates, since users with an ideological bias would not respond to the incentivization [22]. The conducted study in this paper aims to provide a light-weight computation framework that could be applied to different mitigation contexts without a total re-engineering effort. However, we consider this paper as the first step for our proposed system structure by evaluating a network of LAs where a light-weight LA is the core of our framework. Therefore, the political bias of users is not investigated so far.

A LA is an adaptative learning method which can be viewed as a stochastic model operating in the framework of RL. A LA has been found to be a robust method for solving many complex and real-world problems where randomness is a primary characteristic of the problem. Previous applications of LAs have been introduced for social network analysis problems. For instance, a stochastic learning-based weak estimator for learning and tracking a user's time-varying interest was practiced for SM-based recommendation systems [31]. A LA-based framework was also employed for online service selection in a stochastic environment where the latter has unfair service reviews [32]. A stochastic constraint optimization problem such as the one approached in this paper could be defined as a stochastic knapsack problem, where LA has been tasked for by employing the Learning Automata Knapsack Game (LAKG) [33].

LAKG is a game between n finite automata that interact with a scheduler and a stochastic environment. The stochastic environment consists of a set of stochastic material unit volume value functions. If an amount of a certain material is suggested to the environment and favored, the associated value function takes the value 1 with probability p and the value 0 with probability $1 - p$. Besides, the environment provides a signal ϕ, which indicates whether the knapsack is full or not, which also tells if the optimization constraint was reached or not. On the other hand, the scheduler takes material amounts as its input. The purpose of the scheduler is to perform the access to the stochastic environment, sequentially. Besides, it makes sure that the unit volume value functions are accessed with frequencies proportional to all materials amounts. Such a problem description is similar to an incentivization across n users, where the incentivization budget is limited. Therefore, we consider a learning scheme similar to [33]. However, in that knapsack problem, n materials can be evaluated through the whole problem space. Still, on a large social network, it would be impossible to assess a value function across the entire problem space. Therefore, we adapted a different structure to face such an obstacle. Moreover, due to our problem specification, we distinguished between the learned and the evaluated material amount for the value function.

B.3 Methodology

B.3.1 Learning Automata

A LA is a model of intelligent computation where learning is accomplished by exploring and its consequences, in an iterative and reinforcement manner. The objective is to decide the optimal action to select between all the possible actions. A LA learns by interacting with an environment. The environment sends a regular feedback for when the LA explores a particular action. The LA estimates its preferred action selection in future explorations with regard to the environment recent response.

SM platforms and mainly in an emergency setting, are considered as random mediums where uncertainty and outliers exist. Such uncertainty and randomness motivated this paper to practice an LA-based optimization framework as a network of LA, while each LA is assigned to a user on the social network to learn about its authenticity probability of being part of a *misinformation* mitigation strategy. Figure B.1 shows how an individual LA works by interacting with an environment. Each LA task is to learn about the best action between two possible moves. That is, each LA has two possible moves (α_0, α_1) over an associated random walk line, representing moving in the direction of assigning less incentivization and the direction of assigning more incentivization from a mitigation budget C, respectively. For a given LA_i and its chosen action α_1 and incentivization amount x_i at an epoch i, the environment sends the feedback V_i^{i} according to Equation B.11.

$$V_i^{\mathrm{i}} = \begin{cases} 1, & \text{if } \mathcal{F}^{\mathrm{i}}(x_i) < \mathcal{F}^{\mathrm{i}}(x_j) \\ 0, & \text{otherwise} \end{cases}, \text{ where } i \neq j, \tag{B.11}$$

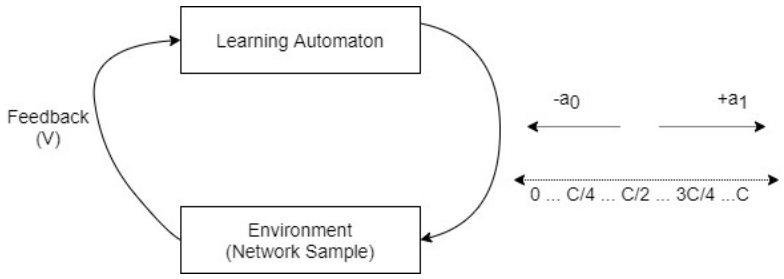


Figure B.1: LA interaction process

where $\mathcal{F}^{\mathrm{i}}(x_i), \mathcal{F}^{\mathrm{i}}(x_j)$ are some investigated objective functions at epoch i for LA_i and LA_j at the epoch iterations i and j, respectively. An objective function calculation is done as per our mitigation objective function definition in Equation B.8. However, with adapting such calculation in our framework, and for a faster computation, we practically reduce the size of the network by randomly sampling over a subset of users U^-, while $|U^-|$ is selected according to the minimum subset size which does not sacrifice the accuracy of the results, since larger subsets might improve the process of user evaluation but will slow the computation. For that, we conducted a grid search to estimate the best value of $|U^-|$. A detailed grid search result and the final selected hyper parameters values are demonstrated in section B.4,

with respect to the convergence rate and the mitigation performance metric. Considering a network sample $|U^-|$ instead of the whole network size $|U|$, an objective function at a given epoch i can be redefined as the one below.

$$\mathcal{F}^{\mathrm{i}}(x_i) = \frac{1}{|U^-|}\sum_{i=1}^{|U^-|} F_i^{\mathrm{i}} - \frac{1}{|U^-|}\sum_{i=1}^{|U^-|} T_i^{\mathrm{i}}. \tag{B.12}$$

For each an epoch i inside a specific time stage t_s, all LAs are sequentially visited one time. Therefore, the function $\mathcal{F}^{\mathrm{i}}(x_j)$ is also calculated and compared with $\mathcal{F}^{\mathrm{i}}(x_i)$, to evaluate how two mitigation functions are different in terms of a minimum value, since the target is to minimize the overall network difference $(F^{t_s} - T^{t_s})$ while learning the optimum subset of users $U^*(t_s)$.

The environment feedback V is stochastic since the mitigation sub-functions are a result of a stochastic process, namely, a MHP. Therefore, the challenge of our stochastic optimization framework is to learn how to minimize $\mathcal{F}(X)$ under the constraint C as discussed in Equation B.8, and Equation B.9. During the learning process, the value X is determined by a learning rate η, as a constant per all users, time stages, and epochs. For instance, while $i \neq j$, $x_i = x_j = \eta$, since evaluating different users should be fairly applied by assigning the same incentivization amount. The hyper parameter η can be estimated through our grid search as well. On the other hand, and per each user, the determined X for the mitigation will be the final converged state of each LA, since that indicates an amount determined through the LA interaction with the environment. Hence, the final converged network incentivization values is a random vector which represents the converged states of the individual LA over a random walk line for each. section B.4 gives more details about how different the learning rate η could be on each dataset and what are the factors that dictates its values.

We designed the network with a uniform learning scheme for the individual LA. Furthermore, we followed a simple random walk as in [33] to represent the LA state transitions. However, our individual LA are different in the way they interact with the environment and in their structure as well. In our case, the environment is partially observed, we refer to that by network sample size $|U^-|$. We have evaluated three main learning schemes for our framework, we refer to them as *random walk reward-penalty in action* RW_{RP}, *random walk reward in action* RW_{RI}, and *random walk penalty in action* RW_{PI}. Figure B.2 demonstrates the components of our framework.

As one of the components in our framework, a shuffler, which is triggered every an epoch after all LAs are visited, to ensure the comparison pair $\mathcal{F}^{\mathrm{i}}(x_i), \mathcal{F}^{\mathrm{i}}(x_j)$ will be different each time. The sampler component selects a subset of the overall network with size $|U^-|$, determined by a hyper-parameter in our configuration. The scheduler component maintains sequential and equal visits to each LA to guarantee equal potential state transitions and actions probabilities update behaviour. The memory component at each computation time step inside an epoch, helps the sampled partial environment to send its feedback according to the evaluation of the functions $\mathcal{F}^{\mathrm{i}}(x_i), \mathcal{F}^{\mathrm{i}}(x_j)$.

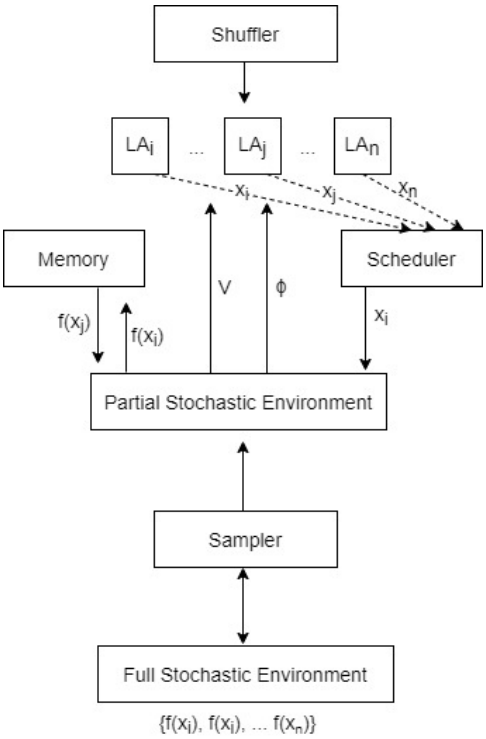


Figure B.2: Mitigation framework structure

For our optimization framework to work efficiently, each LA should have two updating rules, the first is the LA action probability update rule, the second is the state transition mechanism, which eventually decides the amount of incentivization that will be assigned for each user. Therefore, the incentivization amounts are considered shifts (left/ right) in a line where a random walk is exercised. The line represents the LA state space. The shifts are probabilistic and subject to the LA action probabilities and the signal ϕ. The latter tells the LA if the budget constraint C has been met or not yet. Each LA state space has its own boundaries from 0 to C, indicating a minimum and a maximum allowed state values, respectively. Hence, the actions (moves) probabilities are updated according to the environment feedback V at each individual LA visit (epoch). Equation B.13 describes the action probability updating rule. The higher the probability of moving to the right (α_{i_1}), the more likely the user i is a good candidate for the mitigation.

$$\mathcal{P}^{\mathrm{i}}(a_{i_1}) = \frac{W^{\mathrm{i}}_{i_1}}{Z^{\mathrm{i}}_{i_1}}, \text{ and } \mathcal{P}^{\mathrm{i}}_{i}(a_{i_0}) = 1 - \mathcal{P}^{\mathrm{i}}_{i}(a_{i_1}), \text{ where } W^{\mathrm{i}}_{i_1} = \sum_{e=1, e\leq \mathrm{i}}^{\mathrm{i}} V^{e}_{i}, \tag{B.13}$$

where $W^{\mathrm{i}}_{i_1}$ and $Z^{\mathrm{i}}_{i_1}$ are counters for how many times the action a_{i_1} was rewarded and selected starting from first epoch e and till the current epoch i, respectively. The RW_{RP} learning scheme is in action when LA_i moves are either rewarded or penalized. Hence, the actions probabilities are updated all the times when LA_i interacts with the environment. Algorithm A.1 gives the complete details of how RW_{RP} works and how the state transition is applied.

Differently, the RW_{RI} learning scheme is in action when LA_i moves are only rewarded. Hence, the actions probabilities are updated only when $V_i^{\mathrm{i}} = 1$, according to the recent interaction feedback of LA_i with the environment. Algorithm A.2 demonstrates how RW_{RI} works and its state transition.

The RW_{PI} learning scheme works similar to RW_{RI} except the former is in action when LA_i moves are only penalized. Which causes the actions probabilities to be updated only when $V_i^{\mathrm{i}} = 0$. Algorithm A.3 gives the complete details of the state transition and how RW_{PI} works. For any LA, since our purpose is to evaluate for the incentivization, the action α_1 is always performed for all our learning schemes, and $\mathcal{P}(\alpha_0)$ is only updated as per Equation B.13.

B.3.1.1 Random Walk Learning

As a result of learning the incentivization amounts per users, by the end of all computation steps of $\mathcal{I}$ epochs on each time realization t_s of the HP, each LA suggests if its associated user would be part of a proposed subset $U^*(t_s)$ of mitigation candidate(s). Where $\forall u_i \in U : u_i \in U^*(t_s), x_i = S_i(t_s)$, if $\mathcal{P}(\alpha_{i_1}) > \mathcal{P}(\alpha_{i_0})$ by the end of the computation, and $S_i(t_s)$ is the final converged LA state value of user i at the time realization t_s. Therefore, $S_i(t_s)$ will be the final decided assigned value to the variable x_i for the intervention process as demonstrated in Equation B.10.

In the core of our proposed framework, there is a decentralized LA learning model, which learns such final incentivization amount for a user. The learning model learns by performing stochastic moves (actions) over a state space (possible incentivization amounts). The stochastic moves $\mathcal{P}(\alpha_{i_1}), \mathcal{P}(\alpha_{i_0})$ are determined as per Equation B.13, which is dependant on the environment feedback which is measured according to Equation B.11, and Equation B.12. Despite how the random walk moves probabilities are being updated through the different learning schemes (RW_{RP}, RW_{RI}, RW_{PI}), at a specific MHP time stage, and an epoch i, the user i associated LA model learns by conducting random walk moves as the below formal description.

$$
\begin{aligned}
S_i^{t_s}(\mathrm{i}+1) &:= S_i^{t_s}(\mathrm{i}) + \frac{C}{M}, && \text{if } \mathcal{P}(\alpha_{i_1}) > \mathcal{P}(\alpha_{i_0}) \text{ and } 0 \leq S_i^{t_s}(\mathrm{i}) < C \\
& && \text{and } \neg\phi, \\
S_i^{t_s}(\mathrm{i}+1) &:= S_i^{t_s}(\mathrm{i}) - \frac{C}{M}, && \text{if } \mathcal{P}(\alpha_{i_1}) < \mathcal{P}(\alpha_{i_0}) \text{ and } 0 < S_i^{t_s}(\mathrm{i}) \leq C, \\
S_i^{t_s}(\mathrm{i}+1) &:= S_i^{t_s}(\mathrm{i}), && \text{otherwise,}
\end{aligned}
$$

$$
\text{where } \phi = \begin{cases} true, & \text{if } \frac{C}{M} + \sum_{i=1}^{|U|} S_i^{t_s}(\mathrm{i}) > C, \\ false, & \text{otherwise,} \end{cases}
$$

where M is the constant memory depth of each LA state space bounded by C, therefore $\frac{C}{M}$ describes the shift value resulted by the random walk. Since all users should be evaluated through their own random walk model, the above description is applied on all LAs, sequentially through the scheduler component as indicated in Figure B.2. Moreover, Figure B.3 gives a toy example of how a joint random walk learning process from two sequential LAs moves constructs the network converged incentivization vector S^*. Where the horizontal line indicates user i state space, and the vertical line indicates user j state space, given that the latter ended up being allocated all the incentivization budget ($C = 2$) after two epochs (i = 2) and 4 time steps. The example then can be generalized for as many number of users LAs. Eventually, it is important to highlight that at each time stage of the MHP, the individual LA moves probabilities $\mathcal{P}(\alpha_{i_1}), \mathcal{P}(\alpha_{i_0})$ are reset, to ensure learning new temporal influential users over different time stages, if exist.

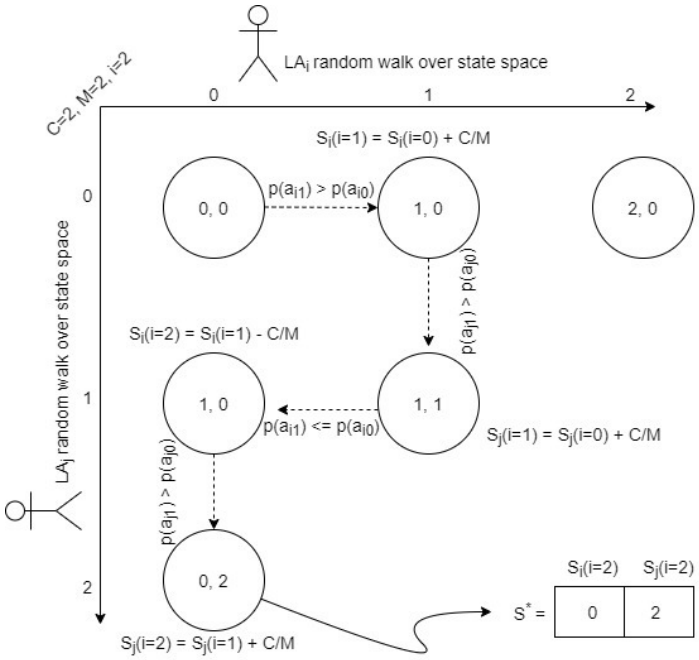


Figure B.3: Toy Example of two LA-based joint random walk

B.3.1.2 Rate of Convergence

For an individual LA_i, its convergence is defined according to the optimum incentivization value (random walk converged point) S_i^*. Hence, the whole LA network optimum random vector S^* is considered the optimization minimizer vector. Therefore, the rate of convergence or the asymptotic error of the LA network can be defined as per Equation B.14, where i represents the current epoch (LA visit).

$$error := \frac{||S_{\mathrm{i}+1} - S^*||}{||S_{\mathrm{i}} - S^*||}. \tag{B.14}$$

The above definition is then used for evaluating the network hyper parameters. Additionally, from our observation, at least for one of the applied LA learning

schemes, we obtained a *Q-superlinear* rate of convergence for all datasets, where the network asymptotic error approached 0. A detailed explanation for the network rate of convergence using different hyper parameters values is given in section B.4.

B.3.2 Datasets

B.3.2.1 Twitter15

The *Twitter15* dataset has initially been collected and created to debunk rumors on *Twitter* [18]. The original dataset had both a political and a more general context. Two rumor tracking websites (*snopes.com and emergent.info*) were used to verify the trustworthiness of the content before categorizing the data into true, false, and unverified rumors. The results consisted of 94 true and 446 fake stories. Accordingly, all relevant and matched tweets were collected and labeled. After downloading and post-processing the original *Twitter15* dataset, we obtained 27,547 users, which contributed to 21,279 of true events (*tweets/ retweets*), and 6,268 of false ones. However, for our experiments and due to the current limitation in our computation power, we scaled down the size of the network to only 1,039 users, and 1,188 events, considering scaling that up in future experiments. Our scaled network also focused only on the American political context. Therefore, the final network was a result of extracting main tweets with relevant keywords and hashtags from the standard dataset. Hence, we only extracted main tweets that contain the keywords and hashtags as demonstrated in Table B.1. The final network had approximately 94.02% of *misinformation*.

For all datasets, we define the term $u_{influencer}$ as the user with the highest node degree on the network, with regard to the number of edges which represent retweeting from her. In *Twitter15* dataset, the top *misinformation* influencer node motivated 301 users to spread false content. Also, 13 users were motivated to spread true news by retweeting valid content, since the top influencer user had generated some trustworthy content as well, that opens a judgemental question if such a user is spreading *misinformation* on purpose or not, and how much spread is enough to measure that in our study. We believe, such question is irrelevant to this study, however, we will involve these numbers when looking at the results and $U^*(t_s)$, since we aim to have an independence from the network high centrality node(s) when it is necessary.

B.3.2.2 Twitter16

The *Twitter16* dataset was collected initially for a recurrent neural network for rumor detection in SM [19]. The data was evaluated by the online rumor debunking service (*snopes.com*), where 778 events were investigated during March-December 2015, and 64% of the data samples were actual rumors. Similar to *Twitter15*, the context of the events are broader than only political struggles. Hence, and after our post-processing, we ended up with $45,566$ incidents and $44,114$ users, which contributed to $38,686$ non-rumors, and $6,880$ rumors. However, and after scaling

Table B.1: Filtered *Twitter15/16* datasets tweets

Hashtag/ Keyword	**(*Twitter15*) Number of main tweets**	**(*Twitter16*) Number of main tweets**
hillary	16	21
trump	27	37
obama	44	30
america	4	3
americans	12	7
american	19	3
democrat	5	5
republican	1	1
clinton	16	27
white house	2	39

down and focusing only on political struggle related events, our final dataset version was $1,206$ users and $1,362$ cases, from which there were around 45.59% considered as *misinformation*. Table B.1 shows the hashtags and keywords used for filtering the main tweets from the standard dataset.

The top *misinformation* influencer node incentivized 230 users to spread malicious content. However, the same user has also incentivized 52 users for retweeting correct information, which means she might not be spreading false content on purpose. Like in *Twitter15* dataset, that insight is useful when evaluating the performance of our method, since the latter should not be driven by such top nodes, especially when they are circulating fake content on purpose. Therefore, some relevant statistical measures would be useful to set a boundary for how we could accept the learned $U^*(t_s)$, in cases when $u_{influencer} \in U^*(t_s)$.

B.3.2.3 Twitter-COVID19

Our new proposed dataset *Twitter-COVID19* was collected during the 28th of March 2020 for the *COVID-19* infodemic on *Twitter*, the dataset had $1,164$ users and $1,180$ events, from which there were around 92.03%, manually labeled as *misinformation*. The dataset focused on the circulated irrational content about some found cures like "silver liquids" and the "anti-malaria" medication. The latter was reported as a cause of severe harmful side effects for people who tried it without consulting a health expert [2]. To show the flexibility of our approach, the mitigation resolution can be seen as applicable in any case where there are two opposite campaigns, and the task is to mitigate one in favor of the other. In our case, we consider a reduction of the effect of believing these false crisis reliefs, since there was no approved cure yet, by the time of collecting the dataset. That can be viewed as an exercise for risk reduction during an infodemic.

In *Twitter-Covid19* scenario, the top *misinformation* influencer node motivated 766 users to spread irrational content. The influencer node had no effect on spreading

[2]https://www.theguardian.com/world/2020/mar/24/coronavirus-cure-kills-man-after-trump-touts-chloroquine-phosphate

any other type of contents during the time window of collecting the data. Therefore, that is considered a perfect example to evaluate how our method would avoid such user before suggesting $U^*(t_s)$. Table B.2 demonstrates some statistical differences between the three datasets.

Table B.2: Datasets statistics

Dataset	***Twitter15***	***Twitter16***	***Twitter-COVID19***
Num of users	1,039	1,206	1,164
Number of events	1,188	1,362	1,180
Misinformation	94.02%	45.59%	92.03%
Network density	.001943	.001778	.001687
Misinformation by $u_{influencer}$	28.87%	19.07%	65.81%

B.4 Empirical Results

B.4.1 *Twitter15*

Our first experiment was the *Twitter15* dataset, which had around 94.02% of *misinformation*. This is considered an important example for evaluating our algorithms on such high percentage. For the HPs, we set the decay factors $w = .75$ and $w = 1$ as in [10] for false and true events, respectively. Besides, we set an hourly interval between time stages ($\Delta = 1$ hour). From the dataset events timestamps, we used the first 10 hours for learning the Hawkes parameters μ and A, before simulating the next 30 hours. Therefore, we used the next 30 hours from the real data for testing, by comparing with events arrivals which were generated from the simulation. We obtained a relatively good simulation behaviour. Figure B.4 indicates the average absolute difference error [5] as the performance metric used for the simulation on both true and false events from *Twitter15* and *Twitter16* datasets. Equation B.15 demonstrates how the average absolute difference error was calculated.

$$\mathcal{E}_{t_s+\Delta} = \frac{1}{|U|}\sum_{i=1}^{|U|} |[N_i^{\mathcal{H}}(t_s+\Delta) - N_i^{\mathcal{H}}(t_s)] - [N_i^{\mathcal{R}}(t_s+\Delta) - N_i^{\mathcal{R}}(t_s)]|, \tag{B.15}$$

where $|U|$ is the number of users and $N^{\mathcal{H}}$, $N^{\mathcal{R}}$ represent the counts of the arrived events from Hawkes simulation and real data, respectively. The calculation is made between the time stages $t_s + \Delta$ and t_s. It is important to highlight that we consider improving the simulation process in future, so that we could maintain a more stable error over time. However, we believe an error up to 1 is still a good indicator since for 1,000 users, that means, on average, there is only 1 event arrival difference per user and prior to a certain time stage. Additionally, we define a random count range for more convenient simulation results, the count range $N^t_{\mathcal{E}\pm}$ interprets $\mathcal{E}$ as a possible discount in the number of generated events. For instance, if $N^t = 50$ from

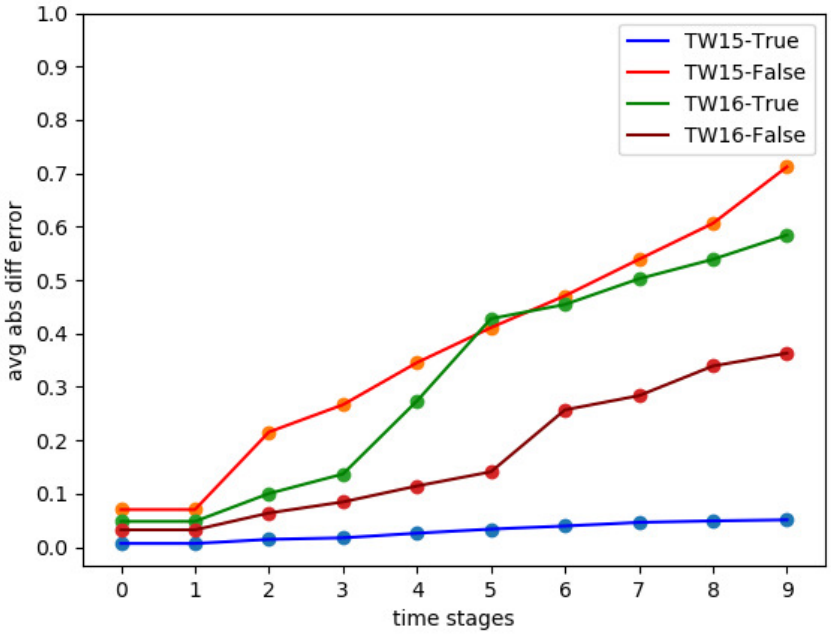


Figure B.4: MHP simulation vs real data on scale of 1 less or more tweet per user difference on a time stage

the simulation, we might be more confident to say $N^t_{\mathcal{E}^\pm} = [(1-\mathcal{E})\cdot 50, 50]$. In case of $\mathcal{E} > 1$, it should be normalized between 0 and 1.

For *Twitter15* dataset, Figure B.5 shows the optimization performance of the three suggested learning schemes with three other considerable performance baselines. The three baselines represent three different measures we sought to outperform, these are *misinformation* before mitigation, uniform distribution of incentivization budget, and random allocation of incentivization budget. Our optimization framework outperformed the three baselines with the three learning schemes with a budget $C = .05$. The latter is considered a limited budget according to its overall effect on the MHP. Moreover, we observed approximately similar performance between RW_{PI} and RW_{RP} on longer epochs, but RW_{RP} was the one with a remarkable early convergence. Eventually, Figure B.6 shows the performance for difference minimization between *misinformation* and true events after learning U^* for the first three time stages (next three hours).

B.4.2 *Twitter16*

The simulation driven from the *Twitter16* dataset can be evaluated as in Equation B.15. And as indicated in Figure B.4, the false events simulation seemed to be slightly enhanced compared to *Twitter15*.

Our version of the *Twitter16* dataset is considered as an interesting case, since it is a situation where *misinformation* is around 45.59% over the network, that is considered too low, relatively to both *Twitter15* and *Twitter-COVID19* datasets.

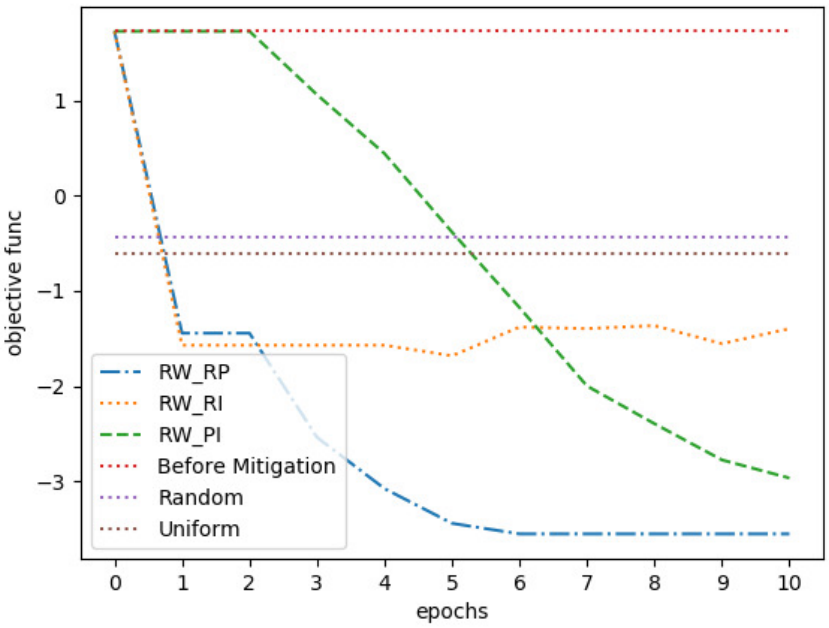


Figure B.5: *Twitter15* mitigation performance on 1'st time stage, $C = .05$.

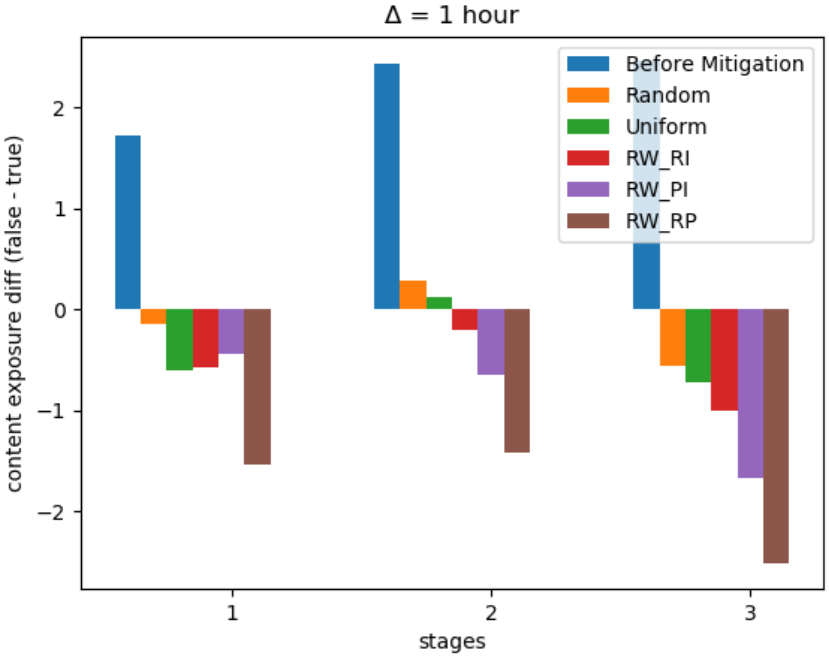


Figure B.6: *Twitter15* mitigation performance on first three time stages, $C = .05$.

However, it seemed that such scenario was more challenging. That is because when an extreme level of *misinformation* exists, it becomes straight forward to distinguish between nodes with temporal negative influence. On the other hand, when there is a balance between both campaigns in the network, it is more vague to distinguish between the authenticity of nodes, since all nodes might be contributing to both *misinformation* and true information diffusion. Moreover, it becomes interesting to see how our method was independent from the network central node(s) in such cases. The latter perspective is fundamental, as it would indicate how flexible and intelligent our method is. For instance, sometimes in a scenario like *Twitter15* and *Twitter16*, a top influencer node might be spreading false content, but still, it is circulating true content.

For the HPs, we set the decay factors $w = .6$ and $w = 1$ as in [10] for false and true events, respectively. Besides, we set an hourly interval between time stages ($\Delta = 1$ hour). We used the same duration as in *Twitter15* for both learning Hawkes parameters and testing.

With a budget $C = .05$, Figure B.7 indicates how our optimization framework with the learning scheme RW_{RP} performed well with more stability, compared to other learning schemes in addition to another two baselines. However, the uniform distribution method performed approximately the same. Moreover, Figure B.8 shows how the random distribution outperformed all other methods for the next two time stages. We consider this as an interesting example of how RW_{RP} continued to be more robust compared to other LA-based learning schemes, but it failed to compete with both uniform and random strategies. However, a slight improve in the difference between RW_{RP} and the uniform and random distribution strategies can be noticed in Figure B.9, after repeating the experiment with only 25% of the original budget, where $C = .0125$. Therefore, our LA-based method showed more robustness when a more strict budget was used.

B.4.3 *Twitter-COVID19*

As in *Twitter15* and *Twitter16*, the Hawkes simulation performance for *Twitter-COVID19* was measured according to Equation B.15. However, we observed more density in the events arrivals timestamps. Therefore, we set $\Delta = 10$ minutes for a more convenient simulation. Figure B.10 explains the HP performance on both irrational content and valid content, respectively.

The *Twitter-COVID19* dataset is also an interesting case for this study, since it has only one user who was spreading the undesired content and at the same time such user had the highest node centrality degree. That is, in such special situation, we would like to evaluate how our method was independent from the graph structure.

For the simulation, we set the decay factors $w = .7$ and $w = 1$ for the irrational and rational content, respectively, while estimating such values following the same technique as in [10]. With a budget of $C = .05$, Figure B.11 shows how most of the LA-based methods outperformed the three baselines. However, it became obvious that RW_{RP} is the most reliable learning scheme for our optimization framework,

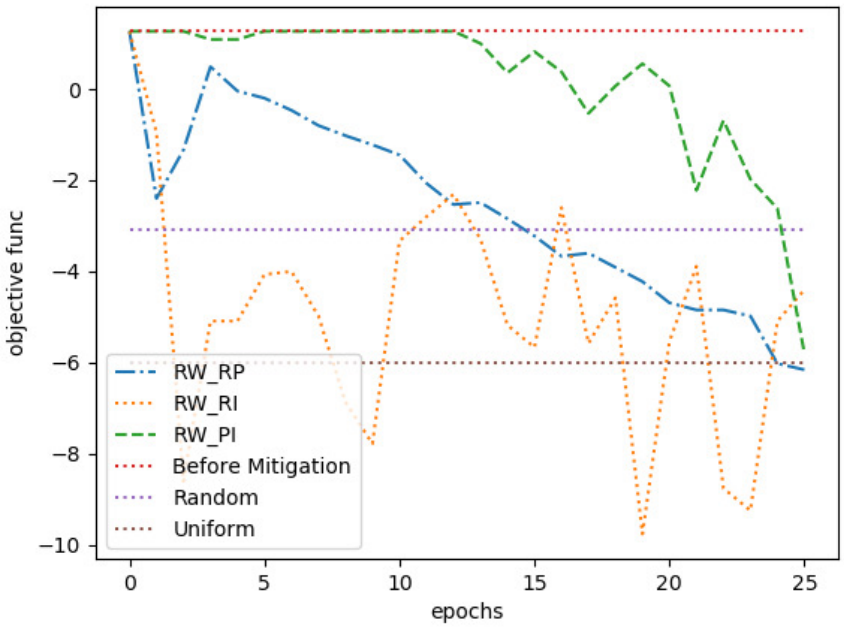


Figure B.7: *Twitter16* mitigation performance on 1'st time stage, $C = .05$

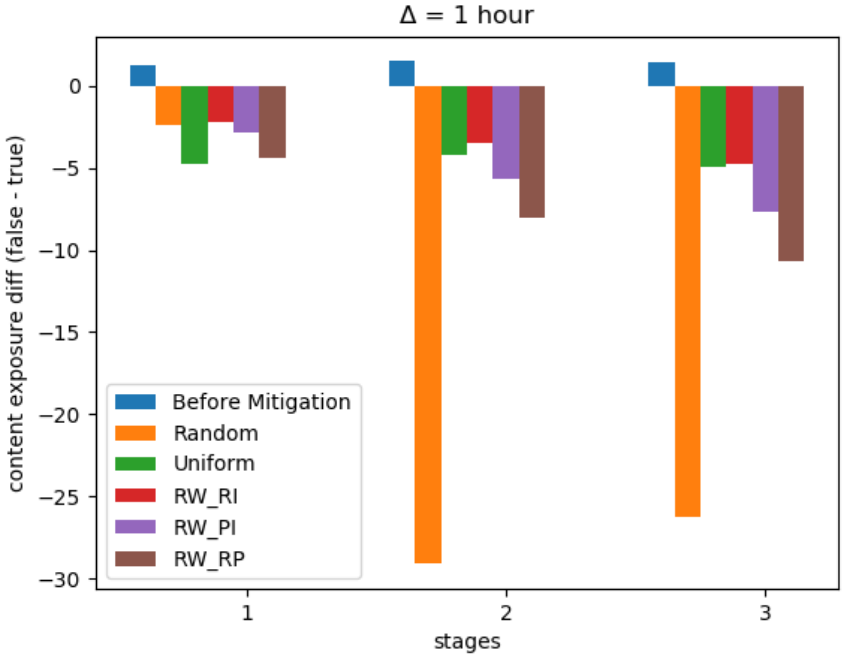


Figure B.8: *Twitter16* mitigation performance on first three time stages, $C = .05$

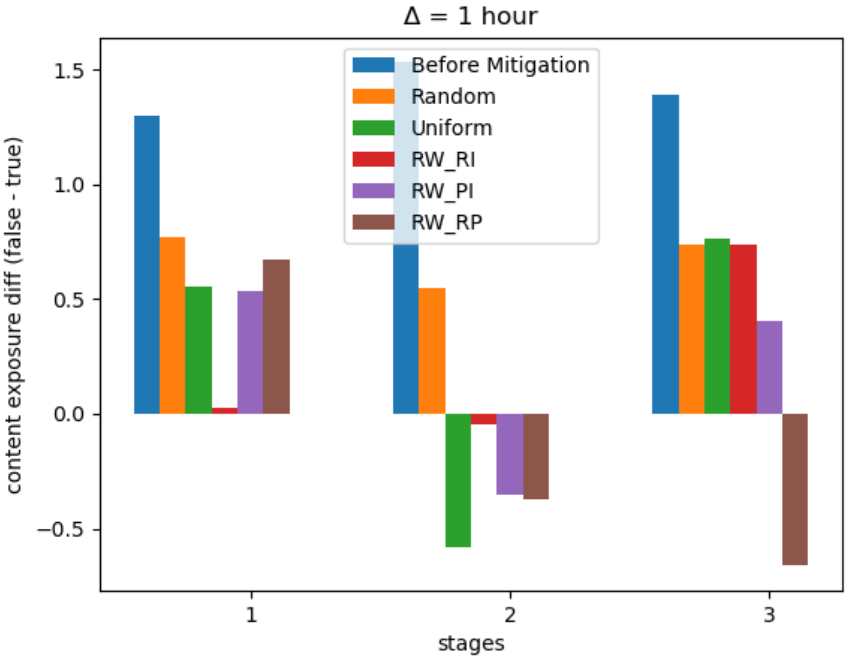


Figure B.9: *Twitter16* mitigation performance on first three time stages, $C = .0125$

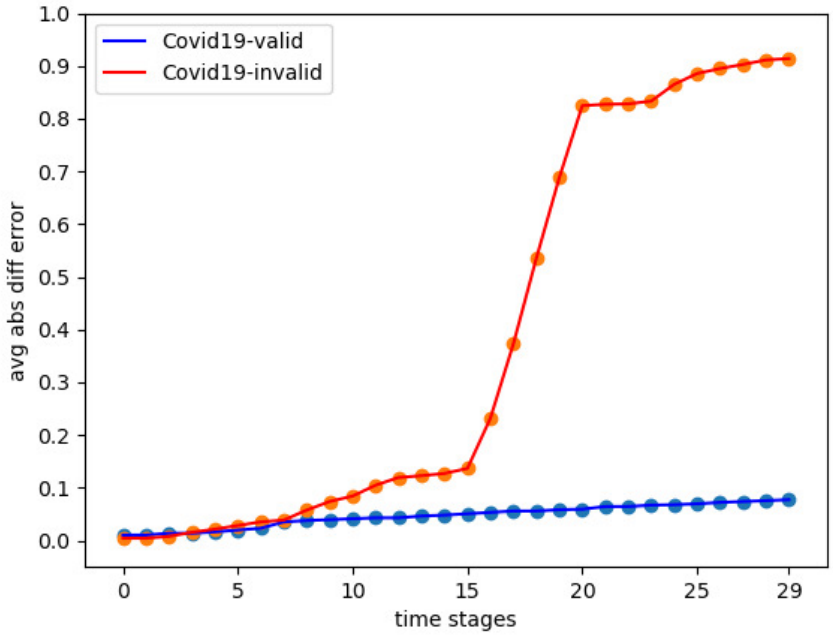


Figure B.10: MHP simulation vs real data on scale of 1 less or more tweet per user difference on a time stage

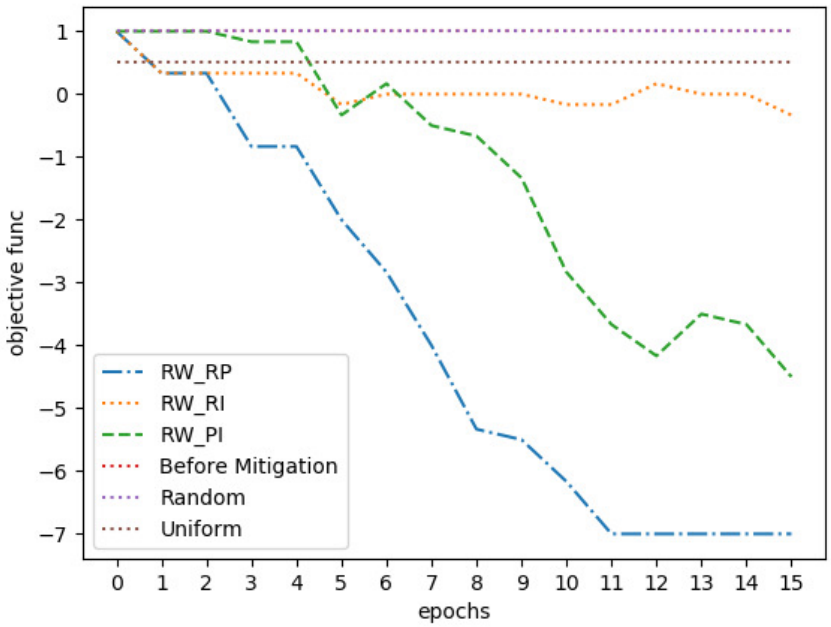


Figure B.11: *Twitter-COVID19* mitigation performance on 1'st time stage, $C = .05$

since it converged earlier with better results in most of our experiments.

B.4.4 Grid-search results

Our method is dependant on three hyper parameters, the learning rate η, the random walk line (LA states) memory depth M, and the sample size $|U^-|$. Therefore, we conducted a grid search to determine their best values. We evaluated the grid search results with respect to how different values of these parameters decreased the asymptotic error for convergence, and increased the risk reduction metric $\mathcal{K}$. The latter is discussed in details in section B.5. Figure B.12 shows how we obtained a *Q-superlinear* convergence on the three datasets from the final estimated hyper parameters and the first MHP time stage and with budget $C = .05$. Table B.3, Table B.4, Table B.5 gives a detailed explanation for the performance of different grid search hyper parameters values over all datasets for the first time stage and with budget $C = .05$. The selection criteria was mainly how much an acceptable risk reduction was achieved with the least possible epochs $\mathcal{I}$. Nevertheless, we considered values which had less effect on the computation speed, since even one epoch might be slower than another while using different values for $|U^-|$, and η. That is because the number of calculations inside one epoch will increase when the network size increases. On the other hand, the learning rate parameter controls the density of the MHP generated events, since higher values of incentivization would lead to more generated counts, which would also increase the number of calculation steps. That is, we approached as much higher $\mathcal{K}$ and lower $\mathcal{I}$, while using as much lower values for $|U^-|$, and η.

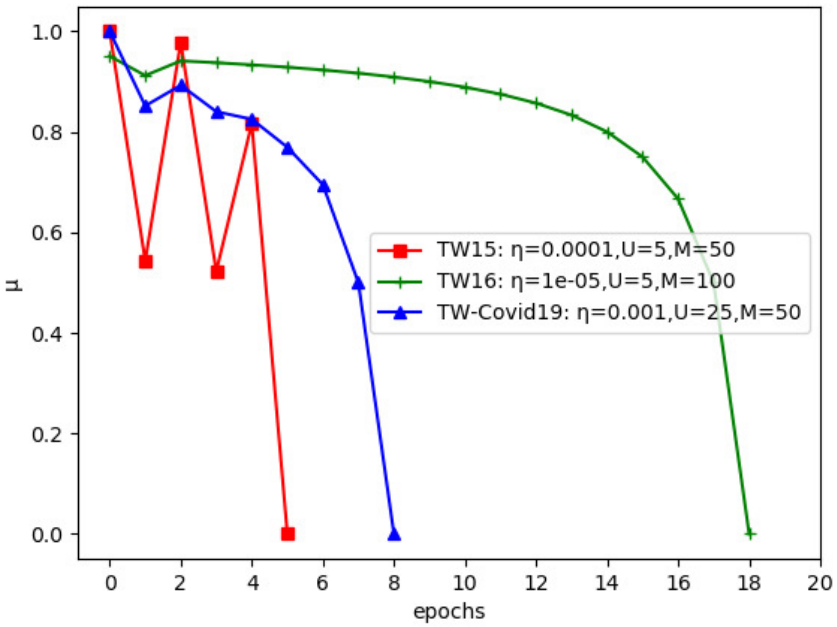


Figure B.12: Convergence plot for the three datasets for $T = t_0, C = .05$

Table B.3: *Twitter15* grid-search hyper parameters for $T = t_0, C = .05$

$\mathcal{I}$	$\mathcal{K}$	$\lvert U^- \rvert$	η	M
≈ 5	$\approx [125\%, 131\%]$	5	.0001	50
≈ 2	$\approx [125\%, 131\%]$	5	.0001	100
≈ 38	$\approx [106\%, 114\%]$	25	.0001	50
≈ 47	$\approx [119\%, 128\%]$	25	.0001	100
> 60	$\approx [100\%, 103\%]$	50	.0001	50
> 60	$\approx [119\%, 128\%]$	50	.0001	100
≈ 25	$\approx [73\%, 79\%]$	5	.00001	50
≈ 19	$\approx [119\%, 128\%]$	5	.00001	100
≈ 11	$\approx [119\%, 128\%]$	25	.00001	50
≈ 13	$\approx [119\%, 128\%]$	25	.00001	100
≈ 18	$\approx [125\%, 135\%]$	50	.00001	50
≈ 34	$\approx [118\%, 125\%]$	50	.00001	100
$N.A$	$[0\%, 0\%]$	5	.000001	50
$N.A$	$[0\%, 0\%]$	5	.000001	100
≈ 12	$\approx [120\%, 129\%]$	25	.000001	50
≈ 13	$\approx [120\%, 129\%]$	25	.000001	100
≈ 48	$\approx [73\%, 79\%]$	50	.000001	50
≈ 25	$\approx [126\%, 135\%]$	50	.000001	100

Table B.4: *Twitter16* grid-search hyper parameters for $T = t_0, C = .05$

$\mathcal{I}$	$\mathcal{K}$	$\lvert U^- \rvert$	η	M
≈ 4	$\approx [227\%, 241\%]$	5	.0001	50
≈ 5	$\approx [235\%, 253\%]$	5	.0001	100
≈ 30	$\approx [235\%, 253\%]$	25	.0001	50
≈ 30	$\approx [119\%, 128\%]$	25	.0001	100
> 60	$\approx [190\%, 209\%]$	50	.0001	50
> 60	$\approx [184\%, 200\%]$	50	.0001	100
≈ 10	$\approx [219\%, 235\%]$	5	.00001	50
≈ 19	$\approx [250\%, 276\%]$	5	.00001	100
≈ 12	$\approx [168\%, 183\%]$	25	.00001	50
≈ 12	$\approx [242\%, 265\%]$	25	.00001	100
≈ 50	$\approx [506\%, 550\%]$	50	.00001	50
≈ 57	$\approx [242\%, 265\%]$	50	.00001	100
≈ 24	$\approx [243\%, 266\%]$	5	.000001	50
> 60	$\approx [242\%, 265\%]$	5	.000001	100
≈ 24	$\approx [242\%, 265\%]$	25	.000001	50
≈ 13	$\approx [239\%, 260\%]$	25	.000001	100
≈ 48	$\approx [138\%, 150\%]$	50	.000001	50
≈ 18	$\approx [199\%, 210\%]$	50	.000001	100

Table B.5: *Twitter-Covid19* grid-search hyper parameters for $T = t_0, C = .05$

$\mathcal{I}$	$\mathcal{K}$	$\lvert U^- \rvert$	η	M
≈ 21	$\approx [330\%, 340\%]$	5	.001	50
≈ 29	$\approx [329\%, 339\%]$	5	.001	100
≈ 10	$\approx [330\%, 340\%]$	25	.001	50
≈ 17	$\approx [327\%, 338\%]$	25	.001	100
≈ 15	$\approx [331\%, 342\%]$	50	.001	50
≈ 19	$\approx [300\%, 310\%]$	50	.001	100
$N.A$	$[0\%, 0\%]$	5	.0001	50
$N.A$	$[0\%, 0\%]$	5	.0001	100
$N.A$	$[0\%, 0\%]$	25	.0001	50
$N.A$	$[0\%, 0\%]$	25	.0001	100
$N.A$	$[0\%, 0\%]$	50	.0001	50
$N.A$	$[0\%, 0\%]$	50	.0001	100
$N.A$	$[0\%, 0\%]$	5	.00001	50
$N.A$	$[0\%, 0\%]$	5	.00001	100
$N.A$	$[0\%, 0\%]$	25	.00001	50
$N.A$	$[0\%, 0\%]$	25	.00001	100
$N.A$	$[0\%, 0\%]$	50	.00001	50
$N.A$	$[0\%, 0\%]$	50	.00001	100

The learning rate value had also some other effects on the results, for example,

Twitter-Covid19 dataset needed a higher value of η to make its users start to respond. That can be seen in Table B.5, where $\mathcal{K} = 0$ in most of the experiments. Eventually, the memory depth parameter M was also essential to the computation, since it controls how fast the knapsack became full, which in turn, could affect the required number of epochs $\mathcal{I}$.

B.5 Discussion

As discussed in section B.4, the random count range $N^t_{\mathcal{E}^\pm}$ defines an estimate range for the random variable N^t, considering the random output and the average absolute difference error $\mathcal{E}$ that affects the mitigation metrics such as the difference between valid and invalid information. Therefore, the risk reduction percentage $\mathcal{K}$ can be calculated with regard to Equation B.3, Equation B.4 as per the one below. Where $\mathcal{E}$ is the sum of errors for *misinformation* and true events simulated counts, Table B.6 shows the summed error for the datasets.

$$\mathcal{K}^{t_s} = \left[(1-\mathcal{E}) \cdot \left(\frac{\sqrt{[(F_1 - T_1) - (F_2 - T_2)]^2}}{(F_1 - T_1)}\right) \cdot \left(\frac{1}{2}\right), \left(\frac{\sqrt{[(F_1 - T_1) - (F_2 - T_2)]^2}}{(F_1 - T_1)}\right) \cdot \left(\frac{1}{2}\right)\right]^{t_s} \tag{B.16}$$

$$, \text{where } F_1 - T_1 \neq 0,$$

Table B.6: Datasets simulation accumulated error $\mathcal{E}$

	t_0	t_1	t_2
Twitter15	.076	.076	.140
Twitter16	.083	.083	.163
Twitter-COVID19	.028	.028	.037

where at a specific time stage t_s, $F_1 - T_1$ represents the difference between false and true content before mitigation, and $F_2 - T_2$ is the difference after mitigation. Then, $\mathcal{K}^{t_s}$ indicates the estimates where such uncertain output would be, by measuring the distance between the two points $(F_1 - T_1)$, $(F_2 - T_2)$ and creating a range between the weighted and the original calculation results. The estimated range values are divided by 2 in order to consider the output $(F_2 - T_2 = 0)$ as only a 50% reduction, since there is still a 50% chance of being exposed to *misinformation* on the network. For cases when $(F_1 - T_1 = 0)$, we omit it as the denominator, and we omit the division by 2 as well.

We have investigated how each learning scheme performed on the three datasets from the perspective of independence from network centrality measures. For instance, when $T = t_0$, we obtained $\mathcal{P}(u_{influencer} \in U^* = 1)$, $\mathcal{P}(u_{influencer} \in U^* = .93)$, and $\mathcal{P}(u_{influencer} \in U^* = .71)$ for *Twitter15* for the three learning schemes RW_{RP}, RW_{RI}, and RW_{PI}, respectively. On the other hand, for *Twitter16*, we got $\mathcal{P}(u_{influencer} \in U^* = .50)$, $\mathcal{P}(u_{influencer} \in U^* = 1)$, and $\mathcal{P}(u_{influencer} \in U^* = .57)$ for

the three learning schemes RW_{RP}, RW_{RI}, and RW_{PI}, when $T = t_0$, respectively. While in *Twitter-Covid19*, we obtained $\mathcal{P}(u_{influencer} \in U^* = .19)$, $\mathcal{P}(u_{influencer} \in U^* = .20)$, and $\mathcal{P}(u_{influencer} \in U^* = .19)$ for the three learning schemes RW_{RP}, RW_{RI}, and RW_{PI}, when $T = t_0$, respectively. Since, the latter dataset is the case where the top influencer user had contributed only to *misinformation*, while other datasets had top influencers who contributed to true content, we would consider our method showing independence from graph structure when the top influencer user is not showing any potentials for circulating valid content.

B.5.1 Evaluation

As mentioned earlier, we have adopted our own post-processed version of the *Twitter15* and *Twitter16* datasets. Further, it was unfeasible to apply all previous baseline mitigation methods on the same data samples we used. However, on the different datasets versions, Table B.7 demonstrates by how far our LA-based method outperformed random and uniform budget distribution methods, with an analogy to previous Reinforcement Learning (RL)-based mitigation methods (MHP-U [10], V-MHP [5], EXP [5]. We refer to our method as LA-MHP, and the evaluation metric is the ratio between the correlation maximization Y^{t_s} at a given time stage for each baseline and either random or uniform maximized correlation, when applied on the associated dataset version. Where the exposure amounts of both fake and true content are considered the correlation variable and constant, respectively ($Y^{t_s} = T^{t_s}F^{t_s}$). For instance, the ratio that indicates how LA-MHP performed against random distribution with regard to correlation maximization is calculated as $\frac{LA-MHP_Y}{RND_Y}$, where Y is calculated twice for both LA-MHP and RND over their MHP generated amount T. The results given in Table B.7 proves how the LA-MHP model is competing with all baselines.

Table B.7: Relative performance against random and uniform methods

Model	***Tw15*-RND**	***Tw15*-UNIF**	***Tw16*-RND**	***Tw16*-UNIF**
LA-MHP	2.37	2.11	2.35	1.71
MHP-U	2.06	1.93	3.20	1.80
V-MHP	1.54	1.87	2.80	1.50
EXP	1.33	1.21	2.04	1.12

B.6 Conclusion

The emergence of the MHP and their application on SM, have boosted the capabilities of SM analysis domain. Hence, MHP-based models became crucial to understand information diffusion and users actions prediction on social networks. Moreover, SM intervention-based approaches are highly depending on MHP to evaluate and improve the developed methods. MHP can be applied to mimic the users

future behaviour on social networks after learning from some past actions on the network. Furthermore, MHP analyze the behaviour of users with regard to different factors. First, MHP can model an exogenous factor that causes a user to act. Then, the model takes into consideration the endogenous factors such as the network users historical behaviour. Therefore, the MHP construct powerful SM dynamics simulation models, where studying internal and external network motivations became possible and reliable.

Compared to deep RL approaches, which were adapted in similar previous work, the explanation given in this study showed how our LA-based method is more reliable for a proactive *misinformation* mitigation strategy, since a LA is easier to understand and implement. Besides, our demonstrated method converged faster while using a notable smaller sample size, compared to the number of samples needed in similar previous work. Furthermore, and to the best of our knowledge, we were the first to apply LAs for *misinformation* on SM.

Future work would investigate how politically biased users might not respond to a mitigation campaign, which will waste the incentivization budget. Furthermore, different objective functions should be investigated, for instance, in certain scenarios, we should consider fair mitigation for the influenced users, instead of calculating the average of differences between fake and true content, since an average for skewed individual differences distribution would not be enough to achieve optimum mitigation results.

Bibliography

[1] Samantha Bradshaw and Philip Howard. Troops, trolls and troublemakers: A global inventory of organized social media manipulation. 2017.

[2] Fang Jin, Wei Wang, Liang Zhao, Edward Dougherty, Yang Cao, Chang-Tien Lu, and Naren Ramakrishnan. Misinformation propagation in the age of twitter. *Computer*, (12):90–94, 2014.

[3] Karishma Sharma, Feng Qian, He Jiang, Natali Ruchansky, Ming Zhang, and Yan Liu. Combating fake news: A survey on identification and mitigation techniques. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(3):1–42, 2019.

[4] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*, 19(1):22–36, 2017.

[5] Mehrdad Farajtabar, Jiachen Yang, Xiaojing Ye, Huan Xu, Rakshit Trivedi, Elias Khalil, Shuang Li, Le Song, and Hongyuan Zha. Fake news mitigation via point process based intervention. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 1097–1106. JMLR. org, 2017.

[6] Savvas Zannettou, Tristan Caulfield, William Setzer, Michael Sirivianos, Gianluca Stringhini, and Jeremy Blackburn. Who let the trolls out? towards understanding state-sponsored trolls. In *Proceedings of the 10th acm conference on web science*, pages 353–362, 2019.

[7] Kai Shu, Suhang Wang, and Huan Liu. Beyond news contents: The role of social context for fake news detection. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, pages 312–320, 2019.

[8] Marian-Andrei Rizoiu, Young Lee, Swapnil Mishra, and Lexing Xie. A tutorial on hawkes processes for events in social media. *arXiv preprint arXiv:1708.06401*, 2017.

[9] Patrick J Laub, Thomas Taimre, and Philip K Pollett. Hawkes processes. *arXiv preprint arXiv:1507.02822*, 2015.

[10] Mahak Goindani and Jennifer Neville. Social reinforcement learning to combat fake news spread. In *Proceedings of the Thirty-Fifth Conference on Uncertainty in Artificial Intelligence*, 2019.

[11] Alan G Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971.

[12] Ke Zhou, Hongyuan Zha, and Le Song. Learning social infectivity in sparse low-rank networks using multi-dimensional hawkes processes. In *Artificial Intelligence and Statistics*, pages 641–649, 2013.

[13] Tohru Ozaki. Maximum likelihood estimation of hawkes' self-exciting point processes. *Annals of the Institute of Statistical Mathematics*, 31(1):145–155, 1979.

[14] Isabel Valera, Manuel Gomez-Rodriguez, and Krishna Gummadi. Modeling adoption of competing products and conventions in social media. In *IEEE International Conference on Data Mining (ICDM)*. Citeseer, 2015.

[15] Keith W Ross and Danny HK Tsang. The stochastic knapsack problem. *IEEE Transactions on communications*, 37(7):740–747, 1989.

[16] Sebastian A Rios, Felipe Aguilera, J David Nuñez-Gonzalez, and Manuel Graña. Semantically enhanced network analysis for influencer identification in online social networks. *Neurocomputing*, 326:71–81, 2019.

[17] Quan Fang, Jitao Sang, Changsheng Xu, and Yong Rui. Topic-sensitive influencer mining in interest-based social media networks via hypergraph learning. *IEEE Transactions on Multimedia*, 16(3):796–812, 2014.

[18] Xiaomo Liu, Armineh Nourbakhsh, Quanzhi Li, Rui Fang, and Sameena Shah. Real-time rumor debunking on twitter. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, pages 1867–1870, 2015.

[19] Jing Ma, Wei Gao, Prasenjit Mitra, Sejeong Kwon, Bernard J Jansen, Kam-Fai Wong, and Meeyoung Cha. Detecting rumors from microblogs with recurrent neural networks. 2016.

[20] Jing Ma, Wei Gao, and Kam-Fai Wong. Detect rumors in microblog posts using propagation structure via kernel learning. Association for Computational Linguistics, 2017.

[21] Mandayam AL Thathachar and P Shanti Sastry. Varieties of learning automata: an overview. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 32(6):711–722, 2002.

[22] Hunt Allcott and Matthew Gentzkow. Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2):211–36, 2017.

[23] Qiao Zhang, Shuiyuan Zhang, Jian Dong, Jinhua Xiong, and Xueqi Cheng. Automatic detection of rumor on social network. In *Natural Language Processing and Chinese Computing*, pages 113–122. Springer, 2015.

[24] Chengcheng Shao, Giovanni Luca Ciampaglia, Onur Varol, Kai-Cheng Yang, Alessandro Flammini, and Filippo Menczer. The spread of low-credibility content by social bots. *Nature communications*, 9(1):1–9, 2018.

[25] Ahmed Abouzeid, Ole Christoffer Granmo, Christian Webersik, and Morten Goodwin. Causality-based social media analysis for normal users credibility assessment in a political crisis. In *2019 25th Conference of Open Innovations Association (FRUCT)*, pages 3–14. IEEE, 2019.

[26] Jotham Wasike. Social media ethical issues: role of a librarian. *Library Hi Tech News*, 2013.

[27] Mehrdad Farajtabar, Nan Du, Manuel Gomez Rodriguez, Isabel Valera, Hongyuan Zha, and Le Song. Shaping social activity by incentivizing users. In *Advances in neural information processing systems*, pages 2474–2482, 2014.

[28] Mehrdad Farajtabar, Xiaojing Ye, Sahar Harati, Le Song, and Hongyuan Zha. Multistage campaigning in social networks. In *Advances in Neural Information Processing Systems*, pages 4718–4726, 2016.

[29] Amir Hassan Zadeh and Ramesh Sharda. Modeling brand post popularity dynamics in online social networks. *Decision Support Systems*, 65:59–68, 2014.

[30] Ryota Kobayashi and Renaud Lambiotte. Tideh: Time-dependent hawkes process for predicting retweet dynamics. In *Tenth International AAAI Conference on Web and Social Media*, 2016.

[31] B John Oommen, Anis Yazidi, and Ole-Christoffer Granmo. An adaptive approach to learning the preferences of users in a social network using weak estimators. *Journal of Information Processing Systems*, 8(2):191–212, 2012.

[32] Anis Yazidi, Ole-Christoffer Granmo, and B John Oommen. Service selection in stochastic environments: a learning-automaton based solution. *Applied Intelligence*, 36(3):617–637, 2012.

[33] Ole-Christoffer Granmo, B John Oommen, Svein Arild Myrer, and Morten Goodwin Olsen. Learning automata-based solutions to the nonlinear fractional knapsack problem with applications to optimal resource allocation. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37(1):166–175, 2007.

Appendix C

Paper C

Title:	Modelling Emotion Dynamics in Chatbots with Neural Hawkes Processes
Authors:	Ahmed Abouzeid, Ole-Christoffer Granmo, Morten Goodwin
Affiliation:	University of Agder, Faculty of Engineering and Science, P. O. Box 509, NO-4898 Grimstad, Norway
Conference:	*International Conference on Innovative Techniques and Applications of Artificial Intelligence*, Cambridge, UK.
DOI:	10.1007/978-3-030-91100-3_12.

Modelling Emotion Dynamics in Chatbots with Neural Hawkes Processes

Ahmed Abouzeid, Ole-Christoffer Granmo, Morten Goodwin
Department of Information and Communication Technology
Faculty of Engineering and Science, University of Agder
P.O. Box 509, NO-4898 Grimstad, Norway
E-mails: {ahmed.abouzeid, ole.granmo, christian.webersik,
morten.goodwin}@uia.no

Abstract **— Conversation partners tend to stick to a particular emotional state unless some external motivation excited them to change that state. Usually, the excitation comes from the other conversation partner. This preliminary study investigates how an Artificial Intelligence (AI) model can provide excitation for the other partner during a dyadic text-based conversation. As a first step, we propose a Neural Emotion Hawkes Process (NEHP) for predicting future emotion dynamics of the other conversation partner. Moreover, we hypothesize that NEHP can facilitate learning of distinguishable consequences of different excitation strategies, and thus it allows for goal-directed excitation behavior by integrating with chatbot agents. We evaluate our preliminary model on two public datasets, each with different emotion taxonomies. Our preliminary results show promising emotion prediction accuracy over future conversation turns. Furthermore, our model captures meaningful excitation without being trained on explicit excitation ground-truths as practiced in earlier studies.**

C.1 Introduction

Nowadays, sequential events are in our every day life, we can observe that in the huge amount of online data being generated at every second. Data generated on Social Media (SM) platforms such as Facebook and Twitter is an example. In such mediums, each user shares a sequence of events though self-opinions and interactions with others [1]. Besides SM, sequential events also exist in domains like regular conversations [2]. In general, sequences are characterized by the particular order of their elements and either their occurrences were temporal, spatial, or both. However, what actually distinguishes the characteristic of a sequence from another is the pattern of its behaviour. For instance, some sequences are synchronous where the temporal or spatial arrivals of events are synced together. On the other hand, some sequences are asynchronous [3] which means the time intervals between event arrivals is as important as their order. In the latter scenario, both self and mutual excitation between events exist [4], and that constructs the dynamics of their behaviour. To this end, point processes [5] were utilized to capture the hidden influence caused by event excitation through the different time intervals.

Traditionally, Poisson point process [4] is used as an example for point processes. However, the complicated dynamics usually found in asynchronous events are beyond the capacity of Poisson processes, due to their event history stateless nature. Hence, Hawkes Processes (HPs) [4] were utilized to capture the historical dependencies between events, which led to more accurate prediction and inference of the hidden excitation. However, some generated events could still be unrelated to each other which is not assumed by the classic HP. That is due to the static computation of its parameters where the latter is estimated from the data before prediction. Therefore, alternative approaches were proposed for a non-parametric HP. First, Recurrent Neural Networks (RNNs) and their variants (e.g., Long-Short Term Memory (LSTM)) were utilized [6, 7] to model the HP dynamics through the network hidden state vector, where historical dependencies were implicitly captured. The advantage of this method is how the process parameters were mutable over different time intervals which is more likely in real-world scenarios. However, due to the limitations [8] of the RNN and LSTM, the neural network struggles in longer sequences with some unrelated dependencies. Hence, a self-attentive HP was proposed [9] to overcome the challenge of long-term dependencies. In the latter approach and unlike RNNs, the self-attention mechanism improved the prediction accuracy by eliminating unrelated dependencies.

According to the definition of *emotional inertia* [10], conversation partners tend to stick to a particular emotional state, unless some external motivation excited them to change that state. Usually, the excitation comes from the other conversation partner. Therefore, by considering one partner as a chatbot agent [11], this preliminary study investigates how to learn hidden excitation patterns in emotional conversations, so the agent can control the outcome from other partner's emotions by re-planning its own (chatbot) expressed emotions.

As a first step, we propose a NEHP for the conversation emotion dynamics

prediction. Moreover, we investigate how the NEHP input conversation emotions should be represented to allow for high prediction accuracy, and successful excitation from further developed controller models on top of the NEHP. For instance, a Reinforcement Learning (RL) chatbot agent that interacts with NEHP as its emotional dynamics environment. To evaluate a potential successful integration with chatbot agents in further studies, we manually re-plan the expressed emotions of one partner in the NEHP emotion sequence input. We define the re-planning of one partner's emotions as the replacement of a certain emotion in its associated conversation utterances sequence. Then, we evaluate if such plan will succeed to excite the other partner for a more positive emotional outcome. We conduct a T-test and accept only significant emotion change outcomes where *P-value* ≤ 0.05. Two public text-based conversational datasets are studied for the preliminary experiments, one is from imagined conversations in movies [12], and the other is sentiment-annotated human-to-human conversations [13].

C.2 Proposed Prediction Model

Figure C.1 shows our proposed End-to-End architecture of the NEHP. The model utilizes a LSTM. We believe the utilization of a LSTM would be sufficient as a preliminary study, where the input sequence length is relatively short and the conversation expressed emotions are completely related, due to the complementary nature of emotions [14] in dialogues. However, self-attention mechanisms are recommended for the analysis of long and multi-context conversations.

NEHP is fed by two categories of inputs: (1) the two partners one-hot-encoding vector R^2 over n steps, where n is the emotion change sequence length, and (2) the emotion change feature representation vector R^n. The core idea behind the proposed sequence prediction model is a Multivariate Hawkes Process (MHP). In a MHP, the excitation between different categories of sequential events can be modelled via a parametric intensity function λ. In our case, we are interested in modelling the asynchronous sequential emotion changes from conversation turns. The term asynchronous is adopted to best describe the importance of the time interval for when an emotion change occurs, e.g., excitation planning. Time intervals in a dyadic conversation setting can be viewed as the turns indices, with the clarification that if one partner interacted with consequent adjacent utterances, all these turns will be considered one single interval. That is, the intervals even indices will be associated with one partner while the odd ones will be representing the other partner. The formal description of a parametric classic MHP is given below in Equation C.1 with its conditional intensity function. The conditional intensity λ predicts the intensity of an emotion change event to occur in a particular interval, given some history of a relevant observation. The prediction captures the mutual excitation between two conversation partners by considering hidden patterns in the given history [4].

$$\lambda(e_i, t_i | H_{t_s}) := \mu(e_i, t_i) + \sum g(e_i - e_s, t_i - t_s) : s < i, \tag{C.1}$$

where g is some kernel function over the history with a decay factor over time,

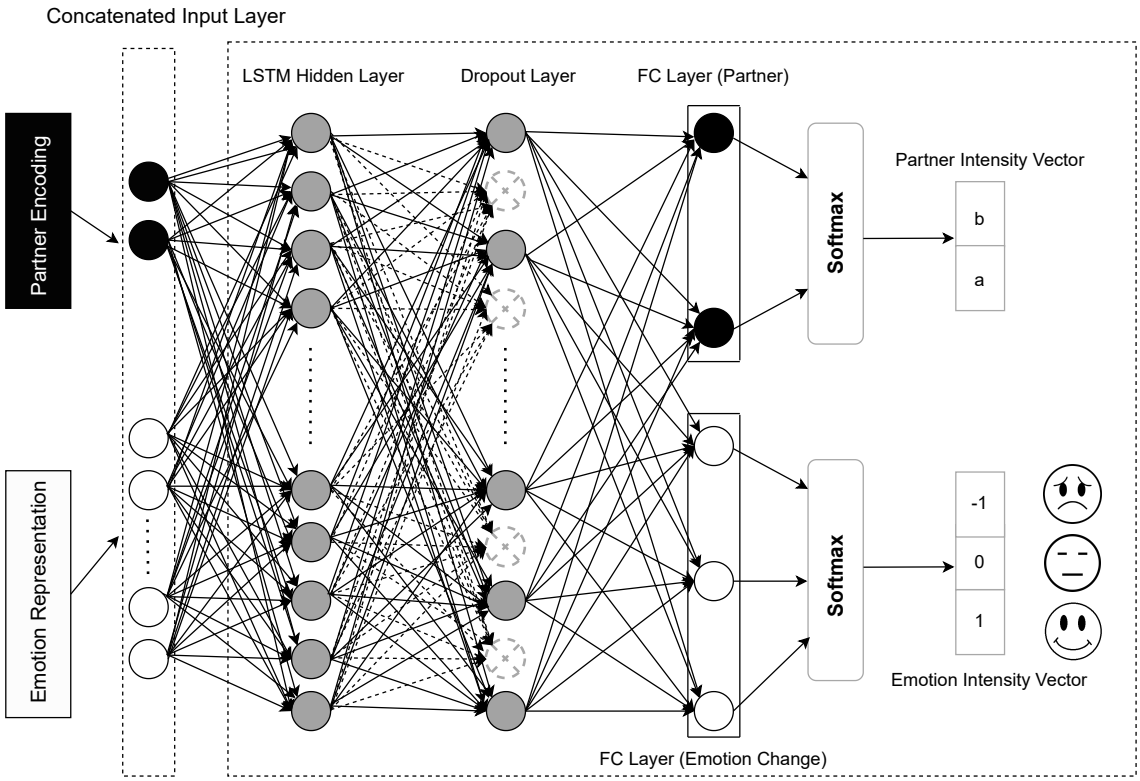


Figure C.1: LSTM-based MHP for dyadic conversation emotion change prediction

and H_{t_s} is the history of emotion change prior to the interval t_i. Equivalently, the index i indicates the current conversation turn index. Typically, g is calculated over an influence matrix A where its entries estimate self and mutual excitation between events, e.g., emotion changes. In a short conversation setting, we can consider no decay of influence since the latter is usually adopted for long temporal sequence analysis and prediction. Moreover, in a neural network setting such as our proposed NEHP, the non-parametric conditional intensities are calculated as per Equation C.2, Where h_t is the hidden state vector from the LSTM network.

$$\lambda(e_{i+1}, t_{i+1}|H_{t_{i+1}}) = \tanh(h_{t_i}). \tag{C.2}$$

C.3 Preliminary Results

Table C.1 shows the preliminary results for our proposed NEHP where four emotion representation techniques were evaluated. Our preliminary results show promising prediction accuracy and potential successful excitation over two future turns of the other conversation partner. The reported excitation results do not involve any RL methods for the control task, but the latter was simplified as a manual modification of the emotion sequence input.

The emotion representation methods vary from each others in terms of how emotion taxonomies scalar values sequence is being prepared as an input to the NEHP. For example, partner difference means that we only focus on the sequence turns

Table C.1: Prediction accuracy of other conversation partner over two turns in future

Dataset	Representation	Acc (1)	Acc (2)	Avg Acc	Excit
Topical	partner difference	0.82	0.51	0.67	**0.16**
Movies	partner difference	0.80	0.35	0.57	0.40
Topical	incremental partner difference	0.89	**0.62**	**0.76**	0.05
Movies	incremental partner difference	0.82	0.20	0.51	**0.45**
Topical	partners difference	0.85	0.54	0.64	0.06
Movies	partners difference	0.81	0.37	0.59	0.24
Topical	incremental partners difference	**0.94**	0.57	**0.76**	0.06
Movies	incremental partners difference	**0.92**	**0.76**	**0.84**	0.16

associated each partner alone and calculate the differences for each. So, we obtain a transformed sequence that represents an emotion change pattern for each partner independently. On the other hand, partners difference means that we calculate the differences from each two adjacent turns emotion scalar values. Since each two adjacent turns in our setting are associated with two different partners, then, the latter technique allows for creating dependencies between partners in the representation. Eventually, the incremental technique allows for accumulated differences, so that it could capture dependencies over time. It is important to highlight that raw emotion scalar values sequence is not a good choice for representing the NEHP input. Since the latter gave a very low prediction accuracy. To reproduce the results, source code and data made available at this link: https://github.com/Ahmed-Abouzeid/rnn_mhp_emotion.

C.4 Conclusion and Future Work

Conversational emotional dynamics has two main properties: (1) self and (2) interpersonal dependencies. We consider such setting is applicable with HPs, where the latter is a point process for modeling mutual and self-excitation in asynchronous events. This preliminary study presented how we could predict emotion changes for n text-based conversation turns in the future. Our NEHP captured hidden excitation of an emotion change point process, and hence could simulate partners' emotion changes as well. Although the NEHP performs a supervised learning for the prediction task, the inferred excitation was not boosted by any supervised learning as practiced in [15]. Therefore, the only given ground-truths were partner ids and their associated emotion change values. The latter capability of the NEHP highlights how the utilization of HPs is beneficial in this domain. Furthermore, the prediction of both partners' emotions means that the NEHP can be either an observer or an actor during a conversation. Observing conversations and predicting their dynamics could be useful in SM analysis to predict the popularity of a post, or the occurrence of critical events in the future. On the other hand, as an actor in a conversation, the NEHP excitation simulation could be considered an environment for a RL-based chatbot, where the latter learns to develop emotional intelligence.

The domain of emotion recognition is evolving rapidly, opening the venue for many research questions. Future work can be applying a self attentive-NEHP model on video recordings-based conversations, where facial and audio features can be extracted. We believe the NEHP could be utilized for other semantic events as well, such as context change in conversations. Moreover, this preliminary study could be a first step towards understanding the behavior patterns of individuals, by studying some semantic shifts in their conversations.

Acknowledgment. This work is part of the Chatbot Interaction Design project, funded by the Norwegian Research Council. The project is a collaboration between SINTEF and the Center for Artificial Intelligence Research (CAIR) in Norway, University of Agder. We would like to thank all colleagues and the project manager Asbjørn Følstad for all the insights and discussions.

Bibliography

[1] Shuang-Hong Yang, Bo Long, Alex Smola, Narayanan Sadagopan, Zhaohui Zheng, and Hongyuan Zha. Like like alike: joint friendship and interest propagation in social networks. In *Proceedings of the 20th international conference on World wide web*, pages 537–546, 2011.

[2] Bram CM Cappers and Jarke J van Wijk. Exploring multivariate event sequences using rules, aggregations, and selections. *IEEE transactions on visualization and computer graphics*, 24(1):532–541, 2017.

[3] Sheldon M Ross, John J Kelly, Roger J Sullivan, William James Perry, Donald Mercer, Ruth M Davis, Thomas Dell Washburn, Earl V Sager, Joseph B Boyce, and Vincent L Bristow. *Stochastic processes*, volume 2. Wiley New York, 1996.

[4] Alan G Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971.

[5] David Roxbee Cox and Valerie Isham. *Point processes*, volume 12. CRC Press, 1980.

[6] Shuai Xiao, Junchi Yan, Xiaokang Yang, Hongyuan Zha, and Stephen Chu. Modeling the intensity function of point process via recurrent neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017.

[7] Nan Du, Hanjun Dai, Rakshit Trivedi, Utkarsh Upadhyay, Manuel Gomez-Rodriguez, and Le Song. Recurrent marked temporal point processes: Embedding event history to vector. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1555–1564, 2016.

[8] Jingyu Zhao, Feiqing Huang, Jia Lv, Yanjie Duan, Zhen Qin, Guodong Li, and Guangjian Tian. Do rnn and lstm have long memory? In *International Conference on Machine Learning*, pages 11365–11375. PMLR, 2020.

[9] Qiang Zhang, Aldo Lipani, Omer Kirnap, and Emine Yilmaz. Self-attentive hawkes process. In *International Conference on Machine Learning*, pages 11183–11193. PMLR, 2020.

[10] Peter Kuppens, Nicholas B Allen, and Lisa B Sheeber. Emotional inertia and psychological maladjustment. *Psychological science*, 21(7):984–991, 2010.

[11] Runnan Li, Zhiyong Wu, Jia Jia, Jingbei Li, Wei Chen, and Helen Meng. Inferring user emotive state changes in realistic human-computer conversational dialogs. In *Proceedings of the 26th ACM international conference on Multimedia*, pages 136–144, 2018.

[12] Cristian Danescu-Niculescu-Mizil and Lillian Lee. Chameleons in imagined conversations: A new approach to understanding coordination of linguistic style in dialogs. *arXiv preprint arXiv:1106.3077*, 2011.

[13] Karthik Gopalakrishnan, Behnam Hedayatnia, Qinglang Chen, Anna Gottardi, Sanjeev Kwatra, Anu Venkatesh, Raefer Gabriel, Dilek Hakkani-Tür, and Amazon Alexa AI. Topical-chat: Towards knowledge-grounded open-domain conversations. In *INTERSPEECH*, pages 1891–1895, 2019.

[14] Jonathan H Turner and Jan E Stets. Sociological theories of human emotions. *Annu. Rev. Sociol.*, 32:25–52, 2006.

[15] Soujanya Poria, Navonil Majumder, Devamanyu Hazarika, Deepanway Ghosal, Rishabh Bhardwaj, Samson Yu Bai Jian, Pengfei Hong, Romila Ghosh, Abhinaba Roy, Niyati Chhaya, et al. Recognizing emotion cause in conversations. *Cognitive Computation*, pages 1–16, 2021.

Appendix D

Paper D

Title:	Socially Fair Mitigation of Misinformation on Social Networks via Constraint Stochastic Optimization
Authors:	Ahmed Abouzeid, Ole-Christoffer Granmo, Christian Webersik, Morten Goodwin
Affiliation:	University of Agder, Faculty of Engineering and Science, P. O. Box 509, NO-4898 Grimstad, Norway
Conference:	*36th AAAI Conference on Artificial Intelligence*, Vancouver, Canada.
DOI:	10.1609/aaai.v36i11.21436.

Socially Fair Mitigation of Misinformation on Social Networks via Constraint Stochastic Optimization

Ahmed Abouzeid, Ole-Christoffer Granmo, Christian Webersik, Morten Goodwin
Department of Information and Communication Technology
Faculty of Engineering and Science, University of Agder
P.O. Box 509, NO-4898 Grimstad, Norway
E-mails: {ahmed.abouzeid, ole.granmo, christian.webersik, morten.goodwin}@uia.no

Abstract **— Recent social networks' misinformation mitigation approaches tend to investigate how to reduce misinformation by considering a whole-network statistical scale. However, unbalanced misinformation exposures among individuals urge to study fair allocation of mitigation resources. Moreover, the network has random dynamics which change over time. Therefore, we introduce a stochastic and non-stationary knapsack problem, and we apply its resolution to mitigate misinformation in social network campaigns. We further propose a generic misinformation mitigation algorithm that is robust to different social networks' misinformation statistics, allowing a promising impact in real-world scenarios. A novel loss function ensures fair mitigation among users. We achieve fairness by intelligently allocating a mitigation incentivization budget to the knapsack, and optimizing the loss function. To this end, a team of Learning Automata (LAs) drives the budget allocation. Each LA is associated with a user and learns to minimize its exposure to misinformation by performing a non-stationary and stochastic walk over its state space. Our results show how our LA-based method is robust and outperforms similar misinformation mitigation methods in how the mitigation is fairly influencing the network users.**

D.1 Introduction

From a computation perspective, there are many approaches to combat the dissemination of misinformation[1]. Recently, [1] illustrated some of the main techniques for classifying misinformation content and how these approaches can be applied in different scenarios. However, classification methods tend to be offline and limited to particular social network features to be learned, such as linguistics and the local political context [2]. Furthermore, such classification models have a potential for *False Positive* matches, which may violate human rights conventions by misjudging and questioning individuals credibility and controlling free speech [3]. On the other hand, recent work proposed intervention-based resolutions as an online approach to mitigate the circulation of misinformation on Social Media (SM) platforms. Such an approach is considered more convenient since it facilitates better collaboration between humans and technology by providing learned misinformation mitigation strategies instead of black-box classification models. For example, [4] proposed a Reinforcement Learning (RL)-based optimization method which provides a strategy to decrease the difference between misinformation and true content exposures in Twitter, given that such misinformation exposure was dominating the network. The purpose was to mitigate the effect of misinformation on network users by incentivizing the latter to spread true information. A similar method was developed to facilitate decentralized and faster computation, as proposed by [5].

The latter approach introduces a light-weight decentralized computation that reduces the optimization sample space and utilizes LAs that learn from reinforcement feedback [6]. However, the method was evaluated according to the decrease in difference between the dominating misinformation and the incentivized true content, averaging over the whole network. The problem with such an approach is that there would be real-world scenarios where some individuals need mitigation efforts more than others, while a sub-network individuals would be already protected from high misinformation exposures. Therefore, we believe a more socially fair intervention and allocation of mitigation resources should be introduced under the framework of [5].

This paper proposes a robust LA-based decentralized mitigation method that addresses a wide range of possible unbalanced exposures to either misinformation or true content, seeking robustness on a variety of a social network's statistics. Our contribution is threefold:

- We propose a novel learning scheme for a LA learning in a stochastic and non-stationary environment. The randomness comes from an information diffusion model based on point processes [7], while the non-stationarity comes from the temporal changes that occur over the whole network when an individual user responds to incentivization. This non-stationarity is particularly intricate due

[1]The term misinformation is sometimes used to refer to all forms of fake news/content. However, in some literature, misinformation is defined as the unintentional spread of false content while disinformation is the on-purpose spread. In this paper, we refer to all forms of false content as misinformation.

to the hidden dependencies in the information diffusion model. The LA task is to construct a network of individual automata on top of the social network. Each individual automaton is associated with a single user and performs a constraint Knapsack optimization via a random walk [8] over the automaton state space.

- We propose a novel loss function to ensure that true content incentivization budget is fairly assigned according to individual users exposure needs. To this end, the problem is defined as a stochastic and non-stationary multi-agent Knapsack [9] optimization problem.
- We introduce two evaluation metrics (Achieved Mitigation and Achieved Fairness) to measure the efficiency and robustness of the proposed misinformation mitigation algorithm on different social network's statistics. And we evaluate how our proposed technique is more socially fair compared to the proposed approach in [5]. We conduct our empirical experiments on both synthetic and real-world social networks. Software source code and data are available here: https://github.com/Ahmed-Abouzeid/MMSS.

D.2 Preliminaries

D.2.1 Information Diffusion Modelling

In order to apply intervention-based resolutions to misinformation mitigation, an information diffusion model is required to simulate the social network which to intervene with. The simulation is considered because intervention with the actual SM platforms is not feasible. We simulate the process of information diffusion by employing a Multivariate Hawkes Process (MHP) as practiced by [10], [5], and [4]. A MHP is a multivariate stochastic process [11] which models the occurrence of temporal or spatio-temporal asynchronous events by capturing the mutual-excitation (dependencies) between these events. To model the social network dynamics, each user is represented by two Hawkes Processes (HPs), one for misinformation dissemination behavior, and the other for true content. The associated user HPs generate estimated random counts for both information types, given some behaviour observation in the past (e.g., estimating number of re/tweeted events given historical dependency). These counts indicate the intensity of the process at a specific time realization. Hence, A HP can be defined with its conditional intensity function λ. The intensity function has two main components: base intensity μ, and exponential decay function g over an adjacency matrix A. The formal explanation of the conditional intensity function is given by:

$$\lambda_i(t_r|H^{t_r}) := \mu_i + \sum_{t_s<t_r} g(t_r - t_s), \tag{D.1}$$

where μ is the base intensity that models some external motivation to propagate some content (independent from inferred relationships in data). On the other hand,

g is some kernel function over the observed history H^{t_r} associated with user i from the discrete time realization t_s prior to time t_r. g is concerned with the history of some influence matrix $A_{i.}$, where $A_{ij} = 1$ if there is an influence indicating that user i influences user j, and $A_{ij} = 0$ if not. We used an exponential decay kernel function $g = A_{i.}e^{-wt}$ as practiced by [4], where w is the decay factor which represents the rate for how the influence is reduced over time. For all users, the base intensity vector μ, and the influence matrix A can be estimated using maximum likelihood as proposed in [12]. To simulate all users behaviours for each content type, a MHP is created, given that different intensity rates are generated at different discrete time realizations. Hence, at each realization, each user behaviour is simulated as an estimated number of events (misinformation or true content) to be generated. We set the interval window between realizations to two hours. The HP simulation algorithm adopted in this study follows the modified thinning algorithm introduced by [13]. See Appendix A.1 [14] for a detailed explanation of the simulation evaluation metric.

D.2.2 Mitigated Diffusion

The core idea behind misinformation mitigation is by introducing the true information to the network through incentivization. Therefore, users associated true content HPs are modified. Hence, let x_i be the incentivization amount decided for user i, and the modified HP for mitigation purposes can be redefined by:

$$\lambda_i(t_r|H^{t_r}) := x_i + \mu_i + \sum_{t_s<t_r} g(t_r - t_s). \tag{D.2}$$

D

D.3 Related Work

D.3.1 Misinformation Impact

According to [15], at least 50% of the world's countries suffer from organized political manipulation campaigns over SM. Other examples of misinformation can be observed during the Ebola outbreak in West Africa, which was believed to be three times more worse than the previous Ebola outbreaks [16]. Therefore, research on the role of online media and border-free passing through messages became an emerging topic of interest in scientific communities. Furthermore, investigation on such a topic is more complicated and requires different perspectives of analysis. For example, recent studies [17] argued that the influence of SM on accepting political misinformation may differ depending on age, culture or gender. Such social studies actively investigated the social impact of misinformation propagation on different SM platforms such as Reddit, Facebook, and Twitter. Novel views on the problem emerged recently. For instance, recent investigations reported that deliberation contexts promoted in SM overcome false information about health [18]. An example of such deliberation can be viewed as a counterfactual campaigns to spread true health

information against the spread of misinformation as practiced for the COVID-19 case on Twitter by [5].

D.3.2 Misinformation Detection

The spread of fake news on SM has been initially considered as the intentional dissemination of false content in news articles [19]. Progressively, others gave attention to the broader range of the problem [20, 21]. Moreover, rumor detection [22], malicious accounts classification [23, 24], and the causal aspects of misinformation [25] have been discussed. However, the majority of these methods are highly depending on linguistic or local features which cause a lack of generality in the final resolution. To the best of our knowledge, it is hard to solve the problem in real-time or without data selection-bias concerns [26]. [27] expressed similar moral concerns since fake news detection resolutions are judgemental by nature. Therefore, the need for safer online strategies that would lead to more generic and authentic resolutions is critically desirable.

D.3.3 Knapsack Optimization

The utilization of LA with Knapsack optimization problems is widely approached in the literature. For instance, [28] worked on optimizing the allocation of polling resources for web page monitoring when the monitoring capacity is restricted. In web page monitoring systems, the system may involve n web pages that are updated on different time intervals. Hence, to avoid involving all web pages including the ones with no updates, the system must determine the most important web pages only, without exceeding the monitoring capacity. The work utilized a team of LAs, where each automaton is involved with a particular web page and learns its importance to a Knapsack total value. Similarly, [29] dealt with a Stochastic Non-linear Fractional Equality Knapsack problem which is a fundamental resource allocation problem based on incomplete and noisy information. In the latter work, they proposed an optimal resolution to the resource allocation problem using a continuous LA without mapping the Knapsack materials onto a binary hierarchy. In such work, the proposed LA had a Reward-Inaction (R-I) learning scheme which only updates the LA actions (transitions) probabilities when rewarded. [30] worked on another combinatorial optimization problem for Knapsack with a proposed Migrating Birds Optimization algorithm to solve a 0-1 knapsack problem [31].

D.3.4 Hawkes Processes

The utilization of HPs-based intervention strategies was effectively presented on minimizing-risk problems. For example, [32] worked on the problem of invasive species spreading to new areas which threatens the stability of ecosystems and causes major economic losses. The latter study proposed a novel approach to minimize the spread of an invasive species given a limited intervention budget, where the spread of

species was modelled by a HP and the minimization task was considered a constraint Knapsack optimization problem.

D.4 Methodology

D.4.1 Learning Automata Network

A LA is a stochastic model suitable for learning in random environments [6]. The LA learns by interacting with the random environment, and updates its actions or state transitions according to the stochastic signal from the environment. Depending on the automaton design and architecture, the task is to find either an optimum/sub-optimum action or state. The LA seeks convergence to such state or action, eventually. The advantage of utilizing a LA-based optimization is due to its decentralized and easy implementation. A LA defined by its stochastic state transitions can be formally defined as a Markov Process [33]. Therefore, to reach equilibrium over all LAs, we build a network of LAs, each performs a random walk over a finite and discrete state space, where the individual optimum or sub-optimum states will be the recommended incentivization values for a misinformation mitigation campaign. The individual random walks together form as a multidimensional joint random walk [34] modelled by a multivariate Markov chain [35]. Figure D.1 demonstrates the proposed LAs network and the underlying multivariate Markov chain (e.g., three automata with M states, each.), where the joint state transitions and their probabilities are derived by the individual automata state transitions which are dictated by a reward signal β.

D.4.2 Learning State Transition

An individual LA_i has a state space with memory depth M, where $M > 0$. If LA_i is in a state S_i^k where $0 < k < M$, then, it has a three possible state transitions: $S_i^{k,k-1}, S_i^{k,k}, S_i^{k,k+1}$ indicating going to left, staying at same state, and moving to the right, respectively. In order to reach an optimum or sub-optimum state S_i^*, LA_i needs to learn the probabilities of its state transitions until it converges. Consequently, the optimum or sub-optimum S_i^* value will be the recommended incentivization value x_i^* to modify the information diffusion model with (See Equation 2). LA_i could have only two possible state transitions: $S_i^{k,k}, S_i^{k,k+1}$ or $S_i^{k,k}, S_i^{k,k-1}$, when $k = 0$ or $k = M$, respectively. At each interaction step t, the probability of LA_i being in a next state depends on its present state and the transition direction a_i^t. With a uniform initial state transitions probabilities, LA_i determines the next state S_i^{t+1} and updates its state transition probability distribution vector π_i according to the below:

$$\delta : S_i^t, a_i^t, \beta_i^t \rightarrow S_i^{t+1}, \pi_i^{t+1}, \tag{D.3}$$

where π_i states probabilities are updated with regard to their rewarded visits frequency, and a_i^t represents the applied state transition $a_i^t = S_i^{k,j}$, where k, j are

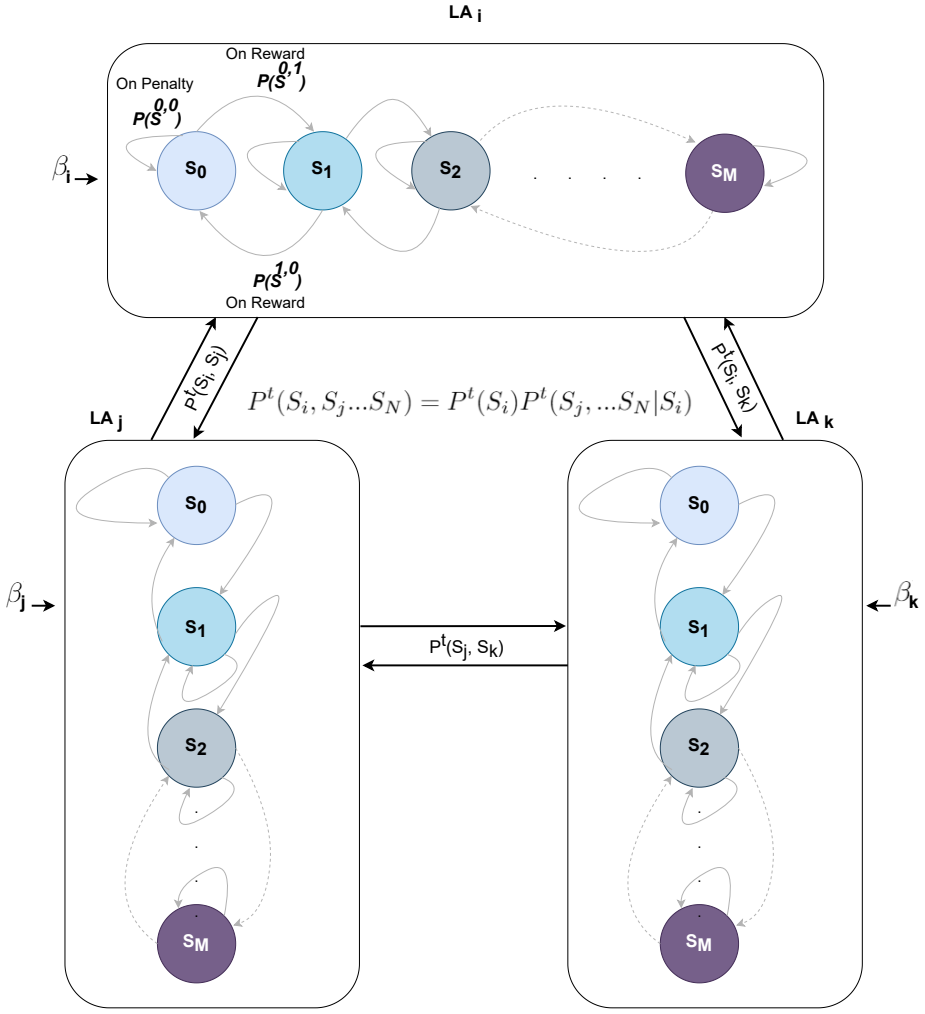


Figure D.1: The proposed LAs network and the underlying multivariate Markov chain architecture for three automata

neighbor state indices and $k = j$ if it was a recurrent state transition. Based on a_i^t and the environment stochastic reward β_i^t, LA_i conducts a random step move over its state space. For instance, if $a_i^{t_0} = S_i^{k,k+1}$, the state transition function δ commits the transition $S_i^k \rightarrow S_i^{k+1}$ only if $\beta_i^{t_0} = 0$, and rolls it back if $\beta_i^{t_0} = 1$. Consequently, $\pi_i^{t_1} = [0^{k,k}, 1^{k,k+1}, ...]^{t_1}$ or $\pi_i^{t_1} = [1^{k,k}, 0^{k,k+1}, ...]^{t_1}$, respectively. We denote v_i^t and w_i^t as how many times a transition was rewarded ($\beta_i = 0$) and performed for LA_i up to interaction step t, respectively. Hence, For the state indices k, j, when $k = j + 1$, state transition probabilities are updated as the below:

$$P^{t+1}(S_i^{k,j}) = \frac{v_i^t(S_i^{k,j})}{w_i^t(S_i^{k,j})}, \tag{D.4}$$

$$P^{t+1}(S_i^{j,k}) = \frac{1 - P^{t+1}(S_i^{k,j})}{2}, \tag{D.5}$$

$$P^{t+1}(S_i^{k,k}) = \frac{1 - P^{t+1}(S_i^{k,j})}{2}, \tag{D.6}$$

$$\text{where } P^{t+1}(S_i^{k,j}) + P^{t+1}(S_i^{j,k}) + P^{t+1}(S_i^{k,k}) = 1. \tag{D.7}$$

Since each LA performs a random walk over its state space through a stochastic state transition, then the optimization problem is solved by the multidimensional joint random walk over the automata network. Furthermore, the individual state transitions are dependent to each others due to the shared knapsack capacity and the inter-connected influence in their environment rewards. Therefore, the probability of a particular automata network state is calculated as the joint probability of the individual automata current states. Hence the joint probability can be calculated as the below, where N is the network size:

$$P^t(S_i, S_j ... S_N) = P^t(S_i) P^t(S_j, ... S_N | S_i). \tag{D.8}$$

D.4.3 Automaton Environment

To learn incentivization values for the social network's users, all users' associated LA interact with a Knapsack which evaluates how valuable the current LA state (incentive) for the mitigation campaign. The Knapsack evaluation is individual to each user behaviour on the network. Users behaviours are modeled through a MHP. Hence, the LA environment has the following main properties.

- **Stochastic:** which is due to the randomness of each HP itself, which generates random counts for each user events (e.g., re/tweets).
- **Non-stationary:** which occurs because of the dependencies between users HP generated events. For instance, when both users i, j have an explicit or implicit dependency, a particular incentivization value $x_i = 0$ might not be optimum for user i but could be optimum when the incentivization value $x_j > 0$. Since the latter could cause user i to be fairly exposed to true content without the need to increase for x_i (incentivize user i).

To reinforce the learning of targeted state values. each individual LA_i will receive a reward signal β_i from its Knapsack environment where $\beta_i \in \{1, 0\}$, indicating a penalty, or reward Knapsack signal, respectively. The final committed state transition for an LA_i is driven by the reward signal β_i. For instance, if LA_i randomly walks towards the right and received a reward, it commits the transition and updates its current state. However, if LA_i receives a penalty, it rolls back the transition and stays at its recent current state before that transition. The state update mechanism also works if LA_i randomly walks to the left direction. These random walks probabilities in both directions are learned according to Equation D.4 and Equation D.5. On the other hand, recurrent state transitions probabilities are updated according to Equation D.6 until converging to a state where the probabilities of performing random walks in both directions became almost 0. The detailed information about how the reward signal β_i is calculated for each direction of an LA_i random walk:

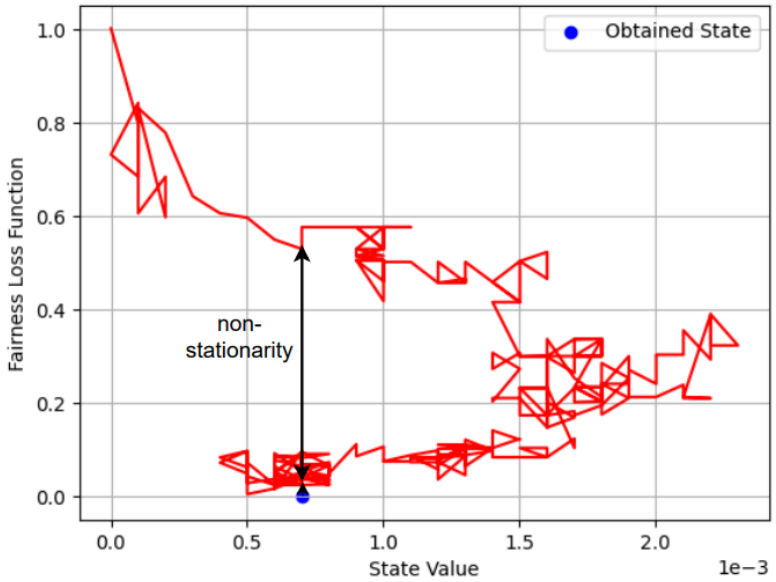


Figure D.2: Finding global minima example for an individual LA random walk over a stochastic and non-stationary HP-based Knapsack response

$$(\rightarrow)\ \beta_i(m_i, \Phi) := \left\{ \begin{array}{ll} 1, & \text{if } m_i > 0 \vee \Phi = 1 \\ 0, & \text{otherwise} \end{array} \right\}, \tag{D.9}$$

$$(\leftarrow)\ \beta_i(m_i, \Phi) := \left\{ \begin{array}{ll} 1, & \text{if } m_i > 0 \\ 0, & \text{otherwise} \end{array} \right\}, \tag{D.10}$$

$$\text{subject to } m_i = \frac{\Delta \mathcal{F}(x_i)}{\Delta x_i}, \text{ where } \Delta x_i > 0, \tag{D.11}$$

where m_i is the slope of a fairness loss function $\mathcal{F}$ for the associated user i and Φ indicates either the Knapsack is currently full ($\Phi = 1$) or not ($\Phi = 0$). The Knapsack initial capacity starts with 0 and increased or decreased according to each individual LA state transition, while the current Knapsack capacity is shared across the LA network. Given that $x_i = S_i : i \leq M$, since the mitigation incentive x_i over time is represented by the current LA state where such an LA has M states. The above definition of the environment reward for the proposed random walk state transitions ensures converging to optimum or sub-optimum mitigation incentive values. Figure D.2 shows an example of our proposed LA state transitions mechanism where the optimization environment is non-stationary and stochastic. However, the LA managed to find a sub-optimal state value.

D.4.4 Fairness Loss Function

To achieve fair mitigation, we need to consider each individual user exposures to both misinformation and true content. Each user exposure associated with a content type is calculated as how much impact that content has on the user. Therefore, the ratio between true and misinformation impact for each user is considered. Hence, a more

skewed initial distribution of these ratios will acquire a fair mitigation strategy to assign the incentivization budget according to user needs, without wasting the budget on users with already high exposures to true content. During the intervention, a ratio $R_i < 1$ means that user i is more exposed to misinformation. Alternatively, a ratio $R_i > 1$ indicates that user i incentivization is not necessary since the latter has already high level of true content exposures. The exposure values used in R_i were calculated as proposed by [163], see Appendix A.2 [14] for more details. Below, we define our proposed fair misinformation mitigation loss function:

$$\min \mathcal{F}(X) := \sum_{i}^{N} \mathcal{F}(x_i), \text{ where } \mathcal{F}(x_i) := \sum_{j=0}^{n} (1 - R_j^{x_i})^2, \tag{D.12}$$

$$\text{subject to } \sum_{i=1}^{N} x_i, \text{ where } x_i \in [0, C], \tag{D.13}$$

where N represents the number of network users and n is the number of adjacent users connected to user i, where user i is considered adjacent to itself. Therefore, j is the index represents i and all its adjacent over the summation. $R_j^{x_i}$ represents the updated ratio between true content and misinformation after applying the incentivization value x_i to the true content HP diffusion model associated with user i. As noticed in Equation D.12, we square the subtraction $1 - R_j^{x_i}$ to maintain positive values in the interval $[0, \infty)$, while the task is to minimize the loss function as much closer to 0 as possible (See Figure D.2. It is important to highlight that the total loss is calculated through the achieved individual loss of each user during the allocation of incentives (e.g., associated LA and its current state value). That means the total loss ensures optimum or sub-optimum assigned incentivization values over X, where X can be viewed as the set of all automata current states. Eventually, the consumption of all incentivization values (LA states) must not exceed the bound C, which represents the Knapsack capacity.

D.4.5 Misinformation Mitigation

To obtain the optimum or sub-optimum learned states vectors of N automata, we initialize each individual LA_i with an initial state transition probability vector $\pi_i^{t_0}$, and the initial ratio $R^{x_i=0}$ where no incentivization values yet to be added to the associated estimated base intensity $\mu_i^{t_0}$ of the relevant HP. Eventually, the initial fairness loss function $\mathcal{F}_i^{t_0}(x_i = 0)$ is calculated while the Knapsack is initially empty $c^{t_0} = 0$. The mitigation algorithm then iterates over the whole LA network until it converges to all optimum or sub-optimum state probability vectors. Then, converged states values are suggested as incentivization values for the underlying associated users on the network. The details of the misinformation mitigation procedure is shown in Algorithm 1.

Algorithm 1 Fair misinformation mitigation.

Input: $\mu_i^{t_0}, \pi_i^{t_0}, R^{x_i=0}, \mathcal{F}_i^{t_0}(x_i = 0), \forall i : u_i \in U, c^{t_0}$, and N where $|U| = N$.

Output: $S_i^*, \forall i : u_i \in U$, where $|S^*| = N$.

Let $t = 1$.
while $\neg(\pi_{all}^t \leftarrow \pi_{all}^*)$ **do**
  for $i \leftarrow 1$ to N **do**
    if $\pi_i^t \neq \pi_i^*$ **then**
      $a_i^t \leftarrow max[P(S_i^{k,j}), P(S_i^{k,k}), P(S_i^{j,k})]^t$.
      $S_i^t \leftarrow a_i^t$.
      $x_i \leftarrow S_i^t$.
      $\Delta x_i \leftarrow abs(S_i^t - S_i^{t-1})$.
      $\sum_{j=0}^{n}(R_j^{x_i}) \leftarrow \lambda(x_i)$.
      $\mathcal{F}^t(x_i) \leftarrow \sum_{j=0}^{n}(1 - R_j^{x_i})^2$.
      $\Delta F(x_i) \leftarrow abs(F(x_i)^t - F(x_i)^{t-1})$.
      $m_i = \frac{\Delta F(x_i)}{\Delta x_i}$.
      $\beta_i^t \leftarrow \beta_i^t(m_i, \Phi)$.
      $S_i^{t+1}, \pi_i^{t+1} \leftarrow \delta(S_i^t, a_i^t, \beta_i^t)$.
    else
      continue.
    end if
  end for
  $t \leftarrow t + 1$.
end while
return S^*.

D.5 Experimental Setup

In our experiments we design six synthetic social networks $\{syn1, syn2, syn3, ..., syn6\}$. Each with a unique statistical misinformation exposure distribution among users. The six networks represent the possible real-world scenarios where some user groups might be highly exposed to misinformation more than other groups on the social network. Moreover, some individuals in these groups might be also highly exposed to misinformation more than others from the same group. Allowing for these possible scenarios in our experiments should stress the evaluation of robustness for a fair misinformation mitigation resolution. We design our synthetic networks by randomly generate variant true information and misinformation event counts on both user and network levels. Then, we set different bounds on these synthetic exposures to maintain a variety of statistics for each network. Eventually, we run our resolution on a real-world social network used in [5] as another benchmark. The real-world network is a COVID-19 social network and annotated for ordinary and false re/tweets from Twitter on the 28^{th} of March, 2020. The collected re/tweets focused on discussions about COVID-19. The criteria for the misinformation annotation was if any propagated content urged the public for using false drugs [36] without any official statements from the health authorities at that time. Within

Network	Knapsack size	Overall Misinformation
Syn1	0.06	17.00%
Syn2	0.06	58.00%
Syn3	0.06	88.50%
COVID-19	0.06	89.50%
Syn4	0.18	11.75%
Syn5	0.18	47.25%
Syn6	0.18	86.50%
COVID-19	0.18	89.50%

Table D.1: Configuration details of fair misinformation mitigation experiments on the proposed social networks

each of our experiments, we consider different mitigation incentivization budget for the Knapsack capacity to evaluate for different levels of constraints. Due to the randomness of experiments, we run each for multiple times and take the average as an estimate of the final outcome. Table D.1 shows the configuration of our experiments, where all networks have 200 users. For the selection of hyper-parameters values in all experiments, see Appendix A.3 [14].

D.6 Evaluation

D.6.1 Uniform-baseline

To highlight the need for a fair misinformation mitigation method, we make an analogy with a uniform allocation of the incentivization budget. For instance, if all or almost network users are equally exposed to misinformation than true content, a uniform distribution of incentivization budget is theoretically an optimum fair mitigation strategy. We refer to the latter as **Case-0**. However, the more the two content types were unbalanced on the network, the more challenging for a budget uniform distribution to achieve the desired mitigation results. For example if only 20% of network users were exposed to misinformation, a uniform incentivization becomes a waste for 80% of the budget, which might cause no mitigation at all since 20% of the budget becomes insufficient to maintain $R = 1$ for the targeted users. We refer to the latter as **Case-1**. Another form of skewness is when the majority of users are exposed to misinformation but a subset of them are significantly more exposed to misinformation than others, in such scenario, the uniform method will suffer as well, since these subset of users will need more incentivization than others. We refer to the latter as **Case-2**. It is important to highlight that the purpose of the HP information diffusion model is to predict future behaviours. Therefore, the initial distribution of misinformation exposures before any future intervention is unknown, and a robust incentivization is mandatory to overcome all the potential misinformation percentages.

D.6.2 AVG-LA-baseline

We further investigate how our LA-based resolution performs against current existed LA-based methods [5]. We refer to the latter as **AVG-LA**, while we refer to our proposed method as **Fair-LA**.

D.6.3 Mitigation Efficiency

To evaluate for robustness on multiple social networks' scenarios, we introduce a mitigation efficiency metric which is calculated as per the below:

$$1 - \frac{a}{b}, \tag{D.14}$$

where a and b are the misinformation percentages after and before mitigation, respectively. According to our synthetic social networks' different setups (See Table D.1), **Case-1** can be observed in syn1 and syn4, while **Case-2** can be observed in syn3, and syn6. As concluded from Figure D.3, our proposed **Fair-LA** outperforms both **AVG-LA** and **Uniform** methods in most of the scenarios, especially in **Case-1**. Moreover, when **Case-2** occurs, **Fair-LA** still outperforms other methods when the Knapsack capacity C was larger. From our statistical analysis on the COVID-19 network with 200 users, we observed almost a scenario equivalent to **Case-0**. Therefore, the **Uniform** method performs better than others. However, we can observe how the efficiency gap is reduced between **LA** and **Uniform** when the Knapsack capacity is more restricted. Eventually, the STD error in the achieved mitigation efficiency percentages for **Fair-LA** is significantly lower than **AVG-LA** which also shows how our proposed method is more stable.

D.6.4 Fairness Error

Since our proposed loss function (See Equation D.12) is considered a general fairness concept, we measure how fair the distribution of incentivization budget among all methods by calculating a normalized total loss. Figure D.4 shows how our proposed method significantly achieved less fairness error among other methods in all scenarios with stable STD error as well. Consequently, that resulted in not consuming the whole incentivization budget by our method. See Appendix A.4 [14] for more details about how **Fair-LA** is wisely consuming the Knapsack capacity.

D.6.5 Learning Bias

In the context of our work, a learning bias means unnecessary incentivization values to be assigned based on incomplete evaluation of users' needs due to the non-stationary problem. To reduce such bias, we considered a relatively small learning rate (the automaton state increase/ decrease value) that ensured all users will be visited almost equal times before consuming the whole budget. Moreover, the fairness error ensured that no user will consume more than its needs from the budget. Eventually, political polarization would reshape how the learned incentives could

actually cause mitigation. Hence, modeling the polarized responses to incentives should be integrated with our resolution in the future work.

D.6.6 Desired Mitigation Baseline

As demonstrated, the idea of misinformation mitigation is to introduce counter information by incentivizing users to propagate it on the network. However, a question remains about to which extend a mitigation should be considered enough. In other words, what if an equal exposure of counter information to misinformation is not enough to maintain authenticity on the network. In such scenario, we propose a balance factor parameter, where the ratios in Equation D.12 are considered fair only when approaching some balance. For instance, if the desired counter information exposure needed to be twice the amount of misinformation exposure per each user, then, the balance factor is set to 2 and the fairness of the ratio R is interpreted accordingly. See Appendix A.2 [14].

D.6.7 Computation Speed

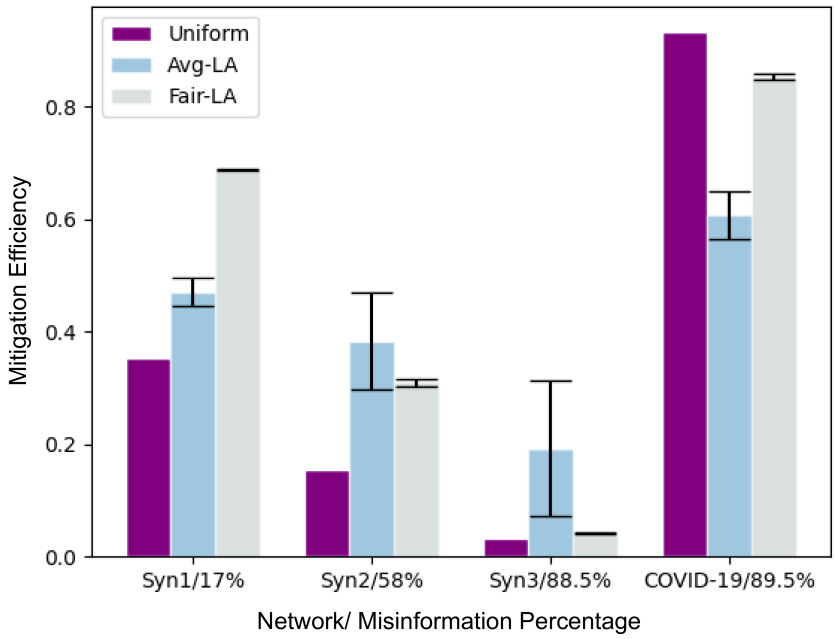
Due to the criticality of the misinformation problem, time is an important factor when evaluating misinformation mitigation resolutions. The complete comparison between **AVG-LA** and **Fair-LA** regarding their computation speed is given in Appendix A.5 [14].

D.6.8 Large Scaled Networks

Appendix A.6 [14] demonstrates how our method could be scaled on larger networks when sampling techniques are adopted to reduce the optimization space without sacrificing the mitigation efficiency.

D.7 Conclusion

This paper proposed a socially fair approach to misinformation mitigation on social networks. We introduced different synthetic social networks to generate diversity in scenarios where fairness will be critical to how we consume mitigation resources. Unlike other methods, where the fairness perspective was not considered and therefore the social networks which were evaluated were not diverse enough. However, as a limitation in our work, we did not consider the problem of non-responding users in a detailed manner. For instance, some users might be extremely polarized to respond to our incentivization even if their associated HP was responsive. Therefore, we believe that a model for political polarization can be integrated with our proposed method in the future.



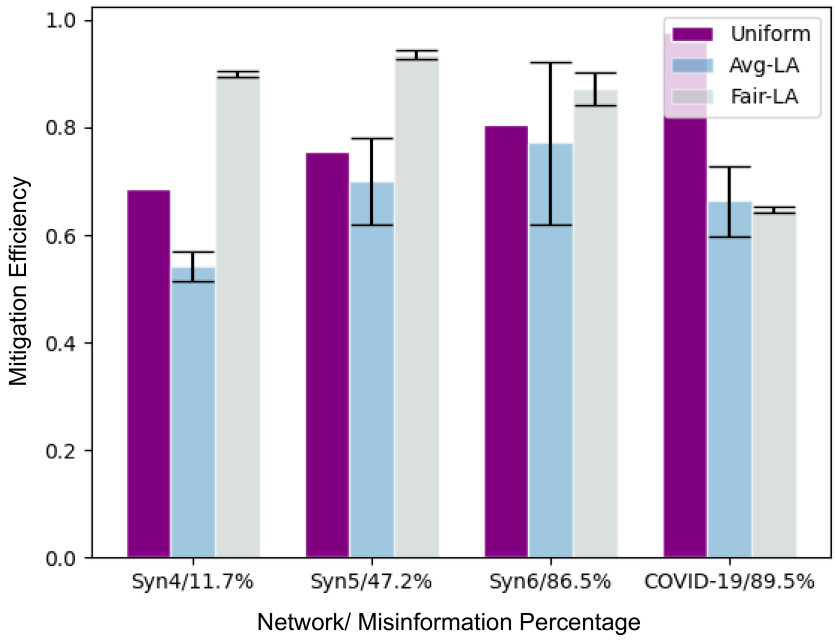


Figure D.3: Mitigation efficiency on different social network scenarios. Left image: C=0.06, right image: C=0.18

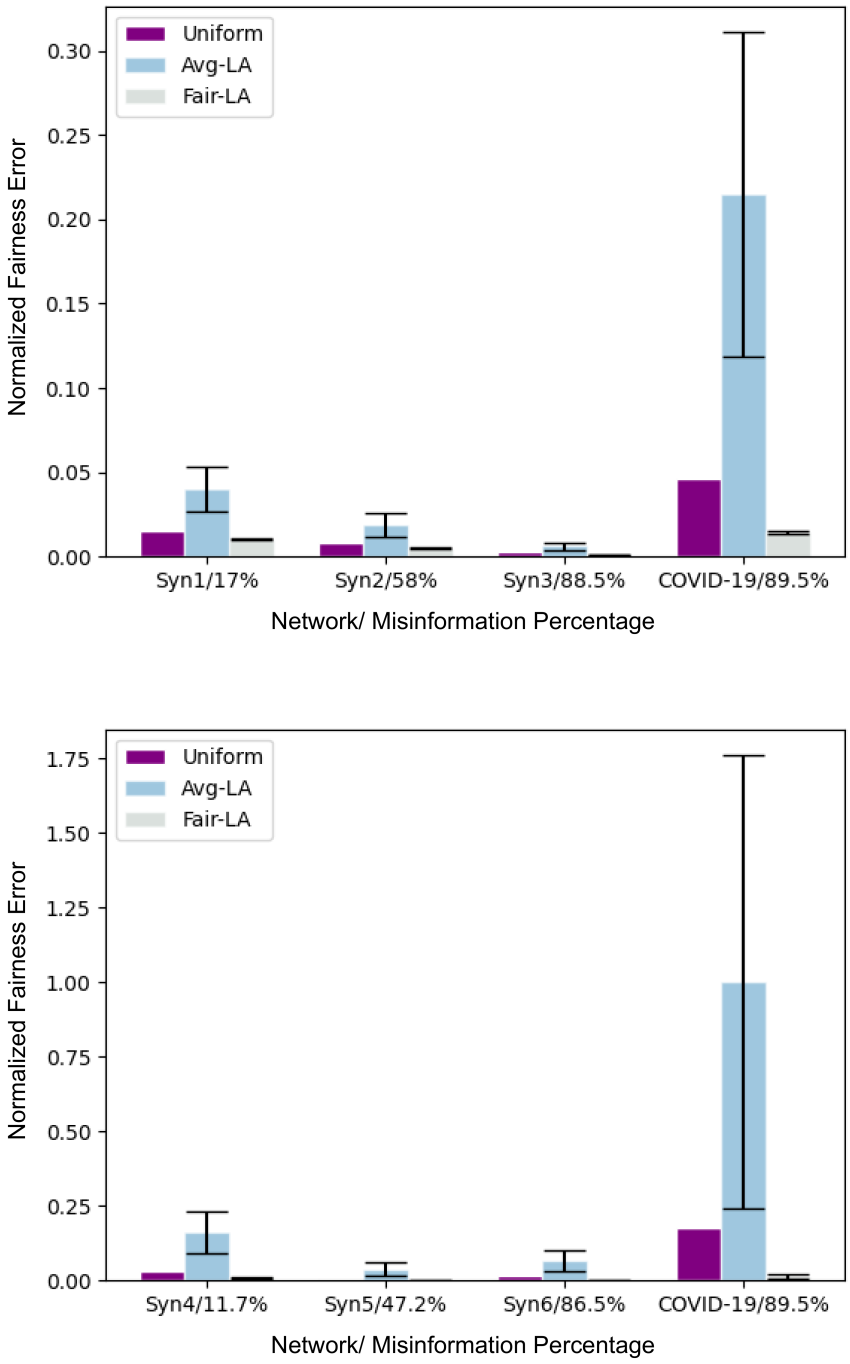


Figure D.4: Normalized fairness error on different social network scenarios. Left image: C=0.06, right image: C=0.18

Bibliography

[1] Dylan De Beer and Machdel Matthee. Approaches to identify fake news: A systematic literature review. In *International Conference on Integrated Science*, pages 13–22. Springer, 2020.

[2] David MJ Lazer, Matthew A Baum, Yochai Benkler, Adam J Berinsky, Kelly M Greenhill, Filippo Menczer, Miriam J Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, et al. The science of fake news. *Science*, 359(6380):1094–1096, 2018.

[3] Selman Özdan. The right to freedom of expression versus legal actions against fake news: A case study of singapore. In *The Epistemology of Deceit in a Postdigital Era*, pages 77–94. Springer, 2021.

[4] Mehrdad Farajtabar, Jiachen Yang, Xiaojing Ye, Huan Xu, Rakshit Trivedi, Elias Khalil, Shuang Li, Le Song, and Hongyuan Zha. Fake news mitigation via point process based intervention. In *International Conference on Machine Learning*, pages 1097–1106. PMLR, 2017.

[5] Ahmed Abouzeid, Ole-Christoffer Granmo, Christian Webersik, and Morten Goodwin. Learning automata-based misinformation mitigation via hawkes processes. *Information Systems Frontiers*, pages 1–20, 2021.

[6] Kumpati S Narendra and Mandayam AL Thathachar. Learning automata-a survey. *IEEE Transactions on systems, man, and cybernetics*, (4):323–334, 1974.

[7] Patrick J Laub, Thomas Taimre, and Philip K Pollett. Hawkes processes. *arXiv preprint arXiv:1507.02822*, 2015.

[8] Karl Pearson. The problem of the random walk. *Nature*, 72(1867):342–342, 1905.

[9] Gaia Nicosia, Andrea Pacifici, and Ulrich Pferschy. On multi-agent knapsack problems. In *CTW*, pages 44–47, 2009.

[10] Mahak Goindani and Jennifer Neville. Social reinforcement learning to combat fake news spread. In *Uncertainty in Artificial Intelligence*, pages 1006–1016. PMLR, 2020.

[11] Yuanda Chen. Thinning algorithms for simulating point processes. *Florida State University, Tallahassee, FL*, 2016.

[12] Tohru Ozaki. Maximum likelihood estimation of hawkes' self-exciting point processes. *Annals of the Institute of Statistical Mathematics*, 31(1):145–155, 1979.

[13] Yosihiko Ogata. On lewis' simulation method for point processes. *IEEE transactions on information theory*, 27(1):23–31, 1981.

[14] Ahmed Abouzeid, Ole-Christoffer Granmo, Christian Webersik, and Morten Goodwin. *Socially Fair Mitigation of Misinformation on Social Networks via Constraint Stochastic Optimization*. arXiv, 2022.

[15] Samantha Bradshaw and Philip Howard. *Troops, trolls and troublemakers: A global inventory of organized social media manipulation*. Oxford Internet Institute, 2017.

[16] Fang Jin, Wei Wang, Liang Zhao, Edward Dougherty, Yang Cao, Chang-Tien Lu, and Naren Ramakrishnan. Misinformation propagation in the age of twitter. *Computer*, 47(12):90–94, 2014.

[17] Giselle Rampersad and Turki Althiyabi. Fake news: Acceptance by demographics and culture on social media. *Journal of Information Technology & Politics*, 17(1):1–11, 2020.

[18] Cristina M Pulido, Laura Ruiz-Eugenio, Gisela Redondo-Sama, and Beatriz Villarejo-Carballido. A new application of social impact in social media for overcoming fake news in health. *International journal of environmental research and public health*, 17(7):2430, 2020.

[19] Hunt Allcott and Matthew Gentzkow. Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2):211–36, 2017.

[20] Karishma Sharma, Feng Qian, He Jiang, Natali Ruchansky, Ming Zhang, and Yan Liu. Combating fake news: A survey on identification and mitigation techniques. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(3):1–42, 2019.

[21] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, 19(1):22–36, 2017.

[22] Qiao Zhang, Shuiyuan Zhang, Jian Dong, Jinhua Xiong, and Xueqi Cheng. Automatic detection of rumor on social network. In *Natural Language Processing and Chinese Computing*, pages 113–122. Springer, 2015.

[23] Savvas Zannettou, Tristan Caulfield, William Setzer, Michael Sirivianos, Gianluca Stringhini, and Jeremy Blackburn. Who let the trolls out? towards understanding state-sponsored trolls. In *Proceedings of the 10th acm conference on web science*, pages 353–362, 2019.

[24] Chengcheng Shao, Giovanni Luca Ciampaglia, Onur Varol, Kai-Cheng Yang, Alessandro Flammini, and Filippo Menczer. The spread of low-credibility content by social bots. *Nature communications*, 9(1):1–9, 2018.

[25] Ahmed Abouzeid, Ole Christoffer Granmo, Christian Webersik, and Morten Goodwin. Causality-based social media analysis for normal users credibility assessment in a political crisis. In *2019 25th Conference of Open Innovations Association (FRUCT)*, pages 3–14. IEEE, 2019.

[26] Nedjma Ousidhoum, Yangqiu Song, and Dit-Yan Yeung. Comparative evaluation of label agnostic selection bias in multilingual hate speech datasets. In *Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP)*, pages 2532–2542, 2020.

[27] Jotham Wasike. Social media ethical issues: role of a librarian. *Library Hi Tech News*, 2013.

[28] Ole-Christoffer Granmo, B John Oommen, Svein Arild Myrer, and Morten Goodwin Olsen. Learning automata-based solutions to the nonlinear fractional knapsack problem with applications to optimal resource allocation. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37(1):166–175, 2007.

[29] Anis Yazidi and Hugo L Hammer. Solving stochastic nonlinear resource allocation problems using continuous learning automata. *Applied Intelligence*, 48(11):4392–4411, 2018.

[30] Erkan Ulker and Vahit Tongur. Migrating birds optimization (mbo) algorithm to solve knapsack problem. *Procedia computer science*, 111:71–76, 2017.

[31] Arnaud Fréville. The multidimensional 0–1 knapsack problem: An overview. *European Journal of Operational Research*, 155(1):1–21, 2004.

[32] Amrita Gupta, Mehrdad Farajtabar, Bistra Dilkina, and Hongyuan Zha. Discrete interventions in hawkes processes with applications in invasive species management. In *IJCAI*, pages 3385–3392, 2018.

[33] Charles Ames. The markov process as a compositional model: A survey and tutorial. *Leonardo*, 22(2):175–187, 1989.

[34] Vitor M Marquioni. Multidimensional elephant random walk with coupled memory. *Physical Review E*, 100(5):052131, 2019.

[35] Katerina Gotzamani, Andreas Georgiou, Andreas Andronikidis, and Konstantina Kamvysi. Introducing multivariate markov modeling within qfd to anticipate future customer preferences in product design. *International Journal of Quality & Reliability Management*, 2018.

[36] Wubshet Tesfaye, Solomon Abrha, Mahipal Sinnollareddy, Bruce Arnold, Andrew Brown, Cynthia Matthew, Victor M Oguoma, Gregory M Peterson, and Jackson Thomas. How do we combat bogus medicines in the age of the covid-19 pandemic? *The American Journal of Tropical Medicine and Hygiene*, 103(4):1360–1363, 2020.

Appendix E

Paper E

Title:	MMSS: A Storytelling Simulation Software to Mitigate Misinformation on Social Media
Authors:	Ahmed Abouzeid, Ole-Christoffer Granmo
Affiliation:	University of Agder, Faculty of Engineering and Science, P. O. Box 509, NO-4898 Grimstad, Norway
Journal:	*Software Impacts*
DOI:	10.1016/j.simpa.2022.100341.

MMSS: A Storytelling Simulation Software to Mitigate Misinformation on Social Media

Ahmed Abouzeid, Ole-Christoffer Granmo
Department of Information and Communication Technology
Faculty of Engineering and Science, University of Agder
P.O. Box 509, NO-4898 Grimstad, Norway
E-mails: {ahmed.abouzeid, ole.granmo}@uia.no

***Abstract* — This paper proposes a modular python implementation of a storytelling simulation. The software evaluates misinformation mitigation strategies over Social Media (SM) and visualizes the investigated scenarios' potential outcomes. Our software integrates information diffusion and control models components. The control model mitigates users' exposure to misinformation with social fairness awareness, while the diffusion model predicts the outcome from the control model. During the interaction of both models, a graph coloring algorithm traces the interaction within specific time intervals. Then, it generates meta-data to construct visuals of predicted near-future states of the social network to help support decision-making and evaluate proposed mitigation strategies.**

E.1 Introduction

Our Misinformation Mitigation Storytelling Simulation (MMSS) software works by receiving social network data. The data must consist of historical observations regarding the information dissemination behavior for the individual users. Such behavior is represented by the timestamps of each user for when a particular type of information was created. The MMSS expects such information to be labeled as either misinformation or regular content. Then the next task is for the information diffusion component to model and learn from these recorded observations. Typically, the information diffusion model task is to predict the future behavior of each user if either misinformation or regular content will be circulated during a specific time realization soon. However, the information diffusion model is controlled first before conducting any predictions, allowing for a misinformation mitigation strategy to alter the present for better consequences in the future. Therefore, each user is picked up and assigned an Artificial Intelligence (AI)-based agent to learn how such a user will influence and be influenced inside the network. The MMSS has a sampler component to shrink vast networks or target a specific community. Then, the picked user's behavior is examined by the controller agent within the sampled network. Sampled networks can be formed randomly or by targetting specific groups. When all users are assigned agents and the latter finalized learning and converged, the controller component reports the results of its interventions to the data visualization layer to view intuitions about how effective a mitigation strategy could be. The MMSS helps investigate different mitigation strategies through system parameters. The experiments and empirical results from adopting our software are illustrated in [1]. Eventually, reported results are standard colored graphs meta-data that can be easily integrated with graph visualization software. Figure E.1 shows the main components and layers of the MMSS software.

E.2 The Motivation for MMSS

Efficient Crisis Management Systems rely on visual analytics to filter and visualize relevant information extracted from SM platforms like Twitter and Facebook. The provided analytics from these tools equip emergency responders with different points of view to explore and better understand the situation and take a specific course of actions [2]. Recent works have proposed data visualization techniques for emergency operators. These tools took advantage of the large amount of data generated on the social networks every second. A non-exclusive list of applications is health monitoring [3], organized crime [4], hate speech [5], and gender bias [6]. On the other hand, simulation framework designs were introduced for crisis management as well. The latter has a distinctive advantage in informing emergency responders of potential risks or preferable actions to mitigate threats in the future. Furthermore, other efforts [7] presented a perennial simulation framework that targets crisis management simulation. Their framework incorporated concepts of dynamic data-driven systems, symbiotic simulation, and human-in-the-loop techniques [8]. Others

worked on visualization-based techniques, [9] and proposed a mass transport system simulation to familiarize the operators on the way to handle emergencies by carrying out virtual drills. To the best of our knowledge, the literature lacks resolutions for simulation-based visualization techniques for the problem of misinformation mitigation on SM. Hence, the proposed software in this paper tries to fill the gap and provide a practical resolution to emergency responders to help support their decision-making when working on an infodemic crisis.

E.3 Methodological Foundation

E.3.1 The Diffusion Model

The information diffusion model is responsible for predicting the behavior of individual users on the social network. The model is based on a Multivariate Hawkes Process (MHP) as practiced by [10], [11], and [12]. A MHP is a multivariate stochastic process [13] which models the occurrence of temporal or spatio-temporal asynchronous events by capturing their mutual dependencies. The behavior of each individual on the network is modeled through two Hawkes Processes (HPs), one for the misinformation dissemination the user is involved with and the other for the regular content dissemination of the same user. The associated user HPs generate estimated random counts for both information kinds, given some behavior observation in the past. For instance, predicting future re/tweeted events given the observed mutual dependencies with other users, where observed dependencies are estimated from the given social network historical and labeled data. These final counts indicate the intensity of each behavior style at a specific time realization. Mathematically, a HP can be defined with its conditional intensity function λ. The conditional intensity function has two significant elements: base intensity μ and exponential decay function g over an adjacency matrix A which represents the estimated mutual dependencies. The formal explanation of the conditional intensity function is given as the below equation:

$$\lambda_i(t_r|H^{t_r}) := \mu_i + \sum_{t_s<t_r} g(t_r - t_s), \tag{E.1}$$

where μ is the base intensity that represents some external motivation to propagate some content. On the other hand, g is some kernel function over the observed history H^{t_r} associated with user i from the discrete time realization t_s prior to time t_r. g is concerned with the history of some influence matrix $A_{i.}$, where $A_{ij} = 1$ if there is an influence indicating that user i follows user j or quotes (with agreement) content from j, and $A_{ij} = 0$ if not. We used an exponential decay kernel function $g = A_{i.}e^{-wt}$ as practiced by [12], where w is the decay factor which represents the rate for how the influence is reduced over time. For all users, the base intensity vector μ and the influence matrix A can be estimated using maximum likelihood as proposed in [14]. To predict all users' behaviors for each content type, a MHP is created, given that different intensity rates are generated at different discrete-time

realizations. Hence, we model each user behavior at each time realization as an estimated number of events (misinformation or regular). We utilized the *tick* library [15] for our implementation of the Hawkes-based information diffusion model.

E.3.2 The Control Model

On the other hand, controlling the diffusion model means introducing additional information to the network to alter the diffusion's future outcomes. The misinformation mitigation strategies could be implemented by introducing specific information to the network. For instance, when an individual is exposed to a certain amount of misinformation and then being exposed to its correction. Such imposed correction can be viewed as mitigating the impact of manipulating content. Therefore, the base intensity μ in the regular content HP is the element being under control. Therefore, For each user, an incentivization value x is added to the external motivation of the HP. The incentivization values are bounded by a predetermined mitigation budget C, representing time limitations or other incentivization constraints. Hence, let x_i be the incentivization amount decided for user i and $x_i \leq C$, the modified HP for the mitigation purposes can be redefined by the below equation:

$$\lambda_i(t_r|H^{t_r}) := x_i + \mu_i + \sum_{t_s<t_r} g(t_r - t_s). \tag{E.2}$$

To intelligently apply incentivizations to the network, we need an intervention model that can learn from the observed social network dynamics. Therefore, we utilized Learning Automaton (LA) [16] for its capabilities, easy implementation, and light-weight computation. We believe easier implementations and lower computation costs are important and needed for the practicality of our proposed MMSS. For each individual user, we assign a LA controller to intervene with the simulated user from the information diffusion model. Therefore, We built a network of user-assigned LAs [11] for learning optimum incentivization value needed for each assigned user in an optimum mitigation strategy.

E.3.3 Results Visualization

The data visualization layer takes advantage of our proposed graph coloring algorithm. The latter mainly produces meta-data for colored and sized graph nodes. The outputted information provides a detailed story that includes temporal changes in the predicted consequences over the network. For instance, when users start to follow, tweet, and retweet specific information types with other users, that includes starting and ending times. The traced temporal information stored for nodes and edges is important to provide dynamic transitioned graphs overtime frame by frame. Also, the colored graph meta-data contains different nodes flags to help distinguish users' nodes from content nodes. For a fully detailed illustration about the dynamic graph transitions generated for the different told stories by the MMSS, please visit the following link: https://www.youtube.com/watch?v=Lqmp4PdWCp4.

E.4 MMSS Impact and Target Domain

As discussed earlier, there is a critical need for simulation-based visualization techniques for the problem of misinformation mitigation on SM. The proposed MMSS software provides a resolution and fills such gap. Moreover, the provided mitigation strategy parameters can help investigating different strategies with different consequences. The latter is a big advantage when working on emergencies and crisis related to political manipulation for instance. Hence, the proposed MMSS software can be utilized by emergency responders when an infodemic is causing social disorder or an extreme level of political polarization based on some political misinformation. Therefore, mitigating the misleading content on SM could be approached by using the MMSS to evaluate and adopt learned misinformation mitigation strategies, where the outcome of the MMSS would be the different consequences of the different evaluated and simulated mitigation strategies. The latter outcome can also be represented by a dynamic social network state transitions on temporal basis, generated by the graph coloring component the MMSS has. The mentioned impact was scientifically evaluated on both real and synthetic data and the work [1] was accepted under the track of AI for social impact at the AAAI22 venue.

E.5 Adoption of MMSS in Misinformation Research

The development of MMSS considered multiple aspects to provide practicality and robustness as a misinformation crisis management system. Some of these offered practicalities are mitigation strategy parameters, traced temporal meta-data while simulating the mitigation strategy, and an easy and decentralized implementation inspired by the capabilities of the LAs network. However, the problem of SM misinformation is intricate and interconnected with other issues like hate speech, political polarization, and the echo chamber effect inherited by SM platforms' technology. Therefore, we believe the MMSS modular implementation could be the first step towards a fully integrated pipeline for a storytelling simulation for misinformation/hate speech/polarization mitigation on SM. The mitigation incentives could be viewed as motivational causes SM providers can apply to specific users at specific time intervals. After evaluating different mitigation strategies and supporting mitigation decisions, how intensive each user's motivation could be is decided from the MMSS learning. Moreover, the modular implementation of the MMSS software allows for reusing a subset of its components while replacing the others for the sake of scientific evaluation. For instance, researchers can adopt different controller models to make an analogy between different controllers within the MMSS pipeline. Similarly, information diffusion model can be replaced with another diffusion methodology.

E.6 Limitations and Future Work

Modeling human behavior is entirely different and more complicated than modeling an artificial system. Hence, formal verification of such social models should be studied along with interpretable AI consideration in the utilized methods. For instance, interpretable information diffusion models are highly needed to avoid mistaken imitation of users' behaviors which is considered a drawback of misinformation classification methods. Additionally, the proposed MMSS depends on labeled historical data from the social network, which makes the MMSS still dependent on such classification methods. Although, the mitigation approach reduces the risk of these classifiers' drawbacks since the former works by incentivization, making a wrong decided incentivization value less harmful. However, integrating an interpretable misinformation classifier with MMSS in future work is also highly recommended for both interpretability and dependability characteristics.

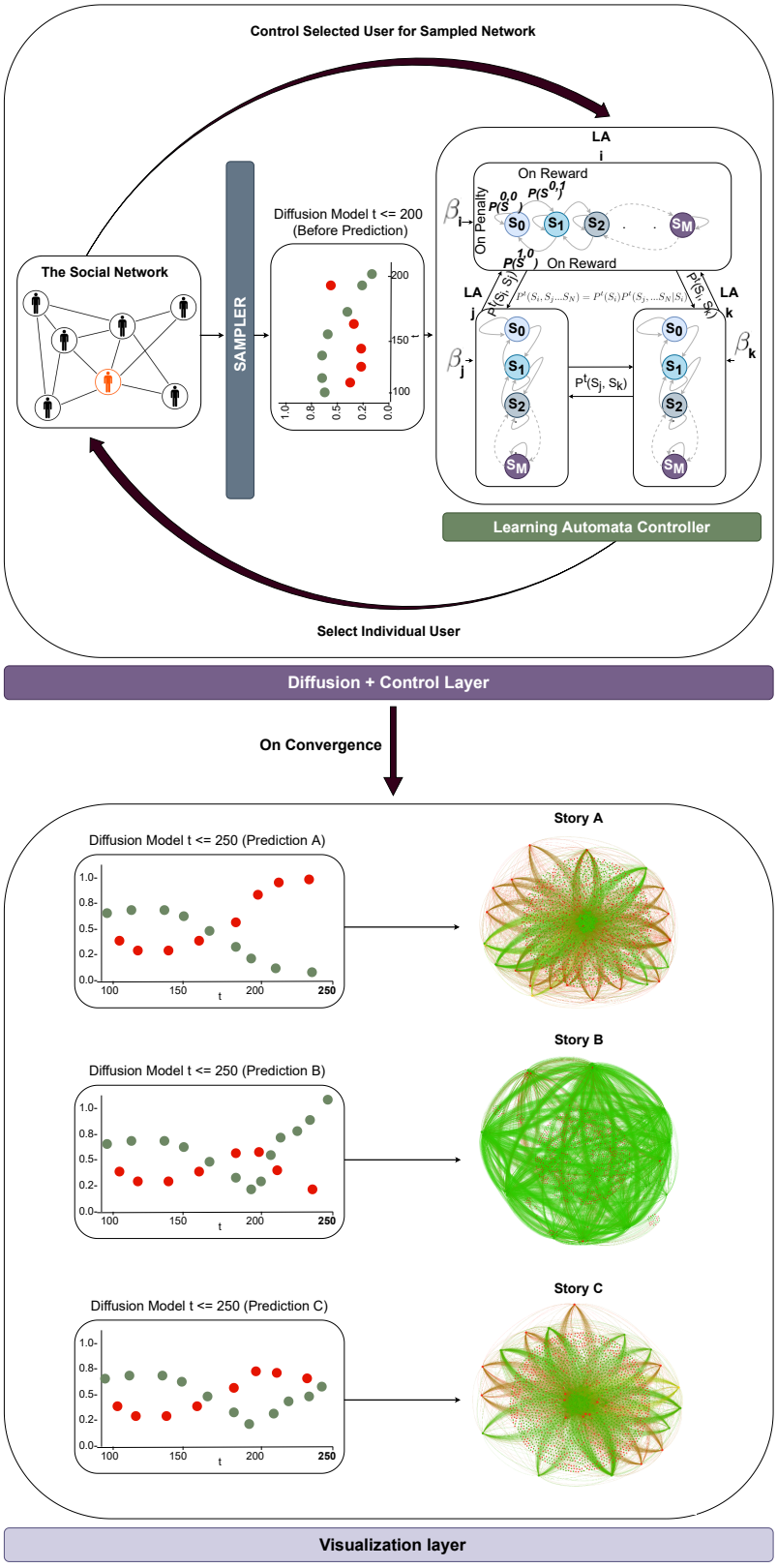


Figure E.1: MMSS architecture and underlying components. Red nodes indicate misinformation, while green nodes indicate regular content

Bibliography

[1] Ahmed Abouzeid, Ole-Christoffer Granmo, Christian Webersik, and Morten Goodwin. Socially fair mitigation of misinformation on social networks via constraint stochastic optimization. *arXiv preprint arXiv:2203.12537*, 2022.

[2] Teresa Onorati, Paloma Díaz, and Belen Carrion. From social networks to emergency operation centers: A semantic visualization approach. *Future Generation Computer Systems*, 95:829–840, 2019.

[3] Rongheng Lin, Zezhou Ye, Hao Wang, and Budan Wu. Chronic diseases and health monitoring big data: A survey. *IEEE reviews in biomedical engineering*, 11:275–288, 2018.

[4] Simon Andrews, Ben Brewster, and Tony Day. Organised crime and social media: a system for detecting, corroborating and visualising weak signals of organised crime online. *Security Informatics*, 7(1):1–21, 2018.

[5] Arthur TE Capozzi, Mirko Lai, Valerio Basile, Fabio Poletto, Manuela Sanguinetti, Cristina Bosco, Viviana Patti, Giancarlo Ruffo, Cataldo Musto, Marco Polignano, et al. Computational linguistics against hate: Hate speech detection and visualization on social media in the" contro l'odio" project. In *6th Italian Conference on Computational Linguistics, CLiC-it 2019*, volume 2481, pages 1–6. CEUR-WS, 2019.

[6] Prashanth Rao and Maite Taboada. Gender bias in the news: A scalable topic modelling and visualization framework. *Frontiers in Artificial Intelligence*, 4, 2021.

[7] Seth N Hetu, Samarth Gupta, Vinh-An Vu, and Gary Tan. A simulation framework for crisis management: Design and use. *Simulation Modelling Practice and Theory*, 85:15–32, 2018.

[8] Fabio Massimo Zanzotto. Human-in-the-loop artificial intelligence. *Journal of Artificial Intelligence Research*, 64:243–252, 2019.

[9] PK Kwok, Bill KP Chan, and Henry YK Lau. A virtual collaborative simulation-based training system. In *Proceedings of the 10th International Conference on Computer Modeling and Simulation*, pages 258–264, 2018.

[10] Mahak Goindani and Jennifer Neville. Social reinforcement learning to combat fake news spread. In *Uncertainty in Artificial Intelligence*, pages 1006–1016. PMLR, 2020.

[11] Ahmed Abouzeid, Ole-Christoffer Granmo, Christian Webersik, and Morten Goodwin. Learning automata-based misinformation mitigation via hawkes processes. *Information Systems Frontiers*, pages 1–20, 2021.

[12] Mehrdad Farajtabar, Jiachen Yang, Xiaojing Ye, Huan Xu, Rakshit Trivedi, Elias Khalil, Shuang Li, Le Song, and Hongyuan Zha. Fake news mitigation via point process based intervention. In *International Conference on Machine Learning*, pages 1097–1106. PMLR, 2017.

[13] Yuanda Chen. Thinning algorithms for simulating point processes. *Florida State University, Tallahassee, FL*, 2016.

[14] Tohru Ozaki. Maximum likelihood estimation of hawkes' self-exciting point processes. *Annals of the Institute of Statistical Mathematics*, 31(1):145–155, 1979.

[15] Emmanuel Bacry, Martin Bompaire, Philip Deegan, Stéphane Gaïffas, and Søren Poulsen. tick: a python library for statistical learning, with an emphasis on hawkes processes and time-dependent models. *J. Mach. Learn. Res.*, 18(1):7937–7941, 2017.

[16] Kumpati S Narendra and Mandayam AL Thathachar. Learning automata-a survey. *IEEE Transactions on systems, man, and cybernetics*, (4):323–334, 1974.

Appendix F

Paper F

Title:	Label-Critic Tsetlin Machine: A Novel Self-supervised Learning Scheme for Interpretable Clustering
Authors:	Ahmed Abouzeid, Ole-Christoffer Granmo, Morten Goodwin, Christian Webersik
Affiliation:	University of Agder, Faculty of Engineering and Science, P. O. Box 509, NO-4898 Grimstad, Norway
Conference:	*2022 International Symposium on the Tsetlin Machine (ISTM)*, Grimstad, Norway
DOI:	10.1109/ISTM54910.2022.00016.

Label-Critic Tsetlin Machine: A Novel Self-supervised Learning Scheme for Interpretable Clustering

Ahmed Abouzeid, Ole-Christoffer Granmo, Morten Goodwin, Christian Webersik
Department of Information and Communication Technology
Faculty of Engineering and Science, University of Agder
P.O. Box 509, NO-4898 Grimstad, Norway
E-mails: {ahmed.abouzeid, ole.granmo, morten.goodwin, christian.webersik}@uia.no

Abstract **— Unlike typical machine learning algorithms such as Artificial Neural Networks, the Tsetlin Machine (TM) is based on propositional logic instead of arithmetic operations, promoting it as a novel machine learning paradigm with interpretable learning outcomes. To this end, this paper proposes a self-supervised learning scheme inspired by the self-correction and interpretability provided by a standard TM. The proposed architecture uses a twin of Label-Critic Tsetlin Automata (TAs). The Label-TA learns the individual samples' correct labels guided by a self-corrected TM logical clause. At the same time, the Critic-TA validates the learning and approves the Label-TA reward. Our empirical results on synthetic and real data show promising capabilities for self-supervised learning and interpretable clustering. Furthermore, the Label-Critic TM architecture demonstrates how propositional logic-based learning provides self-correction with the absence of the ground truths in data.**

F

F.1 Introduction

The TM [1] is based on groups of TAs [2] that collaborate to learn unique patterns in data. Typically, the TM conducts supervised learning where the individual TAs learn from the provided data labels and features values. A single TA is a finite state machine that learns by interacting with an environment and updates its state transition according to a stochastic signal. Depending on the current state a TA is in, a particular decision is favoured. In the context of TM, each feature in the training data is assigned a TA. Hence, each TA interacts with its indexed data feature in all samples and seeks convergence to an optimal decision. In that case, the decision is to include or exclude the associated feature in a learned propositional logic statement. From the collaboration of all TAs, a single TM eventually learns one True clause of propositions, representing a unique sub-pattern with features corresponding to a specific data class.

A learned clause is a conjunction of literals, where a literal is a propositional feature input or its negation. The final sub-pattern recognition task can then be performed according to a voting scheme from all learned clauses against a given data sample to be classified. Such propositional logic for knowledge representation provides more interpretable logical rules instead of arithmetic computation found in deep learning approaches. Furthermore, the core TAs facilitate only incremental and decremental arithmetic operations during learning, which reduces memory footprints generated by the TM. Figure F.1 demonstrates how a single TA learns to include or exclude a feature in a TM clause, where rewards and penalties are stochastic signals received based on the ground truths in data. The number of states N per each decision represents the memory depth for that decision, and the larger the depth, the higher the confidence.

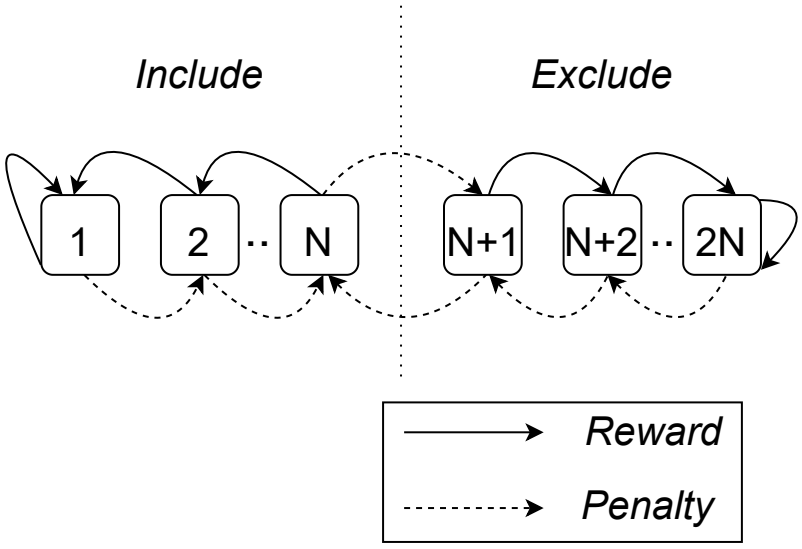


Figure F.1: TA architecture to include or exclude a feature in a TM clause

The TM clause-based inference structure can be explained in Figure F.2, where final sub-pattern recognition is achieved from composing multiple clauses from mul-

tiple TMs before applying the voting. The underlying TAs are coordinated by a game that optimizes accuracy using the so-called Type I and Type II Feedbacks which dictate how the reward-penalty signals will be determined from the ground truths. As indicated in Figure F.2, a standard TM receives a binary vector of features as its input. The number of clauses employed is a user set parameter. Half of the clauses are assigned positive polarity, while the other half are assigned negative polarity. A class positive polarity clauses are the learned rules that support the class sub-pattern(s). On the other hand, a class negative polarity clauses are the learned rules that support that class's negation. The clause evaluations are combined into a final output decision (classification) through summation and thresholding.

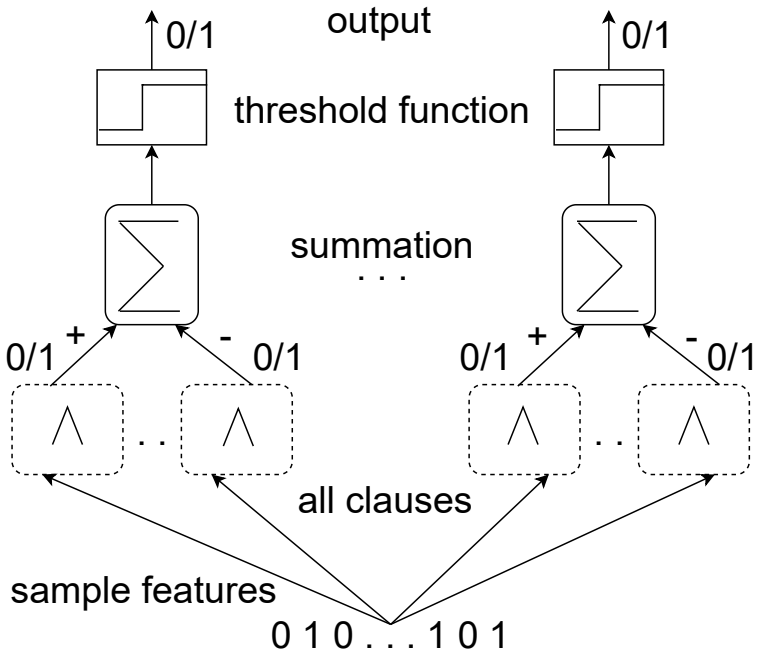


Figure F.2: Standard TM inference structure

During the standard TM learning, teams of TMs are established, where each team learns the logical rules about a particular class. Hence, Type I Feedback is triggered whenever the examined data sample has the same class label the TMs team is tasked for. Hence, Type I Feedback learns frequent features in the sub-patterns. On the contrary, Type II Feedback is triggered whenever the examined sample has different ground truth and then learns to boost the discriminative features between classes. Both Types play the role of rewarding and penalizing all TAs across TMs teams. For example, a TA receives a reward when it selects the *include* action on a feature value of 1 for a sample with the same class the TMs team is assigned for. Another scenario for rewarding is when the feature value is 0, and a TA selected to *exclude* for a sample with the same class the TMs team is assigned for. Furthermore, rewards are gained in counter scenarios when a feature with a value of 1 is excluded, and the sample label is from a different class. Similarly, rewards are gained when a feature with a value of 0 is included as a negated literal, and the sample label

is from a different class. Type I and II Feedbacks work stochastically to determine these rewards when their conditions are met; otherwise, they trigger penalties.

F.1.1 Contribution and paper Organization

The main contribution of this work is threefold:

- We propose a novel Reinforcement Learning (RL)-based architecture[1] on top of a standard TM to facilitate hierarchal clustering [3] based on self-corrected learning [4].
- We provide an interpretable approach to performing clustering tasks.
- We propose an augmentation of the standard TM Feedbacks to self-learn labels in datasets with large number of unique sub-patterns per class.

The remaining of this paper is organized as follows. section F.2 gives an overview on the related work. section F.3 explains the proposed Label-Critic TM architecture and the functionality of its components. section F.4 demonstrates the conducted experiments setup, datasets, evaluation metrics, and empirical results. The limitation of our work is mentioned in section F.5. Eventually, we conclude the paper in section F.6 and propose future directions.

F.2 Related Work

Typically, a pretext task is performed first in self-supervised learning, while solving the pretext allows solving the targeted task. For instance, Zhang et al. [5] colorize grayscale images by detecting the semantics of the scene and its surface texture to propose a cross-channel encoder. Pathak et al. [6] propose an unsupervised visual feature learning algorithm driven by context-based pixel prediction for image inpainting. In our work, the pretext task can be viewed as narrowing down the samples that might belong to the same class or cluster by initially assigning them the same labels. The pretext task usually depends on a supervisory signal generated from the data features by taking advantage of its structured sub-patterns. Zhu et al. [7] introduce Auxiliary Reasoning Navigation to leverage additional training signals derived from the semantic information in the text. Zhang et al. [8] introduce a self-supervised method that generates affordance labels for images by obtaining a segmentation mask for the object of interest, then cut it and inpaint the hole in the context image.

Self-supervised research in speech processing suffers from the absence of vocabulary of speech units over which a self-supervised learning task can be defined. For example, words in Natural Language Processing make establishing a self-supervised learning task easier. Therefore, several prominent models are equipped with mechanisms to learn an inventory of speech units as proposed by [9], a similar problem

[1] https://github.com/Ahmed-Abouzeid/Label-Critic-TM

tackled in [10], [11] for computer vision tasks. Similarly, Caron et al. [12] introduce learning representations invariant to data augmentation. Likewise, this paper presents augmentation techniques where insufficient information from features gave poor supervisory signals to our self-supervised algorithm.

Self-supervised learning is highly dependent on how well the pretext tasks are designed. The pretext tasks could introduce inductive biases during learning, leading to incorrect final labeling. Huynh et al. [13] recognize false negatives and provide strategies to utilize this ability to improve contrastive learning frameworks. However, the majority of these efforts still lack the interpretability in their results. In our work, we further propose a self-correction mechanism to combat both false negative and false positive matches, with provided interpretable clustering.

Interpretability is a fundamental quality for an Artificial Intelligence (AI) system. The TM has shown progress in that matter through recent years. For instance, Berge et al. [14] use a TM-based approach to learn human-interpretable rules from examined medical data. Bhattarai et al. [15] show how TM is able to detect novel content in documents with interpretable linguistic structures. Yadav et al. [16] introduce TM for word sense disambiguation depending on the surrounding context. Abeyrathna et al. [17] propose the Convolutional Regression TM for continuous output problems in image analysis. In all these efforts, the interpretability of TM results depends on the learned logical clauses. We believe the latter is advantageous for the self-supervised problems proposed in this paper, since logical clauses could provide a self-correction mechanism to ensure high-quality supervisory signals during training.

F.3 Methodology

F.3.1 Proposed Architecture

This paper proposes a novel architecture inspired by the established TM. We refer to our method as Label-Critic TM where the standard TM capacity is extended to perform classification or clustering tasks without the guidance of ground truths. To achieve that, as in the standard TM where TAs are attached on a feature level, we further attach TAs on a sample level. Hence, each data sample is assigned a single TA to learn its correct label. The samples' assigned TAs share same structure where one of two possible decisions to be made: either the sample belongs to class A or B. Hence, our approach is similar to top-down hierarchal clustering [18], where all data samples are considered one large cluster before splitting them into smaller ones until a cluster converges without further splitting. Recursively, we split the data into two classes groups: A and B, where each group shares similar sub-pattern(s) learned by a standard TM. Then, another TM training loop will learn discriminative sub-patterns for A and B, separately. If no more discriminative sub-patterns were found, the cluster converges, otherwise, more training loops are performed to further split. Eventually, data samples in the learned clusters can be labeled according to the converged cluster. For instance, for six learned clusters, their underlying samples

labels could be A, A', A'', B, B', B''.

When learning an individual sample label, a Label-TA is a single TA assigned to a sample with a randomly determined initial label. The Label-TA then sends its decision to the TM team, where the latter learn the clauses, given the data sample and its decided label. In parallel, a Critic-TA clones the initially decided labels and sends them to other TM team to perform separate clause learning. The duality of Label-TM/Critic-TM ensures validation of the final learned clauses, that is because of the higher probability of noise in training data. Both Label-TA and Critic-TA validate their decision according to a final True clause generated by their associated TMs after some post-processing. A True Clause is the final learned literals that support a particular class. The latter is an example of self-corrected learning facilitated by the TM propositional logic approach since noise-based literal-contradicted clauses can be corrected and concluded into one logically True clause. Figure F.3 shows an example of how Label-Critic TM invokes a class True clause generation by removing contradicted clauses' literals caused by noisy training data. However, the contradictions would exist in literals between two classes clauses as well, in that case a final True clause per each class is concluded according to the discriminative literals. For instance, if class A has the positive literal L_i, then class B True clause cannot have the same literal, unless it is negated.

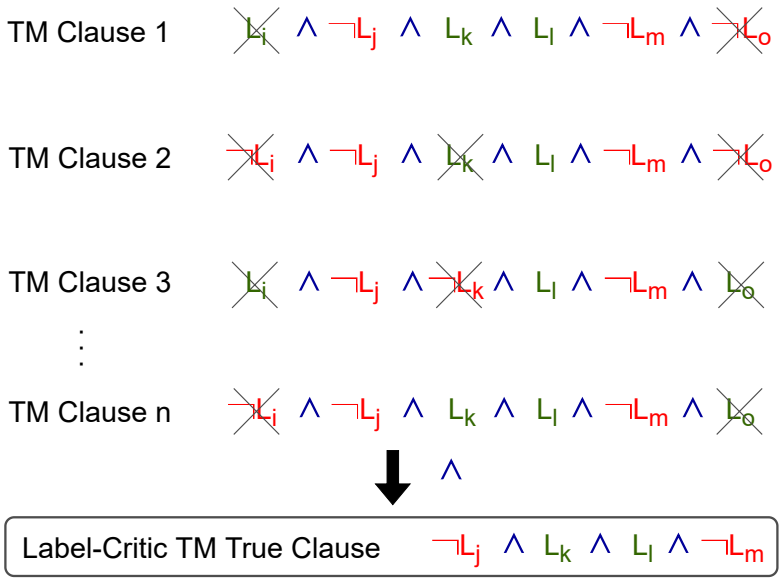


Figure F.3: Novel Label-Critic TM invokes the True clause generation by removing contradicted literals. The figure shows the same class clauses contradiction removal procedure

As indicated in Figure F.3, the generated clauses by the standard TMs would have contradicted literals due to the randomly assigned labels. However, propositional logic-based learning equips the Label-Critic TM with self-correction capabilities. For instance for n learned clauses by a standard TM team for a particular class, if the k^{th} feature has the propositions $L_k, L_k, \neg L_k, L_k$, it is concluded to L_k since L_k was canceled only once by its contradiction $\neg L_k$. At each training epoch

of the Label-Critic TM, the acquired True clause is accumulated to previous ones and same literal can have multiple instances in that case, making the True clause more informative and tuned by adding weighted literals. The generated True clause literals are then used for a similar voting scheme, just like a standard TM but on a literal instead of a clause level. Eventually, each Label-TA evaluates its label selection and is rewarded if its decision matches the class which got the higher votes. However, a Label-TA reward is determined only if its twin Critic-TA has decided the same label, otherwise penalized. Critic-TAs rewards are received the same way as in Label-TAs except that they do not require evaluation of their twin Label-TAs. Figure F.4 demonstrates the fundamental components and recursive cycles of the proposed Label-Critic TM where a recursive binary classification-based splitting is performed over the unlabeled data sample.

F.3.2 Augmenting Standard TM Feedbacks

The self-supervised learning proposed by the Label-Critic TM depends on the standard TM supervised learning mechanism with a minor change. The standard TM Feedback was designed to match the nature of the supervised learning problem, where multiple instances of the same feature are provided during features-assigned TAs learning. However, in our approach, we further attach more TAs to the standard TM, each TA is attached to an unlabeled training sample. Therefore, a dataset of unique samples would challenge the accuracy of the decided labels. Such obstacle is mainly because attached TAs are decentralized per each sample, making one unique sample not enough to learn a decision. Therefore, instead of augmenting data samples with a cost to computation time, we augmented Type I and II Feedbacks as an alternative solution. The latter technique consumed less memory allocation and gave almost the same expected results. The TM Feedbacks augmentation means boosting the *include* and *exclude* literal action's effect, by augmenting the state transitions for each augmented TA_i as demonstrated below.

$$s_i := s_i \pm \lambda, \tag{F.1}$$

where λ is the amount of augmented increase or decrease to the TA_i state when either penalized or rewarded. While in a standard TA and TM Feedback, λ is one step with value of 1.

F.3.3 Self-supervised Learning Scheme

To provide either a reward or a penalty to each TA selected action (label), we collect votes from each class True clause (C_0^T, C_1^T) literals. Each literal would vote +1 or 0 depending on a match with the current sample feature. To reduce voting time, only features with a binary value of 1 is included in the voting process. Given V_0 and V_1 as the final votes for class A and B True clauses, respectively. If the k^{th} feature in the current sample X_i is 1, a True clause literal L_k should increment +1 to V_0. Otherwise, if the literal is negated ($\neg L_k$), then it increments 0. On the contrary, V_1 is

incremented with +1 if the k^{th} literal is negated ($\neg L_k$), and incremented with 0 if the latter is positive (L_k). Each class True clause eventually decides either the favoured class: A IF $V_0 > V_1$, or B IF $V_1 > V_0$, otherwise, a neutral signal is sent back to TA_i. Consequently, the associated TA_i is rewarded with probability $P(r) = 1$ if its decision matches the final voted class, otherwise it is penalized with probability $P(\beta) = \epsilon$. The parameter $0 \leq \epsilon \leq 1$ is a user set value to allow exploration [19] and avoid infinite recursions that could occur due to a potential symmetric change in all TAs decided labels at once. For a sample X_i, given the features set $\{x_i^1, x_i^2, ..., x_i^n\}$, Equations 2 and 3 demonstrate our proposed self-supervised learning scheme, given the Label-Critic TM True clauses favoured class denoted by ϕ_i, where the decided classes by the sample X_i attached Label-TA and Critic-TA are denoted by ϕ_i^l and ϕ_i^c, respectively.

$$V_0 = \sum_{i=0}^{L} 1, \forall x_i \in X_i : x_i = 1 \text{ and } L_i \in C_0^T, \tag{F.2}$$

$$V_1 = \sum_{i=0}^{L} 1, \forall x_i \in X_i : x_i = 1 \text{ and } \neg L_i \in C_1^T, \tag{F.3}$$

$$\text{where } \phi_i = \begin{cases} A & \text{if } V_0 > V_1, \\ B & \text{if } V_0 < V_1, \\ -1 & \text{otherwise,} \end{cases}$$

$$\phi_i^l = \begin{cases} A & \text{if } N \geq s_i^l \geq 0, \\ B & \text{otherwise,} \end{cases}$$

$$\phi_i^c = \begin{cases} A & \text{if } N \geq s_i^c \geq 0, \\ B & \text{otherwise,} \end{cases}$$

$$s_i^l = \begin{cases} s_i^l - \lambda & \text{if } \phi_i = \phi_i^l = \phi_i^c \text{ and } 0 \leq s_i^l \leq N, \\ s_i^l + \lambda & \text{if } \phi_i = \phi_i^l = \phi_i^c \text{ and } N+1 \leq s_i^l \leq 2N, \\ \\ s_i^l + \lambda & \text{if } \phi_i \neq \phi_i^l \text{ and } d \leq \epsilon \text{ and } 0 \leq s_i^l \leq N, \\ s_i^l - \lambda & \text{if } \phi_i \neq \phi_i^l \text{ and } d \leq \epsilon \text{ and } N+1 \leq s_i^l \leq 2N, \\ \\ s_i^l & \text{otherwise,} \end{cases}$$

$$s_i^c = \begin{cases} s_i^c - \lambda & \text{if } \phi_i = \phi_i^c \text{ and } 0 \leq s_i^c \leq N, \\ s_i^c + \lambda & \text{if } \phi_i = \phi_i^c \text{ and } N+1 \leq s_i^c \leq 2N, \\ \\ s_i^c + \lambda & \text{if } \phi_i \neq \phi_i^c \text{ and } d \leq \epsilon \text{ and } 0 \leq s_i^c \leq N, \\ s_i^c - \lambda & \text{if } \phi_i \neq \phi_i^c \text{ and } d \leq \epsilon \text{ and } N+1 \leq s_i^c \leq 2N, \\ \\ s_i^c & \text{otherwise,} \end{cases}$$

where d is a random number between 0 and 1, and the state s_i is bounded by the TA action's memory depth.

F.4 Experiments Setup and Empirical Results

F.4.1 The Data Guess Game

Naturally, data classes can be viewed as a set of unique sub-patterns. There could be one or more unique sub-patterns in data samples, representing different levels of diversity in a single data class. A unique sub-pattern is a distinct features vector where a single or multiple samples would share. We challenged our method to play a data game; we refer to it as the *guess game*. The purpose of the *guess game* is to guess data labels, correctly. The *guess game* has three primary difficulty levels, representing simple to more complex data. Therefore, we adopted different types of unlabeled datasets represented by synthetic, semi-synthetic, and real data. Figure F.5 shows an example of the game level-1 dataset, where each class is represented by only 1 unique sub-pattern, where each of the latter has identical samples. However, the more unique sub-patterns in such a dataset, the more difficult to distinguish between the data classes.

Level-2 dataset had two classes, a class had either 2 or 5 unique sub-patterns. It is a semi-synthetic dataset because we augmented the samples to obtain identical samples per each sub-pattern. The unique sub-patterns were extracted from the MNIST dataset [20] for the digits Zero and One. On the other hand, we used the original MNIST dataset for the game's third level, where also two digits were extracted but with at least 500 unique sub-patterns per digit (no augmentation). The latter scenario is more challenging and replicates real-world datasets where a single class can be represented by hundreds or thousands of distinct samples.

F.4.2 Evaluation Metrics

F.4.2.1 Winning Probability

a straightforward evaluation metric to measure the consistency of the Label-Critic TM in playing the *guess game*. Therefore, the game was played over n rounds, and the probability of winning was simply calculated as how many rounds were won over all rounds. The game is won only if all labels were correctly identified.

F.4.2.2 Silhouette Score

an evaluation metric [21] used to calculate the goodness for a clustering procedure. Its value ranges between -1 and 1, where a value closer to 1 means that clusters are well apart from each other. We believe that metric is more convenient for level-1 experiments where samples are identical for the same class. Moreover, it was convenient to use when evaluating other clustering methods.

F.4.2.3 Visual Interpretation of Clusters

where the visual representation of clusters are the evaluated output. That demonstrates the interpretability of our method. We believe this is more convenient for an image-based dataset like MNIST. Therefore, the visual interpretation is mainly used for the second and third levels of the *guess game*.

F.4.3 Empirical Results

To assess our proposed Label-Critic TM architecture with an analogy to a simpler one, we set the *guess game* for 50 rounds. Then, we evaluated the Label-Critic TM against a standard TM with only attached TAs for guessing the labels while dropping the Critic-TAs. Figure F.6 demonstrates how the Critic-TAs architecture significantly boosted the probability of winning. In that experiment, we employed three synthetic datasets; each had a different number of unique sub-patterns where a class is represented by only 1 unique sub-pattern. In all datasets, the number of samples per class was 100, and the number of features was 400, where the number of examined sub-patterns was 2, 4, and 6.

On the other hand, to assess the Label-Critic TM architecture on a more challenging scenario with an analogy to clustering benchmarks [22], we evaluated eight datasets, each has different number of sub-patterns up to 10, where 1 unique sub-pattern per class. Figure F.7 shows how Label-Critic TM along with the DBSCAN method outperformed other clustering methods, including hierarchal-based clustering such as the Agglomerative method. In a game of 10 rounds, two synthetic data groups were evaluated. The first data group represented small balanced datasets with 400 features and 300 samples per class. The second data group represented relatively larger and unbalanced datasets with the number of features up to $2,000$, and the number of samples varied among classes. For instance, one class had $1,000$ samples while another had only 300 only.

The learning loops of Label-Critic TM can be observed in Figure F.8 where the learning is achieved through optimizing the total number of penalties from all the Label-TAs. The number of original classes in data dictates how many standard TM training loops are required to keep splitting the data into smaller clusters. Loops would eventually converge after some epochs. In some cases, due to the randomness of label initialization, a loop would be repeated if it reaches a particular epoch (threshold) without convergence.

For a higher level of difficulty in playing the *gues game*, we considered a class to be represented by a combination of unique sub-patterns instead of a single one. Hence, we introduce a semi-synthetic dataset with up to 5 unique sub-patterns per class. For that, The digits Zero and One were extracted from the MNIST. The MNIST images were converted to binary features vector where the values 0 and 1 were decided upon some threshold in the original dataset feature values. The 5 unique sub-patterns were augmented into $1,500$ samples per digit, 300 per sub-pattern. The Label-Critic TM managed to guess correct labels with classification accuracy of 1.0. Figure F.9 shows the accurate fetched clusters and interpretability

cabailities of the Label-Critic TM. As observed in the visual clusters, only 3 sub-patterns were extracted from 10 sub-patterns as in the original data, demonstrating how Label-Critic TM merged clusters that seemed similar.

The cluster visuals were obtained by reshaping the clusters associated True clauses literals to a 28×28 matrix using the literal indices to place them in their associated image pixel position. Then, a positive clause literal was intrepreted as a black pixel while negated literals were left white. For non-image-based datasets, measuring distances between learned True clauses would help in merging clusters as well.

Eventually, Figure F.10 shows how the Label-Critic TM managed to recognize 18 unique sub-patterns from the real MNIST dataset filtered by only the digits Zero and One and with 500 samples per digit. Therefore, the $1,000$ samples represented $1,000$ unique sub-patterns of hand-written Zeros and Ones. By setting $\lambda = 300$ instead of augmenting the samples as in previous experiments where $\lambda = 1$, the Label-Critic TM labeled the data samples with an accuracy of around 0.70 and almost accurate clusters. However, the underlying learned True clauses can be used to improve the classification accuracy as a further post-processing step. This experiment demonstrates level-3 of the *guess game* where the difficulty is maximized as in real-world scenario datasets where a single data class has hundreds or thousands of unique sub-patterns. Such difficulty can be observed in the visualized clusters with unclear patterns.

F.5 Limitation

According to our observation, the Label-Critic TM did not converge to the final clusters as fast as other clustering methods like Kmeans and DBSCAN. Furthermore, when evaluated on the real MNIST samples without sample augmentation, the total penalties optimal value was a bit above 0, meaning not all Label-TAs converged. The latter behavior demonstrates how the classification accuracy dropped from 1.0 to 0.70 after testing on real MNIST data.

F.6 Conclusion

Self-supervised learning has attracted enormous awareness for its data efficiency and generalization capability. However, self-supervised methods still lack interpretation in their results and suffer from obtaining a convenient evaluation of the identified data groups. The paper proposed a novel self-supervised learning scheme where data samples can be classified without knowing the ground truths and number of classes in training data. The proposed method demonstrated how our TM-based learning scheme can benefit from the logical clauses and self-correct itself. Moreover, the final True clauses were fundamental interpretable elements for the generated clusters. Future work would focus on improving the computation time of the Label-Critic TM architecture. Moreover, investigating the cabailities of our method on other

datasets with larger number of classes and distinct samples per class. For instance the complete MNIST and textual datasets.

F

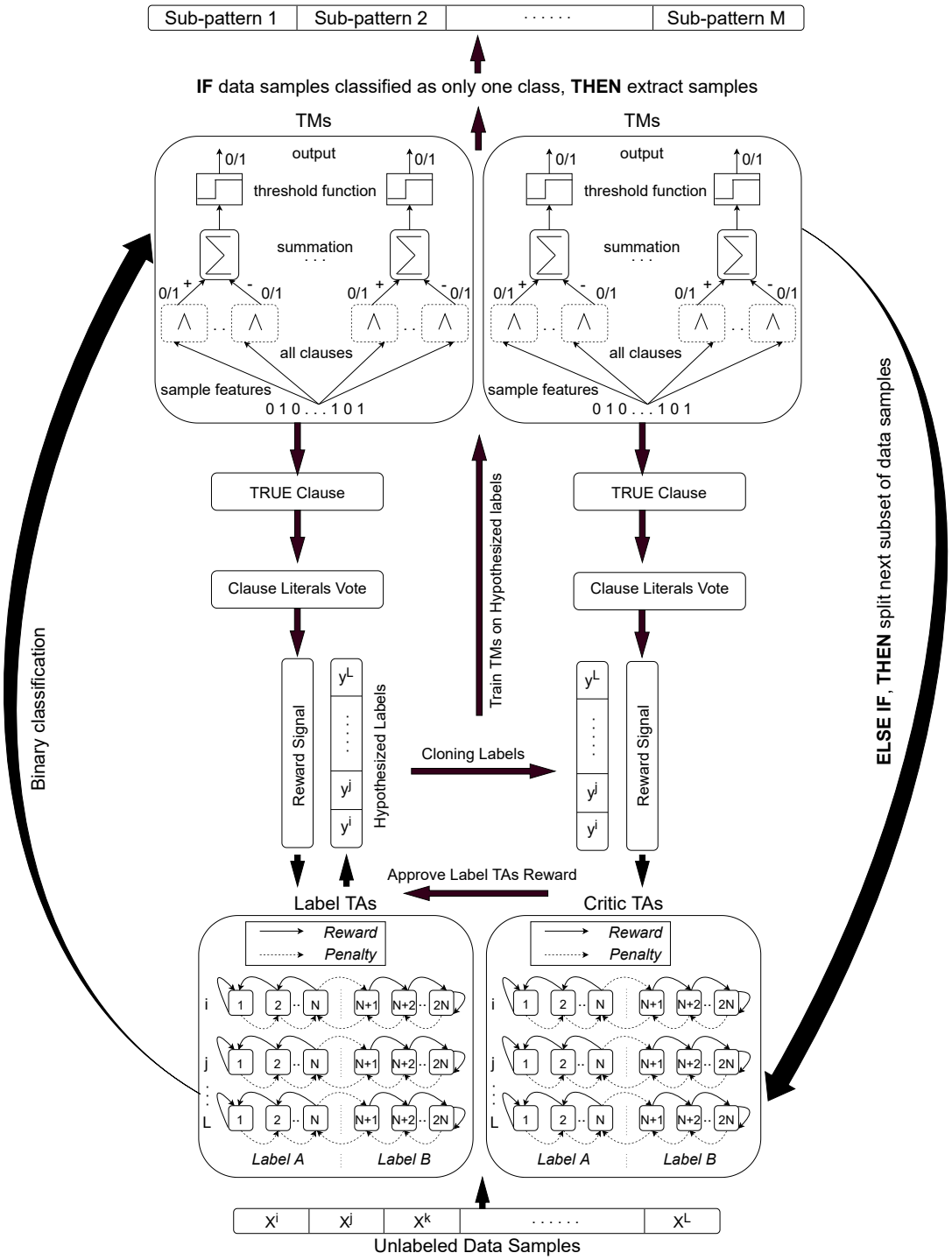


Figure F.4: Novel recursive self-supervised Label-Critic TM architecture

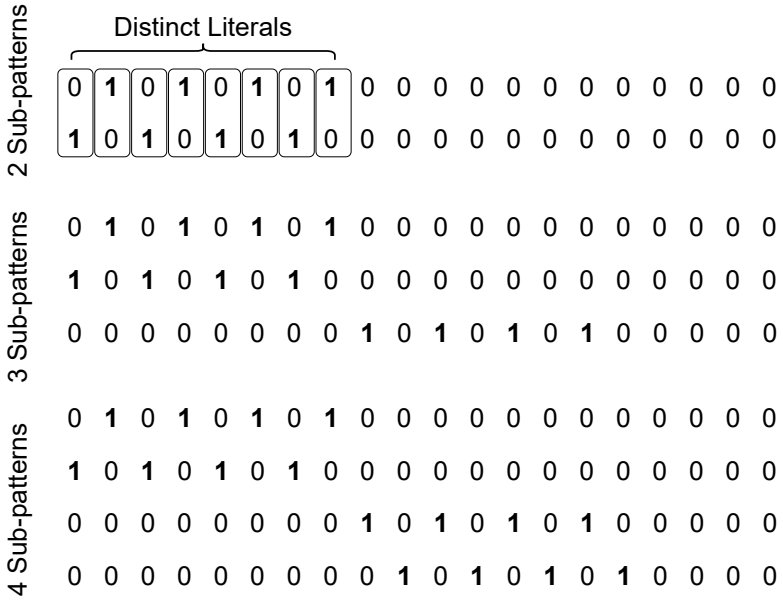


Figure F.5: Synthetic sub-patterns in proposed unlabeled data samples where a class is represented by only 1 sub-pattern. As noted, completely distinct literals only exist on 2 sub-patterns (2 classes)

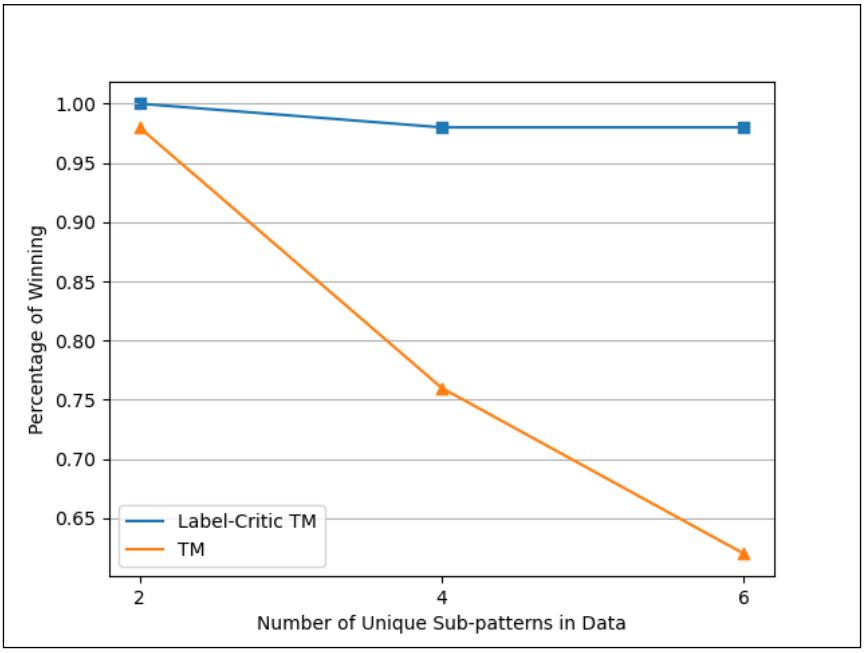


Figure F.6: Performance over 50 game rounds for both Label-Critic TM and the vanilla TM

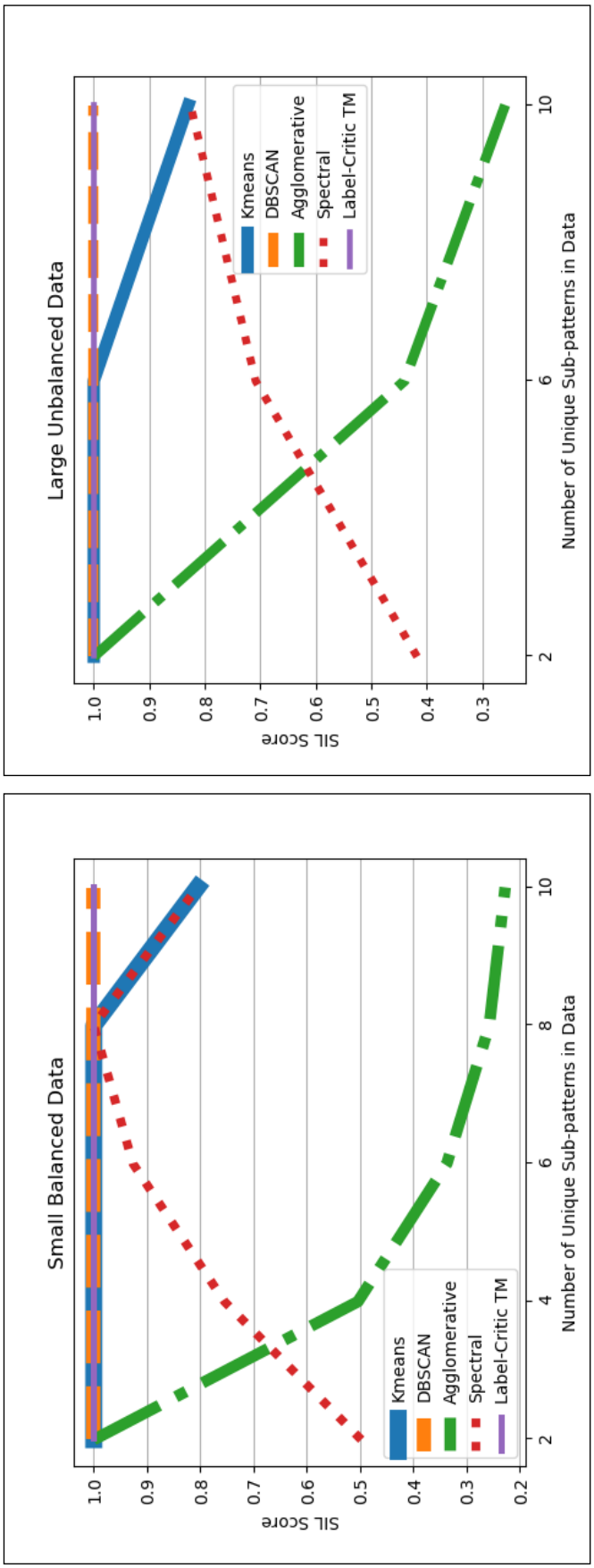


Figure F.7: Label-Critic TM performance comparison with benchmark clustering algorithms over 10 game rounds on balanced and unbalanced synthetic data

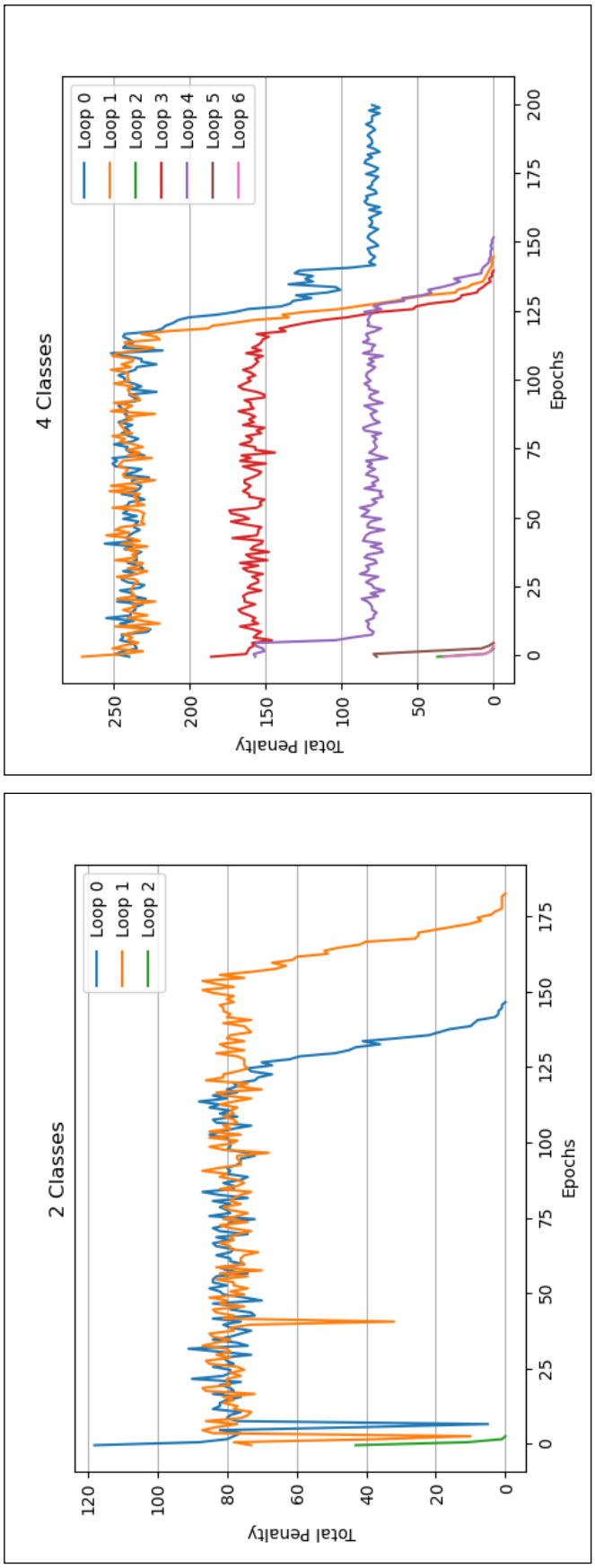


Figure F.8: Recursive self-supervised Label-Critic TM learning curve of conducted loops over different number of sub-patterns, one sub-pattern per class

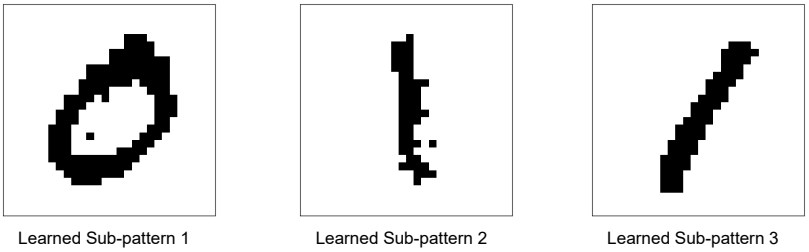


Figure F.9: Label-Critic TM interpreted clustered sub-patterns from augmented MNIST dataset, filtered with 5 unique sub-patterns per Class and augmented

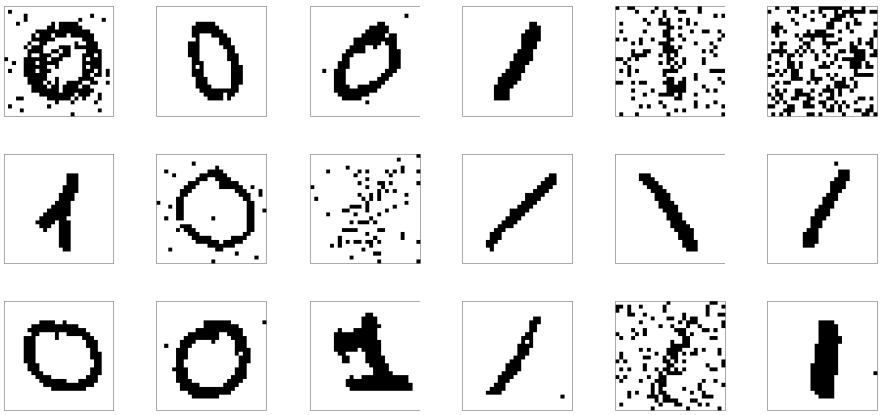


Figure F.10: Label-Critic TM interpretable clustered sub-patterns from original MNIST dataset

Bibliography

[1] Ole-Christoffer Granmo. The tsetlin machine–a game theoretic bandit driven approach to optimal pattern recognition with propositional logic. *arXiv preprint arXiv:1804.01508*, 2018.

[2] Michael Lvovitch Tsetlin. On behaviour of finite automata in random medium. *Avtomat. i Telemekh*, 22(10):1345–1354, 1961.

[3] AA Tarnutzer and KP Weber. Pattern analysis of peripheral-vestibular deficits with machine learning using hierarchical clustering. *Journal of the Neurological Sciences*, page 120159, 2022.

[4] Yunyun Wang, Chao Wang, Hui Xue, and Songcan Chen. Self-corrected unsupervised domain adaptation. *Frontiers of Computer Science*, 16(5):1–9, 2022.

[5] Richard Zhang, Phillip Isola, and Alexei A. Efros. Colorful image colorization. In Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling, editors, *Computer Vision – ECCV 2016*, pages 649–666, Cham, 2016. Springer International Publishing.

[6] Deepak Pathak, Philipp Krahenbuhl, Jeff Donahue, Trevor Darrell, and Alexei A Efros. Context encoders: Feature learning by inpainting. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2536–2544, 2016.

[7] Fengda Zhu, Yi Zhu, Xiaojun Chang, and Xiaodan Liang. Vision-language navigation with self-supervised auxiliary reasoning tasks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10012–10022, 2020.

[8] Lingzhi Zhang, Weiyu Du, Shenghao Zhou, Jiancong Wang, and Jianbo Shi. Inpaint2learn: A self-supervised framework for affordance learning. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 2665–2674, 2022.

[9] Alexei Baevski, Yuhao Zhou, Abdelrahman Mohamed, and Michael Auli. wav2vec 2.0: A framework for self-supervised learning of speech representations. *Advances in Neural Information Processing Systems*, 33:12449–12460, 2020.

[10] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pages 1597–1607. PMLR, 2020.

[11] Jean-Bastien Grill, Florian Strub, Florent Altché, Corentin Tallec, Pierre Richemond, Elena Buchatskaya, Carl Doersch, Bernardo Avila Pires, Zhaohan Guo, Mohammad Gheshlaghi Azar, et al. Bootstrap your own latent-a new approach to self-supervised learning. *Advances in Neural Information Processing Systems*, 33:21271–21284, 2020.

[12] Mathilde Caron, Hugo Touvron, Ishan Misra, Hervé Jégou, Julien Mairal, Piotr Bojanowski, and Armand Joulin. Emerging properties in self-supervised vision transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9650–9660, 2021.

[13] Tri Huynh, Simon Kornblith, Matthew R Walter, Michael Maire, and Maryam Khademi. Boosting contrastive self-supervised learning with false negative cancellation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 2785–2795, 2022.

[14] Geir Thore Berge, Ole-Christoffer Granmo, Tor Oddbjørn Tveit, Morten Goodwin, Lei Jiao, and Bernt Viggo Matheussen. Using the tsetlin machine to learn human-interpretable rules for high-accuracy text categorization with medical applications. *IEEE Access*, 7:115134–115146, 2019.

[15] Bimal Bhattarai, Ole-Christoffer Granmo, and Lei Jiao. Measuring the novelty of natural language text using the conjunctive clauses of a tsetlin machine text classifier. *arXiv preprint arXiv:2011.08755*, 2020.

[16] Rohan Kumar Yadav, Lei Jiao, Ole-Christoffer Granmo, and Morten Goodwin. Interpretability in word sense disambiguation using tsetlin machine. In *ICAART (2)*, pages 402–409, 2021.

[17] Darshana Abeyrathna, Ole-Christoffer Granmo, and Morten Goodwin. Convolutional regression tsetlin machine: An interpretable approach to convolutional regression. In *2021 6th International Conference on Machine Learning Technologies*, pages 65–73, 2021.

[18] Maryna Lukach, David Dufton, Jonathan Crosier, Joshua M Hampton, Lindsay Bennett, and Ryan R Neely III. Hydrometeor classification of quasi-vertical profiles of polarimetric radar measurements using a top-down iterative hierarchical clustering method. *Atmospheric Measurement Techniques*, 14(2):1075–1098, 2021.

[19] Sebastian B Thrun. Efficient exploration in reinforcement learning. 1992.

[20] Li Deng. The mnist database of handwritten digit images for machine learning research. *IEEE Signal Processing Magazine*, 29(6):141–142, 2012.

[21] Ketan Rajshekhar Shahapure and Charles Nicholas. Cluster quality analysis using silhouette score. In *2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*, pages 747–748. IEEE, 2020.

[22] Amit Saxena, Mukesh Prasad, Akshansh Gupta, Neha Bharill, Om Prakash Patel, Aruna Tiwari, Meng Joo Er, Weiping Ding, and Chin-Teng Lin. A review of clustering techniques and developments. *Neurocomputing*, 267:664–681, 2017.

F

Appendix G

Paper G

Title:	Novel Users' Activity Representation for Modeling Societal Acceptance Towards Misinformation Mitigation on Social Media
Authors:	Ahmed Abouzeid, Ole-Christoffer Granmo, Morten Goodwin, Christian Webersik
Affiliation:	University of Agder, Faculty of Engineering and Science, P. O. Box 509, NO-4898 Grimstad, Norway
Journal:	*Journal of Computational Social Science*

Novel Users' Activity Representation for Modeling Societal Acceptance Towards Misinformation Mitigation on Social Media

Ahmed Abouzeid, Ole-Christoffer Granmo, Morten Goodwin, Christian Webersik
Department of Information and Communication Technology
Faculty of Engineering and Science, University of Agder
P.O. Box 509, NO-4898 Grimstad, Norway
E-mails: {ahmed.abouzeid, ole.granmo, morten.goodwin, christian.webersik}@uia.no

Abstract **— Intervention-based mitigation methods have become a common way to fight misinformation on Social Media (SM). However, these methods depend on how information spreads are modeled in a diffusion model. Unfortunately, there are no realistic diffusion models or enough diverse datasets to train diffusion prediction functions. In particular, there is an urgent need for mitigation methods and labeled datasets that capture the mutual temporal incidences of societal bias and societal engagement that drive the spread of misinformation. To this end, this paper proposes a novel representation of users' activity on SM. We further embed these in a knapsack-based mitigation optimization approach. The optimization task is to find ways to mitigate political manipulation by incentivizing users to propagate factual information. We have created PEGYPT, a novel Twitter dataset to train a novel multiplex diffusion model with political bias, societal engagement, and propaganda events. Our approach aligns with recent theoretical findings on the importance of societal acceptance in information spread on SM as proposed by Olan et al. (2022) [1]. Our empirical results show significant differences from traditional representations, where the latter assume users' exposure to misinformation can be mitigated despite their political bias and societal acceptance. Hence, our work opens venues for more realistic misinformation mitigation.**

G.1 Introduction

In the past decade, Farajtabar et al. (2016) introduced a novel intervention-based approach for incentivizing users on social networks to change their activities [2, 3]. In this paper, we focus on applying such intervention to mitigate misinformation[1] on SM with Reinforcement Learning (RL) agents [5, 6, 7]. The purpose is to intervene with the network and learn how to incentivize users to propagate factual information, the latter is also known as truth campaigning. In brief, each RL agent monitors a single user i and must persuade those in her friend or follower zone (e.g., users j and k).

Consider the example when a user k is exposed to misinformation because of manipulative influence from user l. To counter the influence user l has on user k, user i would have to exert effort to persuade user k. Consequently, user i should be incentivized to propagate sufficient counteracting information. Conversely, other users with less victimized friends or followers may, to a smaller extent, need such incentivization.

In practice, the capacity to incentivize users may be limited. Then a problem arises if user i misspends the available incentivization budget, e.g., to convince both j and k when only k needs incentivization [8, 6]. To address this problem, fairness-based mitigation techniques [8] were proposed to ensure that all network users make better use of a limited incentivization budget.

Intervention-based misinformation mitigation approaches usually utilize an information diffusion model [9] to predict the network dynamics and user propagation patterns at a specific discrete time window in the future. Hence, the RL agents can learn about optimal incentivization policies from the simulated environment by the diffusion model [6]. These RL agents, when interacting with the diffusion environment, form a control over the simulated dynamics of users' activities. The control model tries to optimize a loss function by learning optimal incentivization strategies under budget constraints. This task was solved as a multi-agent knapsack optimization problem [10, 5, 8]. However, there are still some research gaps and open questions to obtain a more realistic information diffusion and misinformation mitigation. Therefore, we discuss below some of the main challenges.

(A) Information Diffusion Accuracy: Information diffusion models are necessary for intervention-based misinformation mitigation since applying and evaluating multiple intervention strategies on the real social network is infeasible. In a diffusion model, some users' activities are predicted to simulate and mimic the network dynamics. Traditionally, for the problem of online misinformation mitigation, these activities are the users' temporal propagation of misinformation, and normal information [6]. However, that comes with a challenge, as the model would have an inaccurate prediction of the real-world network propagation of these activities.

[1] The term misinformation is sometimes used to refer to all forms of fake news/content. However, in some literature, misinformation is defined as the unintentional spread of false content, while disinformation is the on-purpose spread [4]. In this paper, we refer to all forms of false content, including political propaganda, as misinformation.

Moreover, such critical drawbacks, if occurred, will also affect the veracity of the optimization decisions made by the control model.

(B) Predicting Users' Engagements: SM users' engagements occur when users like, comment, or repropagate other users' content. Extending an information diffusion model to simulate more patterns, such as user engagements, is a clear advantage. For instance, that would answer an important and open research question: how do we model the incidence of engagement between those who spread misinformation or victimized by it — and those who would be incentivized to propagate counterinformation? The latter question is important because political bias can cause people not to be interested in engaging with other ideas when incentivized to do so. Consequently, the learned incentives would be meaningless and not represent the actual behaviors on the network. To that end, users' activity representation in diffusion models must be studied wisely. Therefore, which network propagation attributes should be included in such representation becomes a fundamental question to obtain a robust solution for the problem.

(C) Insufficient Representation: The currently proposed solutions for mitigating online misinformation [6, 11, 7] suffer from invoking all critical network features in the mitigation and diffusion models. The latter drawback exists because of the limitation in the available datasets [12], from which diffusion models construct the diffusion prediction function as well. That shows how these datasets [13, 14, 15, 16, 17, 18] do not reflect on the advances from the social science, a field where the problem of societal interaction is significantly relevant. For instance, a recent study illustrates how societal acceptance [1] on social networks can determine the level of effectiveness when introducing factual information. In the latter study, Olan et al. (2022) proposed a conceptual framework of how the concept of societal acceptance explains how communities on SM form societal circles and deny the acceptance of outliers. A societal circle can be defined as a societal bubble on SM [19] where a circle is a group of biased users agreeing on a particular opinion or idea.

Thus, modeling the mitigation over sequential misinformation diffusion needs modeling of temporal activities such as societal circles formulations (e.g., when users agree and engage with particular ideas), societal bias incidents (e.g., when users propagate something of a particular bias), and misinformation (e.g., when users propagate false information).

G.1.1 Contribution and Paper Organization

This work addresses challenges **B** and **C** as introduced above, which indirectly contributes to the challenge **A** since more realistic network dynamics representation could lead to more accurate information diffusion. We propose modeling the diffusion of users' engagements, misinformation/normal information, political bias, and societal circles. Hence, our proposed diffusion model is a multiplex diffusion model [20] where multiple interconnected and interdepended diffusion groups interact. Our hypothesis is as follows. What governs users' activity and how their discussions go is a universe of ideas. Previous studies on the so-called SM filter bubbles [21] can

support the latter assumption. These bubbles' associated ideas construct societal circles, where each circle gathers a subset of people inside it by engaging with the concept it represents. Overlapping between circles may exist, but that does not mean no extreme polarization between them could happen simultaneously.

Polarization causes some of these circles to produce misinformation, which persists in such polarization according to how the SM platforms' algorithms are designed [22]. From there, misinformation circulates through these circles with varying degrees of influence. As a result, reducing polarization and misinformation requires weakening the circles that cause or are influenced by misinformation more than others.

In our proposed solution, a harmful circle is weakened when the number of people engaged with its underlying ideology is significantly reduced. Such counts and their variety are obtained from an information diffusion model that predicts temporal activities such as the propagation of authentic content, misinformation, political bias, and societal circle engagements. We highlight this paper's main contributions below while providing open access to the control model source code[2] and dataset with its post-processing scripts and documentation[3]:

- We introduce PEGYPT, a novel misinformation dataset with temporal labels on political propaganda, bias, and societal circle formulation dynamics.
- Based on the above dataset, we introduce a novel technique to represent users' activities on social networks with our proposed Multiplex-Controlled Multivariate Hawkes Processes (MCMHP) diffusion model.
- We propose a novel optimization loss function that takes temporal bias, propaganda, and information from societal circles as part of its domain and guides the control model reward function.
- We extend the recently proposed intervention-based misinformation mitigation algorithm [8] to support scaling up network size through Monte Carlo-based point process simulation with a small sample size. Further, we couple our novel loss function with the algorithm.
- We provide both quantitative and qualitative analysis to show different behavior between a recently introduced misinformation mitigation loss function [8] and our proposed loss function that considers more convenient domain attributes such as societal bias and societal circles.

The rest of this paper is organized as follows. section G.2 provides a literature review, briefly explaining related technical details. Then, section G.3 illustrates the proposed misinformation mitigation loss function, a novel multiplex information diffusion model, a novel simulation technique, and a novel dataset with multiple temporal events. Empirical results, evaluation, and analysis are given in section G.4.

[2]https://github.com/Ahmed-Abouzeid/MMSS_extended
[3]https://github.com/Ahmed-Abouzeid/PEGYPT

A brief discussion about our proposed approach and its limitations and concerns is given in section G.5. Eventually, section G.6 concludes the work and suggests future directions on the topic.

G.2 Related Work

The problem of misinformation propagation on SM has attracted attention in the past decade. Both technical and philosophical efforts were made to investigate the nature of the problem, its fundamental concepts, and potential solutions. For instance, recent studies investigated the negative impact of misinformation on society and how SM providers are taking action to reduce misinformation propagation [1]. The latter study highlighted the importance of the societal acceptance concept and its association with online content and SM platforms. In that manner, they stated how the social network assembles ideological sub-networks or circles which try to attract people who share similar values and increase the propagation and polarization inside these common circles. Further, these circles clamp down on outsiders who question or oppose these circles' values.

Moreover, psychological inoculation improved resilience against misinformation on SM [23]. The latter approach applied interventions to users to inform them about the manipulation techniques so they could distinguish fake content from authentic one. In the latter study, one of the primary purposes was to focus on reducing misinformation susceptibility rather than stopping it. The latter scenario of mitigation rather than stopping is more realistic since the nature of the technology makes it impossible to stop the propagation of misleading content completely. For example, in political contexts and bubbled online discussions on SM platforms, the confirmation bias makes people believe in what is aligned with their political beliefs no matter how authentic it is [24].

As a proposed technique for a wide range of tasks, Artificial Intelligence (AI) was utilized [25] to address the problem of online misinformation. There are two tasks where AI can be utilized for the problem. On the one hand, it is the classification of misinformation, and on the other hand, it is the mitigation of misinformation exposure and its influence on SM users.

There are different approaches being adopted for the misinformation classification task. For instance, content-based [26] Machine Learning classifier approach focused on extracting the textual features of online circulated news articles and their headlines. In the latter approach, word embedding techniques were adopted to represent the semantics of the article's contents. In addition, these features could be derived from visual information like typical images, comics, or deceptive pictures. Such multi-modal approaches took advantage of the combination between text and image-based features and showed more efficient detection for some applications [27].

Further, fake news detection based on contextual information was widely adopted in the literature [28]. In the latter, the content representations considered the co-occurrence between a word i and the context word j instead of only relating words

to a whole article or content. Additionally, the social context was modeled by connecting publisher-news relations and user-news interactions simultaneously [29]. The latter technique improved the detection performance in some applications as well.

Despite the significant enhancements from the above-mentioned efforts, different challenges [26] stand against fake news detection. For instance, detecting unseen events became an obstacle since news events would have unseen features during the training of the original classifiers. Furthermore, noisy multi-modality is possible since fusion mechanisms would generate inefficient representations. More importantly, adopting detection approaches is essential, but more is needed — because judging online content or users' authenticity would violate freedom of speech [30]. That is, it became challenging in political contexts to draw a sharp line between what is fake and what is not. Hence, more democratic approaches were needed and proposed as we discuss below.

Recent utilization of RL methods on the problem of online misinformation showed that learned policies that expose social network users to factual information would significantly mitigate the effect of misinformation [6, 5, 8, 7]. The mitigation approach can be considered an extension of the detection approach since its first task is to learn users' activity patterns from the classified historical events on a SM platform. A common mitigation technique is truth campaigning [8] where the purpose is to learn an optimal mitigation strategy that incentivizes network users to ensure the optimal delivery of factual information to everybody on the network.

There are different proposed incentivization techniques. For example, the latter can be delivering personalized verified news articles to suit users' reading preferences [31]. However, in RL-based truth campaign, the typical way of incentivization is to learn about the number of incentives per user that would acquire her to accept propagating the verified information on the network [6, 5, 8]. In the latter approach, based on these optimal incentives, the mitigation model ensures a maximal delivery of the authentic content which would achieve maximal mitigation.

The utilization of the RL framework means that an intervention with the network users is conducted. The intervention procedure allows the RL agents to learn about the user's activity. These activities are simulated with an information diffusion model, commonly a Hawkes Process (HP) [32].

The temporal activities of users on SM are usually logged with annotations in datasets that are used to train an information diffusion prediction function [6, 5]. The latter function predicts the information type it was trained on, e.g., misinformation or authentic content activities. Unfortunately, the available datasets [13, 14, 15, 16, 17, 18] need more enriched users' activity information. For example, modeling users' activity only through their dissemination patterns of either true or false content does not inform the diffusion model about other aspects, such as political bias and societal engagement. Hence, this paper proposes a novel representation of users' activity, where temporal patterns of bias, societal engagements, and content authenticity were considered when modeling the RL agents' interventions.

G.3 Methodology

This section gives a detailed introduction to our novel users' activity representation and how to utilize that for a solution of misinformation mitigation on SM. Hence, subsection G.3.1 demonstrates our proposed users' activity dataset's collection and annotation processes. Further, subsection G.3.2 explains our proposed MCMHP architecture that encapsulates these representations in a more realistic information diffusion model for how misinformation and other interconnected events circulate over the social network. subsection G.3.3 and subsection G.3.4 respectively illustrate our novel optimization loss function and our proposed idea of controlling users' activity variables to help optimize the loss, achieving misinformation mitigation and a societal acceptance boost. Eventually, subsection G.3.5 explains our simulation technique for large-scale networks and how we determine the truth campaign incentives for each user.

G.3.1 PEGYPT Dataset

The dataset samples were collected from Egyptian Twitter hashtags which discussed the Egyptian presidential election, 2018. The data were extracted using Twitter API between 24^{th} and 27^{th} of March 2018, a few days before and during election voting days. The data samples represented three categorized temporal events: political bias, societal circles engagement, and political propaganda. The forms of these events varied between original tweets, quoted tweets, retweets, and replies from all associated hashtags (check hashtags details here[4]) in the Arabic language. The final numbers for users and events were 10,534 and 36,390, respectively.

Egyptian specialists manually annotated these temporal events while following a systematic approach to ensure consistency when evaluating text or media content in a tweet. That was achieved by establishing predefined keywords and some combinations of the latter (check for details here[5]), so when they exist — a particular judgment (label) is made to the content and overwrites the human given label. Hence, during the annotation process, we aimed to avoid the human bias factor [33] by these predefined keywords that were agreed upon as a code for either propaganda or political bias. For instance, if a content had religious statements and keywords, it was considered as political propaganda, even if the annotator is religious. The latter process yielded around 20% of the events' labels based only on these keywords. We named the dataset as "PEGYPT", abbreviated from the terms: ***P****olarized,* ***Egypt****.*

In the below sub-sections, we describe the criteria for how we annotated the three categorical events. Further, unlike the limited and static labels in the existing datasets [13, 14, 15, 16, 17, 18], we highlight how our novel approach for temporal labeling of social network events opens the venue for an extended analytical capacity of the information diffusion and mitigation tasks, as explained later in section G.4. For example, when labeling temporal changes of political bias and societal engage-

[4]https://github.com/Ahmed-Abouzeid/PEGYPT/blob/main/tags.txt
[5]https://github.com/Ahmed-Abouzeid/PEGYPT/blob/main/annotate_propaganda.py

ments, we could trace how these variables evolve during an intervention-based truth campaign. Finally, we provide statistical properties of the collected social network data to give a better understanding of the context for our experiments.

G.3.1.1 Temporal Bias Label

The temporal bias label had three possible values which evaluated whether the user-created or engaged-with content was neutral (0), biased towards (1) or against (-1) the election process. Unlike assuming a static political bias for users [11], our temporal bias captures whether users changed their opinions when they engaged with or generated content over time and hence — had contradicted bias between different contents. That helped to trace the changes in the frequencies of bias levels during the conduction of the truth campaign and the misinformation mitigation. The latter guided learning more realistic incentives based on traced users' willingness rather than assuming they would accept whatever incentives they would be offered.

G.3.1.2 Temporal Propaganda Label

The temporal propaganda label illustrated the temporal patterns of users with regard to sharing politically manipulative content, in the following sense. The label had two possible values, which described whether the content was political propaganda (1) or not (0). The criteria for the latter were based on whether a user misled readers by using religious expressions or misleading propaganda to manipulate the facts.

G.3.1.3 Temporal Societal Circle Label

We defined societal circles as the finite set of different ideologies in online content, where each user engages with one or more ideological circles over time. In that context, an engaged user of a content ideology means a user who created, quoted, retweeted, liked, mentioned, or replied to that content. The PEGYPT dataset had six societal circles, where a circle ideology in content was defined according to the combination of bias and propaganda labels values. Figure G.1 and Figure G.2 give a better idea of how these combinations constructed the circles.

We extracted the temporal circle information from each content to obtain the temporal incidents of societal circles events (i.e., temporal labels). In that manner, for each content, we extracted the ideology of the content and associated it with the content creation time. On the same content, we further extracted other engagement forms with their creation times and associated ideologies. For example, when a user generates a primary tweet with a particular bias and authenticity level, we consider that as a particular temporal societal circle event. If the latter content had other engagements such as replies, mentions, likes, retweets, or quotes, we consider other temporal societal circles accordingly. Thus, a societal circle becomes a structure that changes its density through time (i.e., the number of users represented in its predefined ideology changes over time).

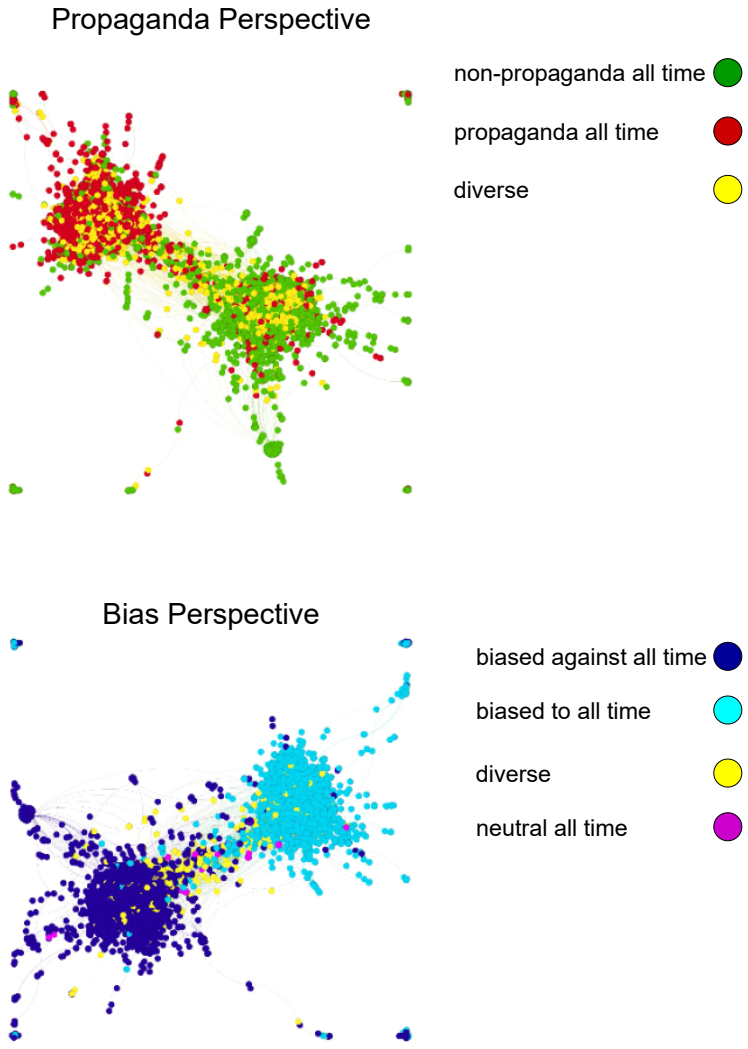


Figure G.1: Colored graph from the PEGYPT network dataset, where nodes and edges represent users and their engagement, respectively. Colors represent the propagation over time of a particular content type

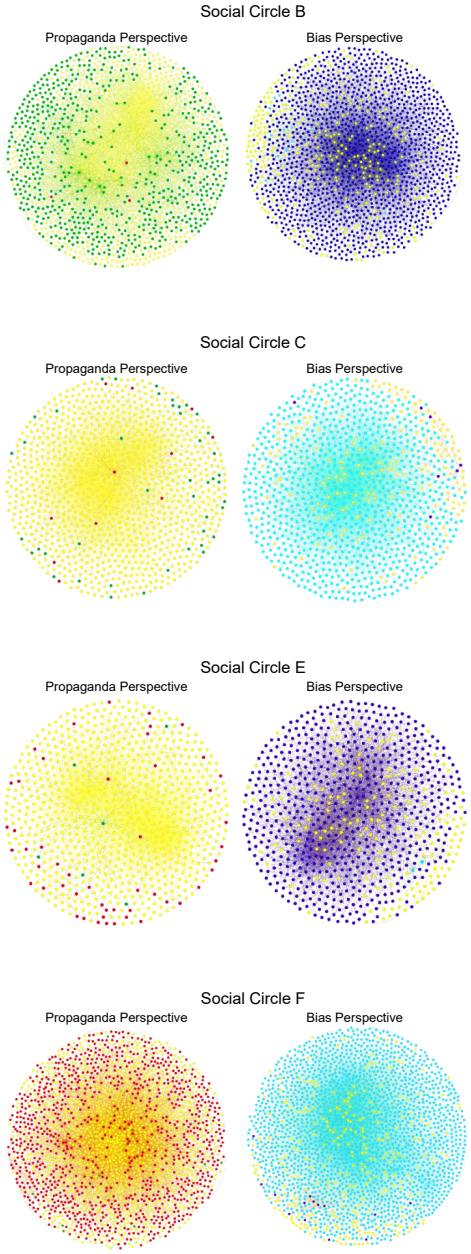


Figure G.2: Extention to Figure G.1: Colored graph from some PEGYPT sub-networks which represent most populated societal circles. Nodes and edges represent users and their engagement, respectively. Colors represent the propagation over time of a particular content type

Modeling the temporal changes of societal circles structure is essential to characterize a wide range of the network's user activities. That is why a user's engagement with a societal circle was defined according to whether that user liked, replied, retweeted, quoted, or was mentioned in a tweet belonging to that particular circle concept. That means, engaging with a societal circle did not necessarily mean agreeing with its underlying idea. The reason behind that approach is that we wanted to trace users' exposures to online content realistically, and timestamped users' interaction was the tangible measure we could have found. That was different from previous misinformation mitigation methods [6, 11, 8] where the exposure measures were considered based on the network connections (e.g., following relationships), regardless of whether an interaction will occur. The latter matter significantly impacts how the mitigation incentives could be decided since the mitigation algorithm highly depends on content exposure calculation and will be learning from unrealistic estimation if the latter is not appropriately modeled.

G.3.1.4 Dataset Details

Table G.1, Table G.2, and Table G.3 show some statistics of the PEGYPT dataset and some example criteria keywords for determining propaganda and bias labels, respectively. According to Table G.2, some keywords were observed in the collected dataset samples and indicated political manipulation and propaganda. These keywords could also be related to different bias directions since manipulation on the network was from both sides. The made-available dataset files provide the complete details of the associated hashtags and all criteria keywords for both propaganda and bias labels.

To show how misinformation manifiested in the collected social network, Figure G.1 shows colored graphs from the two perspectives: political bias and political propaganda, where nodes colors represent how individual users circulated their contents. Figure G.2 also demonstrates an example of the same perspectives for the most crowded societal circles, where some circles were more harmful than others. The complete details of all societal circles and their ideological concepts can be viewed in Table G.4. Hence, we can observe how the societal circle F was the most harmful to the top population.

According to Amnesty's reports on Egypt's human rights situation and witnesses about how fake the election process was[6], we evaluated our mitigation model for the scenario of breaking circle F by incentivizing its members to join other unharmed circles, such as circle B, which was a non-propaganda circle that opposed the election. Thus, in our experiments, we considered the mitigation campaign to oppose the election itself. Further, the mitigation campaign must do that without spreading propaganda to manipulate the public. Hence, a circle with same bias as our mitigation campaign such as circle E was also considered harmful because it is a propaganda circle.

[6]https://www.amnesty.org/en/latest/news/2018/01/egypt-authorities-must-cease-interference-in-upcoming-election-and-set-guarantees-for-free-candidacy/

Table G.1: PEGYPT dataset statistics

Metric	Value
Total population	10,534
Number of events	36,390
Number of societal circles	6
Number of graph edges	22,058
Graph modularity	0.596
propaganda users(%)	0.336
Non-propaganda users (%)	0.471
Variant propagation users (%)	0.193
Only biased to-users (%)	0.435
Only biased against-users (%)	0.513
Only neutrally biased-users (%)	0.009
Variant bias users (%)	0.043
Propaganda events (%)	0.536
Non-propaganda events (%)	0.463
Biased to-events (%)	0.515
Biased against-events (%)	0.478
Neutral bias-events (%)	0.006
Biased to + propaganda events (%)	0.429
Biased to + non-propaganda events(%)	0.087
Biased against + propaganda events (%)	0.107
Biased against + non-propaganda events (%)	0.371

Table G.2: Example keyword(s) for the "Is-Propaganda=1" label

Keyword(s)	Translation
حق الشهداء	For the sake of maryters
حرب اهلية	Civilian war
تتحول لسوريا	To become like Syria
تبقي اد الدنيا	To become superior over all the world
خونة	Betrayals
عميل لأمريكا	American agent
الله تعالى يقول	God says
ناشط عميل	A betrayal political activist
كلام الله	Words of god

Table G.3: Example keyword(s) for the bias label

Keyword(s)	Translation	Label
انزلوا يا مصريين	We Egyptians must go and vote	1
الانتخابات فرحتنا	This election is our joy	1
اختار رئيسك	Choose your president	1
نازلين نكمل المشوار	We will vote to continue the way	1
الزم بيتك	Stay home	-1
بلحة	Sarcastic nickname of the presidential only candidate	-1

Table G.4: Societal circles concepts and population

Circle	Concept	Population
A	neutral bias + non-propaganda	86
B	bias against + non-propaganda	1,581
C	bias towards + non-propaganda	992
D	neutral bias + propaganda	2
E	bias against + propaganda	775
F	bias towards + propaganda	2,084

G.3.2 Information Diffusion Models

To facilitate an intervention environment for the RL agents to learn about users' activity, an information diffusion model is required to simulate the dynamics of social networks. We simulated the latter using a Multivariate Hawkes Process (MHP). A MHP is a multivariate point process [34] that models the occurrence of temporal or spatiotemporal asynchronous events by capturing the self-and/or mutual excitation (dependencies) between these events. In our context, the MHP is multivariate over the network users.

Through users' activity across the temporal information collected from the PEGYPT dataset, each user was represented by a multiplex HP [20] to predict her future activity on different diffusion groups. In that manner, the diffusion groups represented the temporal patterns of propaganda, non-propaganda, bias towards, bias against, neutral bias, and, eventually, all societal circles' engagement events. Therefore, for each diffusion group, there was a MHP for all users, and the relevant group events from PEGYPT data were used to train the diffusion group prediction function over its users.

The associated user HPs are volume-based diffusion models which generated estimated random counts for all event categories, given some activity observations in the past. These counts indicated the intensity of the process at a specific time of realization. Hence, a HP for user i can be defined for any diffusion group with its conditional intensity function λ_i. The intensity function has two main components, base intensity μ_i, and an exponential decay kernel function g over an adjacency matrix A. The formal explanation of the conditional intensity function is given by Equation G.1.

$$\lambda_i(t_r|H^{t_r}) := \mu_i + \sum_{t_s<t_r} g(t_r - t_s). \tag{G.1}$$

Where μ_i represents a base intensity that models some external motivation to propagate some content. g is some kernel function over the observed history H^{t_r} associated with user i from the discrete-time realization t_s prior to time t_r. g is concerned with the history of some influence matrix A, where $A_{ij} > 0$ if there was an inferred influence between user i and user j, and $A_{ij} = 0$ if not. We utilized an exponential decay kernel function $g = A_{i.}e^{-wt}$, where w is the decay factor where $1 > w > 0$, and represents the rate for how the influence decays over time. For all users, the base intensity vector μ and the influence matrix A were estimated using the maximum likelihood algorithm for the HP [35].

To model the intervention-based mitigation across all diffusion groups, a MCMHP was created, where different diffusion groups were controlled and predicted by a network of Learning Automata (LAs) [8] and MHPs, respectively. We discuss the details of the LAs network in subsection G.3.3.

The diffusion groups encapsulated our proposed novel representation of users' activity on SM. They characterized the interdependence between information veracity-related events, societal bias levels, and societal engagements-related events. Analogically to current approaches of misinformation mitigation [6, 11, 7], Figure G.3 shows our design of diffusion and control models interaction, compared to the typical existing design as shown in Figure G.4.

To evaluate the MHPs predictions, we compared the predicted counts for a diffusion group for all users with the real counts on a test dataset. Therefore, and as shown in Equation G.2, an absolute average error ϵ was calculated to measure how close to reality a MHP prediction was. Where n is the number of users and $N^{\mathcal{H}}$, $N^{\mathcal{R}}$ represents the counts of the arrived events from Hawkes prediction and real data, respectively. The calculation was made between the time stages t_s and $t_s + \Delta$.

$$\mathcal{E}_{t_s+\Delta} = \frac{1}{n}\sum_{i=1}^{n} |[N_i^{\mathcal{H}}(t_s + \Delta) - N_i^{\mathcal{H}}(t_s)] - [N_i^{\mathcal{R}}(t_s + \Delta) - N_i^{\mathcal{R}}(t_s)]|. \tag{G.2}$$

Figure G.5 demonstrates how we organized the temporal diffusion group's associated samples from the PEGYPT dataset — to train (estimate μ and A) the MHP where the temporal events counts per user were aggregated into ordered discrete time realizations.

The core idea behind a MHP-based mitigation task is to intensify a particular event in a diffusion group to produce more occurrences against another harmful event type(s). Users-associated HPs for the to-be-intensified event category should be modified to achieve that. Hence, let s_i be the incentivization amount decided for user i, and the modified HP for mitigation purposes can be redefined by Equation G.3.

$$\lambda_i(t_r|H^{t_r}) := s_i + \mu_i + \sum_{t_s<t_r} g(t_r - t_s), \tag{G.3}$$

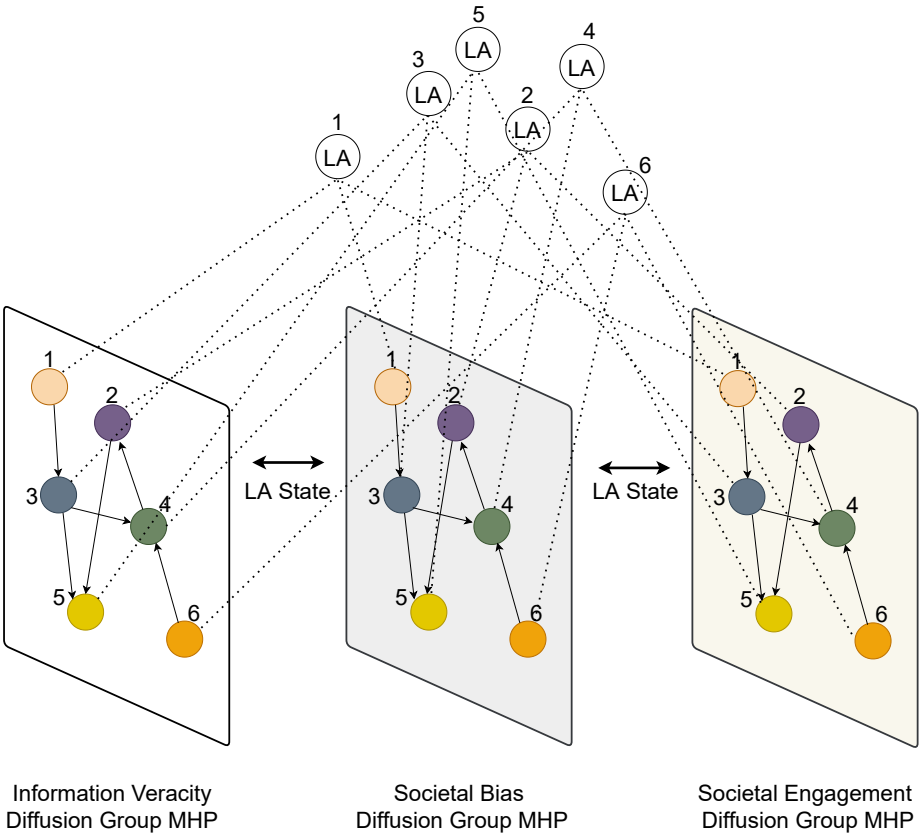


Figure G.3: A toy example of a social network with 6 users and the proposed design of MCMHP interaction, where each LA state is shared between all diffusion groups

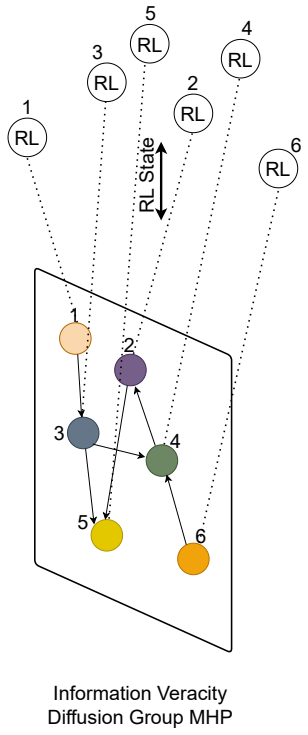


Figure G.4: A toy example of a social network with 6 users and the typical design of MHP interaction with a control model

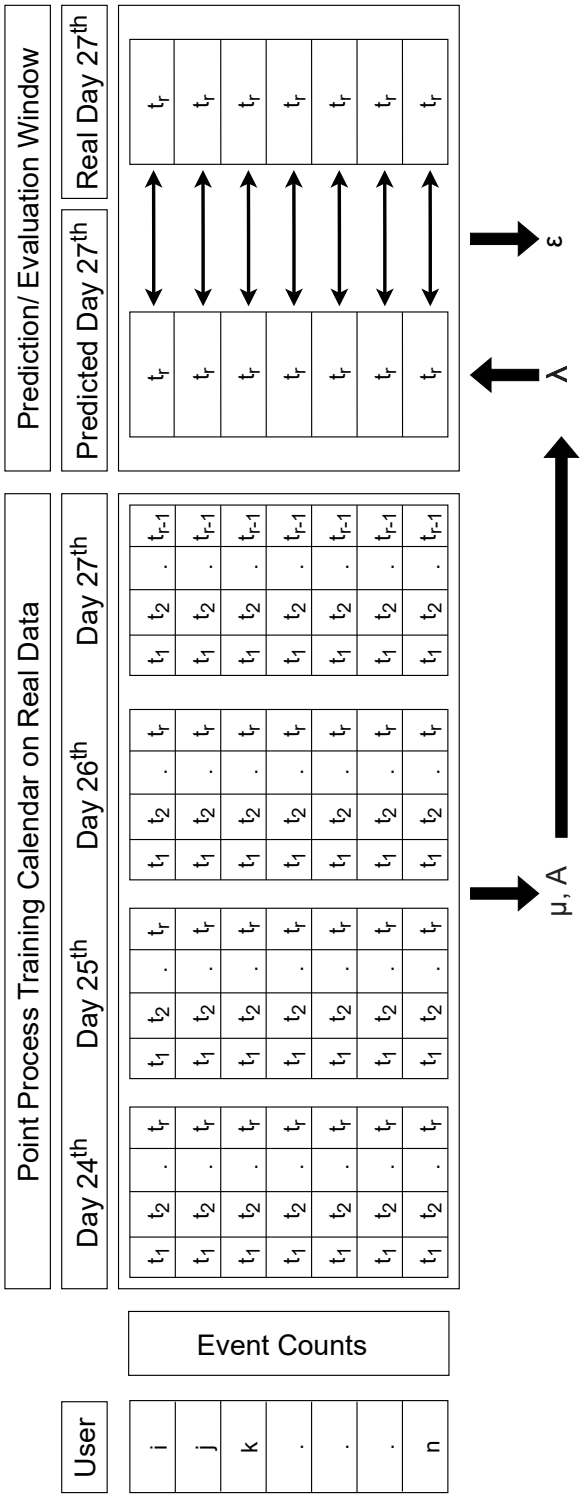


Figure G.5: Feeding the MHP with a diffusion group's samples from the PEGYPT dataset

where s_i is a parameter to be optimized for user i for the mitigation targeted event category, the optimal values for s across all users were governed through a restricted incentivization budget knapsack optimization and a loss function to dictate the rewarding of a RL agent, associated with each user. In our solution, we associated a LA for each user as the individual RL agent that learns the optimal value of s_i.

G.3.3 Controlling of Multiplex Diffusion Groups

As a control model over the stochastic MHP environment, we utilized the LA [36] for its easy decentralized implementation and lightweight computation when compared to traditional RL techniques adopted for the probelm of misinformation mitigation [6, 7]. The LA learns by interacting with the MCMHP and updates its actions or state transitions according to the stochastic signal from a MHP counts-based loss function. In our proposed solution, each LA is attached to each user to learn an optimum/sub-optimum state s^*, where the latter represents a discrete decision value for each user's incentive in the mitigation campaign. As indicated in Figure G.3, these incentives (LAs states) are shared across all diffusion groups to embrace the interdependencies between the different aspects of users' activity. The LA seeks convergence at such an incentive value by optimizing the latter through the loss function. The latter dictates the potential reward or penalty of the LAs, and the LAs updates their states accordingly. The loss function evaluates its gradient when a LA increases its state and causes new predicted volumes from the different diffusion groups in the MCMHP. Hence, if the loss slop declined, then the LA should be rewarded. If inclined, the LA should be penalized.

Figure G.6 demonstrates how challenging optimizing such loss function through each associated LA state transition, where optimal states could be non-stationary due to the interdependencies and complexity between all diffusion groups, i.e., some users' optimum value s^* will determine the optimum s^* for others. The latter property persists due to the mutual dependencies between users on the MHP-generated dynamics. That means how many incentives a user i would need could make it unnecessary for an engaged user with user i to have many incentives when both share the same other engaging users. Therefore, for user i, if s_i is optimum and consequently, for user j, s_j is not. Then, given $s_i^{'}$ as another possible incentive value for user i, $s_i^{'}$ could still be optimum while s_j is also optimum. This non-stationarity occurs then because the LAs while consuming the incentivization budget — have to intervene with their associated users in sequential order, not simultaneously.

When saddle points occur, particular LA state transitions and rewarding techniques are applied as proposed by Abouzeid et al. (2022) [8] since we utilized the same LAs network-based control model. The complete details of how each LA learns its incentive value and updates its state transition probabilities are given in Appendix H of this paper.

We extended the LAs network-based control model with a Monte Carlo simulation technique [37] over multiple interventions $\{e^1, e^2, ..., e^*\}$. Hence, the con-

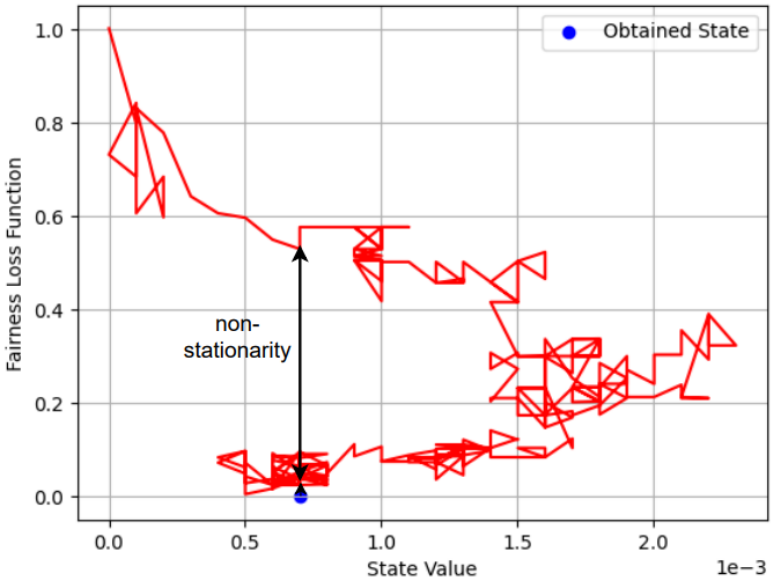


Figure G.6: Non-stationarity of an optimal automaton state in its individual loss function trajectory over time

trol model could interact with random samples instead of the whole network. We repeated the latter procedure over hundreds of sampled networks, and then we calculated the expected values of individual converged s^* values over all samples. We believe this approach opens the venue for scaled-up misinformation mitigation frameworks where the real network size would reach hundreds of thousands of users and more. Moreover, obtaining a probability distribution for a user's decided incentive allowed us to measure the level of uncertainty in the solution (see subsubsection G.4.3.1).

G.3.4 Optimizing Societal Acceptance with Fairness

The criteria for successful mitigation were based on how eventually the sampled network users would be less exposed to the harmful content since the incentivization should boost the amount of authentic content on the network. Therefore, the optimization task was to reduce a total loss function during the intervention. To achieve the latter, a LA per user conducted the intervention by suggesting a shared incentive value to modify the associated HP diffusion group by which the political manipulation would be mitigated. Thus, we wanted to incentivize the diffusion groups' events of non-propaganda, bias against, and societal circle B with the same shared incentive value.

During learning, after an intervention step e, a dedicated individual loss function is responsible for the evaluation of the current incentive of its user. At the same time, the MCMHP predicted temporal events information was given as the function domain. Hence, for an individual user i, all other users' predicted activities (from all diffusion groups) were passed to the individual loss function i. Ideally, the total loss function should converge to a steady point after multiple interventions across

all users.

We extended the fair mitigation loss function introduced by Abouzeid et al. (2022) [8]. In the latter, the distribution of incentives was conducted according to user needs. In that manner, we keep maintaining the concept of fair incentivization. However, in addition to representing only temporal events for misinformation and authentic content, we propose additional information on the temporal societal circles and temporal bias, to model the occurrence of engagement and its nature, respectively. That means we predict the propagation of authentic content (e.g., non-propaganda), propaganda, engagement of users with societal circles, and eventually, the bias directions of users at a specific intervention step e and time realization t_r. We think such a combination gives more close-to-reality dynamics from diffusion modeling and characterizes the societal acceptance concept that governs social networks [1]. Equation G.4 and Equation G.5 demonstrate our novel loss function.

$$\min \mathcal{F}(s_U) := \sum_{i=1}^{N} \Lambda_i^{'}(s_i) + \mathcal{F}(s_i), \text{ where } \mathcal{F}(s_i) := \sum_{j=0}^{n} (2 - R_j^{s_i} - \Lambda_j(s_i))^2, \quad \text{(G.4)}$$

$$\text{subject to } \sum_{i=1}^{|s_U|} s_i, \text{ where } s_i \in [0, C]. \quad \text{(G.5)}$$

The term s_U represents the set of passed users' incentives, where the sum of the latter set cannot exceed the incentivization budget C as demonstrated in Equation G.5. Further, as indicated in Equation G.4, an individual loss function for a user i is evaluated first by measuring how the incentive value s_i affected all other users with an engagement relationship to i. Hence, the term $R_j^{s_i}$ defines the exposure counts ratios between non-propaganda npg and propaganda pg events for all engaged users with user i at a particular time realization t_r.

A user i exposure to a particular event category (e.g., non-propaganda) at a particular time realization t_r is the count of all events from users engaged with user i at t_r. Equation G.6, Equation G.7, and Equation G.8 show how $R_j^{s_i}$ is calculated, while a user-associated ratio closer to 1 means a boosted non-propaganda exposure. A ratio that exceeds 1 means an unnecessarily high incentive value assigned to that user, which indicates unfairness. In Equation G.6, ξ is a tiny smoothing positive value close to 0 to avoid division by Zero when propaganda events do not exist for some users. Also, b is a mitigation balance factor to satisfy a mitigation campaign threshold. For example, $b = 2$ if successful mitigation means the exposure to non-propaganda should be at least twice the exposure to propaganda, and hence the unfairness happens if the ratio exceeds 2.

$$R_i^{t_r}(s_i) := \frac{\xi + npg_i^{t_r}(s_i)}{(\xi + pg_i^{t_r}) \cdot b}, \quad \text{(G.6)}$$

$$pg_i^{t_r} := \sum_{s=0}^{t_r} \sum_{j=1}^{n} A_{ij} \cdot pg_j^{t_s}, \quad \text{(G.7)}$$

$$npg_i^{t_r}(s_i) := \sum_{s=0}^{t_r} \sum_{j=1}^{n} A_{ij} \cdot npg_j^{t_s}(s_i). \tag{G.8}$$

The term $\Lambda_j(s_i)$ is calculated according to Equation G.9, and it represents a joint probability of two events. First, c_j, which is the probability that an engaged user j with user i would engage with the societal circle to which the mitigation campaign tries to attract people to. Second, the probability that j being in the same bias of the mitigation campaign and is denoted as $bias_j$. It is essential to highlight that such probabilities are calculated after applying the incentives s_i and s_j from the associated interventions, which would change the generated counts for bias and societal engagement HP events. Hence, $\Lambda_j(s_i)$ measured the probability of the societal engagement with the circle we seek acceptance of its concept, and the probability of agreeing with that circle during such engagement.

$$\Lambda_j(s_i) := P(c_j) \cdot P(bias_j). \tag{G.9}$$

While interventions cause different incentives and accordingly different diffusion volumes, given an increased value of $\Lambda_j(s_i)$ will decrease the loss function, and the associated LA_i will be rewarded.

The individual loss for user i could also be increased by $\Lambda_i^{'}(s_i)$, which represents the probability of user i not being in the same bias direction of the mitigation campaign. That means no matter how engaging users with i would agree and engage with the circle we seek — the loss will always be high if user i's bias disagrees with the mitigation campaign. The latter mechanism means that users will consume incentives wisely and according to their probability of accepting the incentives instead of naively assuming they would. Equation G.10 shows how the latter probability is calculated.

$$\Lambda_i^{'}(s_i) := 1 - P(bias_i). \tag{G.10}$$

G.3.5 Monte Carlo Simulation

Let us assume the network sample has n users, where $n = 3$ (see Appendix H of this paper), and each associated LA has the state depth M ($M + 1$ possible incentive values). Then, we can demonstrate the following procedure. Let the users i, j, and k be the sampled network users at intervention step e. Then, $s_U = \{s^e{}_i,\ s^e{}_j,\ s^e{}_k\}$ is the discrete state values of the associated LAs at e. Hence, the converged states and final obtained results from the interventions can be assigned to the below-modified HPs diffusion prediction functions for the three given users. The modified HPs should suggest an optimum or sub-optimum predicted activity on the network if the obtained state values were passed as incentives. The latter should satisfy the minimization of the total loss function $\mathcal{F}(s_U^*)$ in Equation G.4. Where $\sum_{i=1}^{|s_U^*|} s_i \leq C$.

$$\lambda_i(t_r|H^{t_r}) := s^{e^*}{}_i + \mu_i + \sum_{t_s<t_r} g(t_r - t_s), \tag{G.11}$$

$$\lambda_j(t_r|H^{t_r}) := s^{e^*}{}_j + \mu_j + \sum_{t_s<t_r} g(t_r - t_s), \tag{G.12}$$

$$\lambda_k(t_r|H^{t_r}) := s^{e^*}{}_k + \mu_k + \sum_{t_s<t_r} g(t_r - t_s). \tag{G.13}$$

Equation G.11, Equation G.12, and Equation G.13 together construct the incentivized users on the network for a particular event type in a specific diffusion group. Thus, the MCMHP can be viewed as replicating these incentives for all desired events in targetted diffusion groups. For instance, the optimal value s_i is shared across non-propaganda, bias-against, and circle B engagement events to incentivize their associated HPs. Same concept applies for all users.

Since multiple samples are taken during the Monte Carlo sampling procedure, the final determined value for any s is the expected value of the random variable s on its distribution. Hence, for the user i, given a converged random variable $s_i^{e^*}$ from w Monte Carlo samples, the vector $s_i^* = \{s^*{}_1, s^*{}_2, ..., s^*{}_w\}$ represents an example for the possible obtained values from converged automaton LA_i state over w samples, where the user i was sampled w times. Further, the distribution vector $d_i = \{p(s^*{}_1), p(s^*{}_2), ..., p(s^*{}_w)\}$ represents the probabilities for s_i^* entries. Therefore, the final incentive value for a given user i is the expected value for s_i^* over its samples.

Equation G.14 shows how the final incentives were determined, where the final incentivization vector for all users is a vector of all expected values calculated over all sampled networks, where U is the set of all users.

$$s_U^{**} = \{\forall s_i^* \in s_U^* : E[s_i^*] = \sum_{l=1}^{w} s_l^* p(s_l^*)\}. \tag{G.14}$$

G.4 Empirical Results

G.4.1 Experiment Setup

In our experiments, we considered a subset of the PEGYPT dataset where only users with high engagement frequencies were selected. The avoiding of sparsity was necessary for the MHP parameters estimation since the latter requires a sufficient number of events per user. Furthermore, high engagement was essential to study typical social network dynamics where extreme political polarization and propaganda govern the network. Hence, we extracted users with at least six temporal events while keeping similar percentages of propaganda and bias levels as in the original PEGYPT dataset. The final social network had 940 users and 20,084 temporal events. Table G.5 shows the complete details of the obtained social network for the experiments.

We used a sample size of 100 users to construct the sampled networks during the Monte Carlo simulation. We also ran the sampling 100 times to ensure each user will have a probability distribution of the obtained incentives to calculate its expected value. We utilized a time realization period of 180 minutes for the time realization structure. That means events per user (see Figure G.5) were grouped every three

hours and passed to the MHP model. The latter structure helped estimate the MHP parameters as the grouped event counts were enough to infer the influence matrix A and the base intensity μ. We set the Knapsack budget $C = 2$ and LA state depth $M = 500$.

Table G.5: The social network used in the experiments as a subset of PEGYPT

Metric	**Value**
Number of Users	940
Number of Events	20,084
Only Biased Towards-Users (%)	0.44
Only Biased Against-Users (%)	0.50
Variant Bias Users (%)	0.06
Propaganda Events (%)	0.55
Non-propaganda Events (%)	0.45
Max Number of Events per User	177
Min Number of Events per User	6
Number of Societal Circles	5
Number of Graph Edges	6,619
Graph Modularity	0.519
Graph Density	0.015

From the final 940 users' network, we established eleven MHPs to model the different behavioral aspects of the network via a multiplex diffusion. We ran the LA control model on two different environments setup based on two utilized loss functions for the optimization. The latter setup allowed us to monitor how our proposed societal acceptance representation constructed another environmental behavior for the LA control while learning the incentives.

To mitigate the misinformation caused by political propaganda, we incentivized a group of MHPs through the shared incentive value being learned. For example, when a user is incentivized to create or retweet non-propaganda content, the same content declares a particular bias direction. The latter, combined with the non-propaganda content, belong to a specific societal circle as introduced earlier in Table G.4. Hence, the non-propaganda event category was not the only modified MHP — but the relevant events categories for both bias and the particular societal circle were also intensified with the exact amounts represented by the LA state.

To replicate the results and clarify how each MHP was configured and established, we demonstrated the eleven MHPs simulations in Appendix I of this paper, with their customized configurations and purposes. Further, the experimental network had only five circles, as indicated in Table G.5 since circle D had two members only, which was challenging to simulate. Nevertheless, that did not influence the validity of our experiments.

G.4.2 MHP Simulation Evaluation

We ran the eleven simulations and reported their results in Table G.6. We adopted two evaluation metrics to measure how each MHP was reliable enough for the prediction. First, we calculated the average absolute difference error $\mathcal{E}$ as explained earlier in Equation G.2. We then applied the Z statistics to compare the predicted counts with the actual counts.

As indicated in Table G.6, we obtained lower $\mathcal{E}$ and Z values. For more detailed information about the MHPs simulation performance, see Appendix J C of this paper.

Table G.6: MHP simulations performance evaluation with a flag indicating the incentivized MHPs

MHP	**Z**	$\mathcal{E}$	**Incentives**
Bias-towards	0.59 ± 0	0.30 ± 1.31	No
Bias-against	0.55 ± 0	0.35 ± 2.32	No
Bias-against-sampled	0.38 ± 0.25	0.37 ± 1.44	Yes
Propaganda-sampled	0.30 ± 0.24	0.52 ± 1.25	No
Non-propaganda-sampled	0.32 ± 0.28	0.48 ± 1.41	Yes
Circle A	0.02 ± 0	0.02 ± 0.26	No
Circle B	0.27 ± 0	0.23 ± 1.47	No
Circle B-sampled	0.39 ± 0.25	0.37 ± 1.50	Yes
Circle C	0.02 ± 0	0.06 ± 0.62	No
Circle E	0.09 ± 0	0.00 ± 0.07	No
Circle F	0.01 ± 0	0.03 ± 0.28	No

G.4.3 Control Model Evaluation

In this section, we evaluated and compared our proposed loss function to the previously introduced mitigation fairness loss function [8]. We refer to our proposed loss as *Societal Acceptance + Fairness* since the latter still holds the fairness concept when distributing the incentives. At the same time, it is essential to highlight that it was not feasible to evaluate other control models [6, 7] since their structure depends on an entirely different dataset and representations, where temporal bias and societal circles were not modeled. Further, this work's main focus was to assess the novel representation of users' activities. Hence, we utilized the same control model proposed in [8] to extend its fairness loss function with the societal acceptance concept. We employed three evaluation metrics as below.

- **Propaganda Mitigation:** a traditional mitigation evaluation metric [5, 8] to calculate the percentage of how much reduction happened on the users' exposure [6] to propaganda through the engagement relationships with each other. Equation G.15 illustrates how this metric was calculated, where x and y are the political propaganda percentages after and before mitigation, respectively. Hence, the higher this metric, the better.

$$\text{Propaganda Mitigation} := 1 - \frac{x}{y}, \text{ where } x \leq y : y \neq 0. \tag{G.15}$$

- **Polarization Mitigation:** evaluated the percentage of how much harmful polarization was mitigated on the network. For instance, since the campaign task was to convince users against the manipulation in the election, that metric measured how the bias-towards the election among users was lessened after optimizing the incentives. To measure that, the probability distribution over the three bias levels of each user was calculated first, and then we calculated an average percentage of the bias-towards over all users. The latter was calculated twice: once when using the learned incentives to calculate the distribution from the MHP predictions and once when there was no intervention at all. Thus, the polarization mitigation was calculated following the same concept as in Equation G.15.

- **Societal Acceptance Boost:** similar to the above metric, but measuring the percentage of how much the societal acceptance increased on the network. We defined societal acceptance as the breaking of circle F by letting its users accept ideas from circle B. That means we measured the joint probabilities of being engaged with circle B and being biased-against the election (see Table G.4). Similarly to the above metric, we calculated the probability distributions over circle F members to measure how far the intervention succeeded in breaking circle F and allowing its users to accept the societal circle B concept.

Table G.7 shows how our proposed societal acceptance representation outperformed the traditional fairness-only when mitigating polarization and boosting societal acceptance during the misinformation (i.e., propaganda) mitigation. However, we can observe that the percentage of propaganda exposure mitigation was significantly higher than the percentages in both polarization and societal acceptance. We believe that was due to the traditional less strict definition of propaganda content exposure and its mitigation metric [6, 8]. The latter usually consider the counts of events a user is assumed to access through a following/ engaging relationship on the network [6]. However, we believe that would be a naive assumption since following relationships or past engagements do not guarantee actual exposure in the future. Therefore, it was essential to adopt more strict metrics from our proposed representation, such as the actual dynamics of societal acceptance and polarization, which estimated how likely an engagement would occur and to what degree it would be an agreeing engagement inherited from its associated bias. Therefore, our proposed novel representation allowed for calculating the three metrics together, which gave a better justification of the performance.

One of this paper's main motivations and purposes was to analyze the achieved mitigation efficiency to verify what it represented and how the control model learned the incentives. That is because a computational social model evaluation is considered one of the most challenging tasks [38] since the latter lacks a systematic pattern to

Table G.7: Control model obtained performance on utilizing different optimization loss functions. The result is an average over 3 independent runs

Metric	Fairness	Societal Acceptance + Fairness
Propaganda Mitigation	**0.89**±.05	0.88 ± .05
Polarization Mitigation	0.23 ± .10	**0.26** ±.09
Societal Acceptance Boost	0.16 ± .03	**0.19**±.05

consider as ground truth. Therefore, we propose the below analysis to help apply some quantitative and qualitative analysis on the model's performance.

G.4.3.1 Analysis

Figure G.7 and Figure G.8 demonstrate the difference in behavior between the two mitigation loss functions despite their similar propaganda mitigation performance captured in Table G.7. For instance, Figure G.7 on the left side explains how the top 200 incentivized users' engagement was distributed among the different societal circles, i.e., the top most users who consumed the incentivization budget and their societal engagements. We observe that the fairness loss function-based mitigation consumed most of the incentivization budget on users who contributed to around 60% of the engagement in circles F and C. Although circle F was the most harmful circle and circle C also had a different bias than the incentivization campaign (see Table G.4). On the contrary, after considering the societal acceptance representation, our proposed loss function consumed most of the budget on less than 30% of these circles' contributors. The latter behavior indicates how the temporal bias and societal circles' information matured the mitigation more and incentivized users based on the probabilities of accepting the incentive while the fairness loss function incentives were given irrationally.

Moreover, on the right side of Figure G.7, lesser distribution of incentives was the other way around for users engaged with the circles F, C. That indicates how vital these circles were for the incentivization campaign to target and assign more incentives. However, that was done more wisely by the societal acceptance loss function.

Further, Figure G.8 gives an example of how we break the other circles to push users to join circle B by engaging with it and accepting its ideology, not engaging by disagreeing. Hence, we observe how the probability of being biased-against and being engaged with circle B increased. The latter represents modeling the engagement occurrence, while the former models the acceptance of that engagement since circle B represents a bias-against concept. That also demonstrates how representing the temporal bias and societal circles' engagements allowed for tracing and analyzing the associated users' activities of these events.

Eventually, on the left side of Figure G.9, we show how the LAs environment, characterized by our societal acceptance representation, was more strict and gave fewer rewards. We believe that was due to the more interdependent variables considered in the societal acceptance-based loss function. However, such rigidity helped

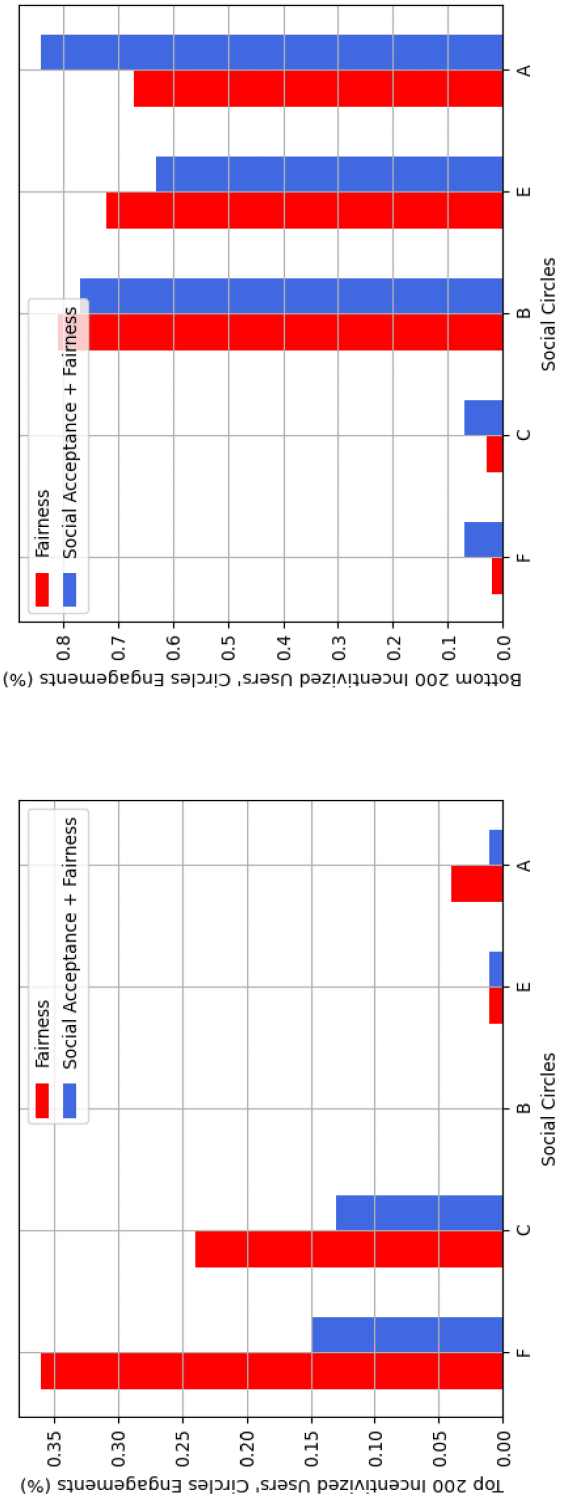


Figure G.7: Incentivized users' circles engagement

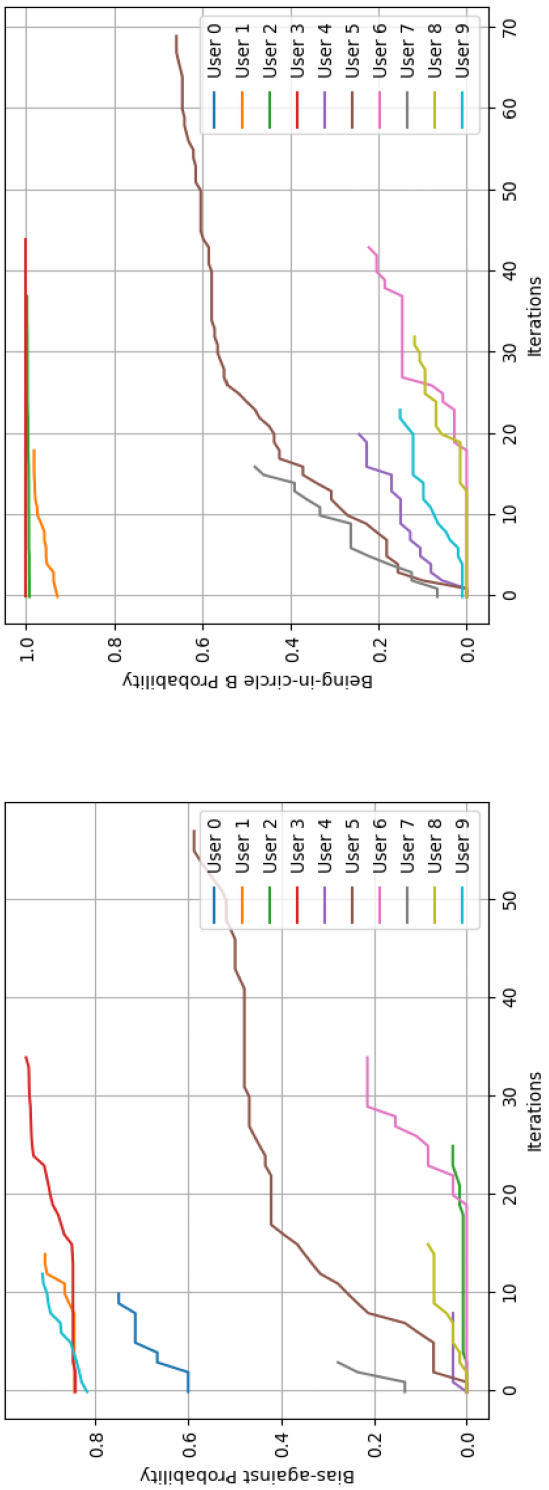


Figure G.8: Example of breaking the societal circles by incentivizing some users to circle B

achieve higher polarization mitigation and societal acceptance in addition to slightly more certainty of the learned incentives. The latter can be viewed on the right side of Figure G.9, where we calculated the Shannon entropy of the individual incentives' probability distribution which was obtained over the Monte Carlo sampling. We can observe there were more users with significant zero entropies when the societal acceptance loss function was applied. The entropies values in Figure G.9 are only for users with different obtained entropies between the two loss functions.

G.5 Discussion

Unlike the recently proposed work [6, 5, 8, 11, 7], instead of directly modeling the misinformation volumes and exposures, and learn incentives accordingly, we first model the relevant network dynamics that derive these exposures. The latter extended the analytical capacity of the solution as deomnstrated in Figure G.8. However, the reader might wonder about the reason behind not modeling societal acceptance directly instead of modeling the bias and engagement separately. That means defining the temporal societal circles based on acceptance rather than engagement in general. Then, modeling the temporal societal circles' acceptance by a HP to predict the acceptance in the next time realization. In the latter scenario, we will lose the capability to trace and analyze the detailed users' activity, such as the interaction with contents, either by agreeing or disagreeing. The latter information is crucial for any further analysis required on the network.

We extracted user engagements from the direct engagement relationship in the historical data to evaluate the influence of incentives during the intervention. However, indirect engagement or influence could also be considered in future attempts. For instance, if user i engages with user j, and user k engages with user j, then users k and i could be considered indirectly engaging together. This influence-cascading technique could also be applied when we consider other influence patterns instead of engagements, such as the following relationships.

G.6 Conclusion and Future Work

The social sciences are studying the societal acceptance concept on social networks to extract the key elements that can describe human behavior regarding information dissemination. Recent efforts revealed the importance of understanding the relationship between fake news, social network platforms, and societal acceptance. Therefore, this paper considers the interdependencies between the latter in a proposed computational social model for mitigating online misinformation. Our proposed model encapsulates novel representations of users' activity, such as temporal polarization patterns, community engagement, and propaganda dissemination. Derived from the latter three temporal patterns, we establish a more realistic information diffusion and mitigation models.

Future work should include a self-supervised detection [39] of the different tempo-

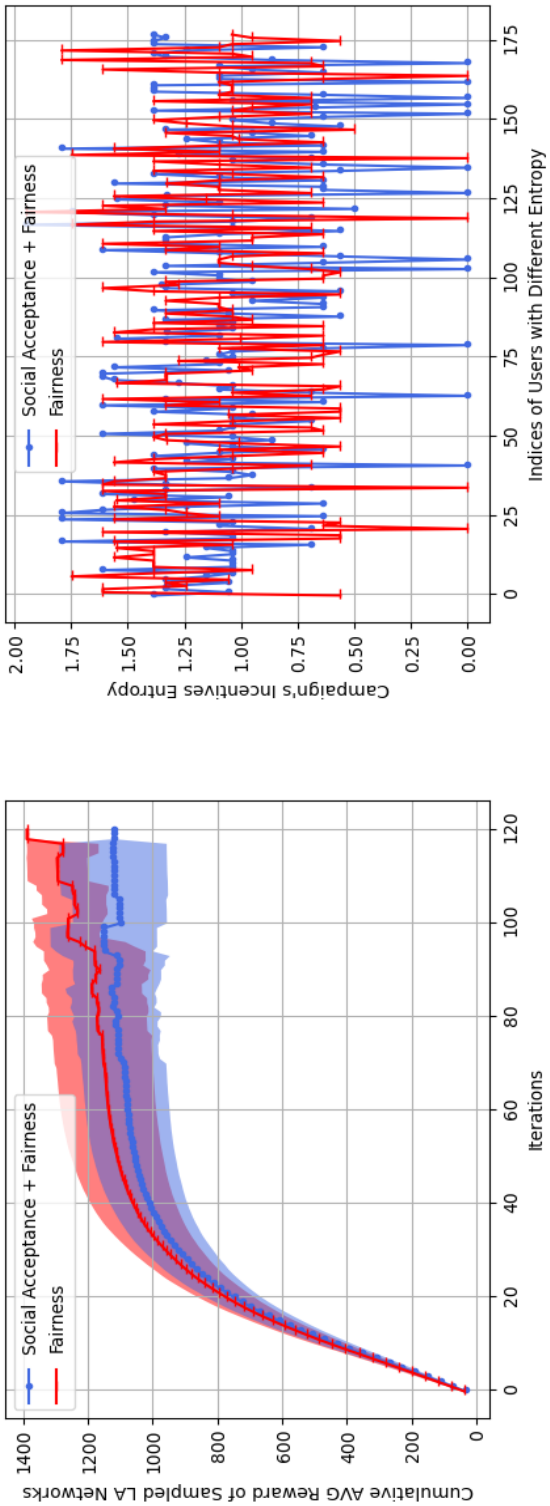


Figure G.9: Average cumulative rewards during incentive learning with entropy of the finally decided incentive value

ral events instead of the manual annotation. Moreover, more verification techniques should be studied to ensure realistic obtained incentives that would help in the real world. Eventually, the information Twitter API could provide to researchers is considered a limitation since the timestamps of likes are not provided.

Bibliography

[1] Femi Olan, Uchitha Jayawickrama, Emmanuel Ogiemwonyi Arakpogun, Jana Suklan, and Shaofeng Liu. Fake news on social media: the impact on society. *Information Systems Frontiers*, pages 1–16, 2022.

[2] Mehrdad Farajtabar, Nan Du, Manuel Gomez Rodriguez, Isabel Valera, Hongyuan Zha, and Le Song. Shaping social activity by incentivizing users. *Advances in neural information processing systems*, 27, 2014.

[3] Mehrdad Farajtabar, Xiaojing Ye, Sahar Harati, Le Song, and Hongyuan Zha. Multistage campaigning in social networks. *Advances in Neural Information Processing Systems*, 29, 2016.

[4] Kai Shu, Suhang Wang, Dongwon Lee, and Huan Liu. Mining disinformation and fake news: Concepts, methods, and recent advancements. In *Disinformation, misinformation, and fake news in social media*, pages 1–19. Springer, 2020.

[5] Ahmed Abouzeid, Ole-Christoffer Granmo, Christian Webersik, and Morten Goodwin. Learning automata-based misinformation mitigation via hawkes processes. *Information Systems Frontiers*, 23(5):1169–1188, 2021.

[6] Mehrdad Farajtabar, Jiachen Yang, Xiaojing Ye, Huan Xu, Rakshit Trivedi, Elias Khalil, Shuang Li, Le Song, and Hongyuan Zha. Fake news mitigation via point process based intervention. In *International conference on machine learning*, pages 1097–1106. PMLR, 2017.

[7] Xiaofei Xu, Ke Deng, and Xiuzhen Zhang. Identifying cost-effective debunkers for multi-stage fake news mitigation campaigns. In *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, pages 1206–1214, 2022.

[8] Ahmed Abouzeid, Ole-Christoffer Granmo, Christian Webersik, and Morten Goodwin. Socially fair mitigation of misinformation on social networks via constraint stochastic optimization. *arXiv preprint arXiv:2203.12537*, 2022.

[9] Mei Li, Xiang Wang, Kai Gao, and Shanshan Zhang. A survey on information diffusion in online social networks: Models and methods. *Information*, 8(4):118, 2017.

[10] Ole-Christoffer Granmo and B John Oommen. Optimal sampling for estimation with constrained resources using a learning automaton-based solution for the nonlinear fractional knapsack problem. *Applied Intelligence*, 33(1):3–20, 2010.

[11] Mahak Goindani and Jennifer Neville. Social reinforcement learning to combat fake news spread. In *Uncertainty in Artificial Intelligence*, pages 1006–1016. PMLR, 2020.

[12] Tal Schuster, Roei Schuster, Darsh J Shah, and Regina Barzilay. Are we safe yet? the limitations of distributional features for fake news detection. *arXiv preprint arXiv:1908.09805*, 2019.

[13] Xiaomo Liu, Armineh Nourbakhsh, Quanzhi Li, Rui Fang, and Sameena Shah. Real-time rumor debunking on twitter. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, pages 1867–1870, 2015.

[14] Jing Ma, Wei Gao, Prasenjit Mitra, Sejeong Kwon, Bernard J Jansen, Kam-Fai Wong, and Meeyoung Cha. Detecting rumors from microblogs with recurrent neural networks. 2016.

[15] William Yang Wang. " liar, liar pants on fire": A new benchmark dataset for fake news detection. *arXiv preprint arXiv:1705.00648*, 2017.

[16] Fatima K Abu Salem, Roaa Al Feel, Shady Elbassuoni, Mohamad Jaber, and May Farah. Fa-kes: A fake news dataset around the syrian war. In *Proceedings of the international AAAI conference on web and social media*, volume 13, pages 573–582, 2019.

[17] Kai Shu, Deepak Mahudeswaran, Suhang Wang, Dongwon Lee, and Huan Liu. Fakenewsnet: A data repository with news content, social context and spatial-temporal information for studying fake news on social media. *arXiv preprint arXiv:1809.01286*, 2018.

[18] Sonal Garg and Dilip Kumar Sharma. New politifact: a dataset for counterfeit news. In *2020 9th International Conference System Modeling and Advancement in Research Trends (SMART)*, pages 17–22. IEEE, 2020.

[19] Gregory Eady, Jonathan Nagler, Andy Guess, Jan Zilinsky, and Joshua A Tucker. How many people live in political bubbles on social media? evidence from linked survey and twitter data. *Sage Open*, 9(1):2158244019832705, 2019.

[20] Peiyuan Suny, Jianxin Li, Yongyi Mao, Richong Zhang, and Lihong Wang. Inferring multiplex diffusion network via multivariate marked hawkes process. *arXiv preprint arXiv:1809.07688*, 2018.

[21] Axel Bruns. Filter bubble. *Internet Policy Review*, 8(4), 2019.

[22] Brent Kitchens, Steven L Johnson, and Peter Gray. Understanding echo chambers and filter bubbles: The impact of social media on diversification and partisan shifts in news consumption. *MIS Quarterly*, 44(4), 2020.

[23] Jon Roozenbeek, Sander Van Der Linden, Beth Goldberg, Steve Rathje, and Stephan Lewandowsky. Psychological inoculation improves resilience against misinformation on social media. *Science advances*, 8(34):eabo6254, 2022.

[24] Patricia Moravec, Randall Minas, and Alan R Dennis. Fake news on social media: People believe what they want to believe when it makes no sense at all. *Kelley School of Business research paper*, (18-87), 2018.

[25] Mustafa A Al-Asadi and Sakir Tasdemir. Using artificial intelligence against the phenomenon of fake news: a systematic literature review. *Combating Fake News with Computational Intelligence Techniques*, pages 39–54, 2022.

[26] Sakshini Hangloo and Bhavna Arora. Content-based fake news detection using deep learning techniques: Analysis, challenges and possible solutions. In *2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT)*, pages 411–417. IEEE, 2022.

[27] Dhruv Khattar, Jaipal Singh Goud, Manish Gupta, and Vasudeva Varma. Mvae: Multimodal variational autoencoder for fake news detection. In *The world wide web conference*, pages 2915–2921, 2019.

[28] Eslam Amer, Kyung-Sup Kwak, and Shaker El-Sappagh. Context-based fake news detection model relying on deep learning models. *Electronics*, 11(8):1255, 2022.

[29] Kai Shu, Suhang Wang, and Huan Liu. Beyond news contents: The role of social context for fake news detection. In *Proceedings of the twelfth ACM international conference on web search and data mining*, pages 312–320, 2019.

[30] Alvin I Goldman and Daniel Baker. Free speech, fake news, and democracy. *First Amend. L. Rev.*, 18:66, 2019.

[31] Shoujin Wang, Xiaofei Xu, Xiuzhen Zhang, Yan Wang, and Wenzhuo Song. Veracity-aware and event-driven personalized news recommendation for fake news mitigation. In *Proceedings of the ACM Web Conference 2022*, pages 3673–3684, 2022.

[32] Ryota Kobayashi and Renaud Lambiotte. Tideh: Time-dependent hawkes process for predicting retweet dynamics. In *Tenth International AAAI Conference on Web and Social Media*, 2016.

[33] Eirini Ntoutsi, Pavlos Fafalios, Ujwal Gadiraju, Vasileios Iosifidis, Wolfgang Nejdl, Maria-Esther Vidal, Salvatore Ruggieri, Franco Turini, Symeon Papadopoulos, Emmanouil Krasanakis, et al. Bias in data-driven artificial intelligence sys-

tems—an introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3):e1356, 2020.

[34] Terry L Amburgey. Multivariate point process models in social research. *Social Science Research*, 15(2):190–207, 1986.

[35] Marian-Andrei Rizoiu, Young Lee, Swapnil Mishra, and Lexing Xie. A tutorial on hawkes processes for events in social media. *arXiv preprint arXiv:1708.06401*, 2017.

[36] Anis Yazidi, Nourredine Bouhmala, and Morten Goodwin. A team of pursuit learning automata for solving deterministic optimization problems. *Applied Intelligence*, 50(9):2916–2931, 2020.

[37] Samik Raychaudhuri. Introduction to monte carlo simulation. In *2008 Winter simulation conference*, pages 91–100. IEEE, 2008.

[38] Steven Bankes, Robert Lempert, and Steven Popper. Making computational social science effective: Epistemology, methodology, and technology. *Social Science Computer Review*, 20(4):377–388, 2002.

[39] Ahmed Abouzeid, Ole-Christoffer Granmo, Morten Goodwin, and Christian Webersik. Label-critic tsetlin machine: A novel self-supervised learning scheme for interpretable clustering. In *2022 International Symposium on the Tsetlin Machine (ISTM)*, pages 41–48. IEEE, 2022.

Appendix H

Paper G Supplementary

H.1 Control Model

The individual Learning Automaton (LA) system associated with each user and the whole network system can be viewed as Markov systems through the state transitions of the LAs and the joint probabilities of the latter. Figure H.1 illustrates both the individual LA_i state transitions probabilities matrix S_i and the whole LAs joint probabilities matrix P. Where state transitions and joint probabilities change between an intervention step e until convergence to a steady state and joint probability of being in such states by the intervention step e^*. As indicated in Figure H.1, each LA_i can only perform a state transition by moving one step either to the left, right, or staying at the current state. Eventually, the state transition convergence for each LA_i means it converged to a transition probability of staying at its current state with a value close to 1.

Further, the joint probability of being in a particular state for all LAs determines the final incentive values on a sampled network. These state transitions are governed by a reward and penalty signal β as shown in Figure H.2. Such a signal comes from evaluating the gradient of the total loss function. For instance, if the total loss declined compared to its previous value, the LA which caused that will be rewarded, and its state transition will be committed. Otherwise, it will be penalized and should stay in its current state. We adopted the same probability calculations and reward function of the utilized LAs as proposed in [8].

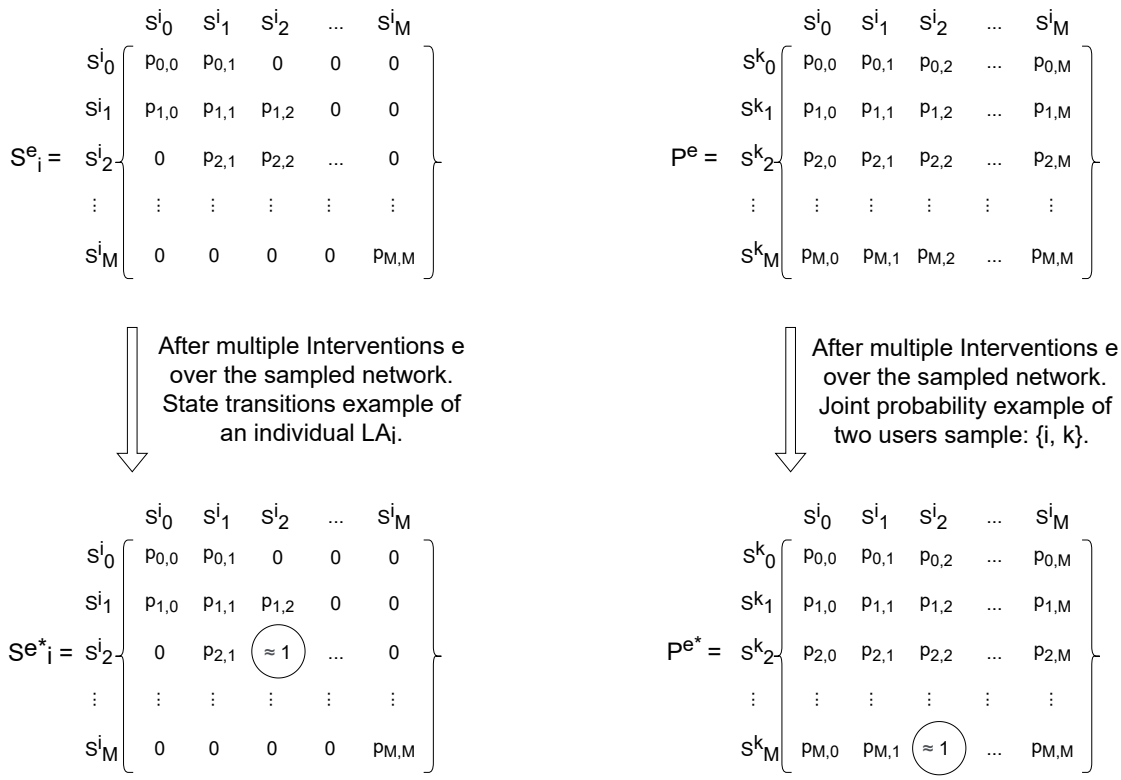


Figure H.1: The individual LA_i state transitions probabilities matrix S_i and the whole LAs joint probabilities matrix P of their joint state transitions from intervention step e until convergence in intervention step e^*

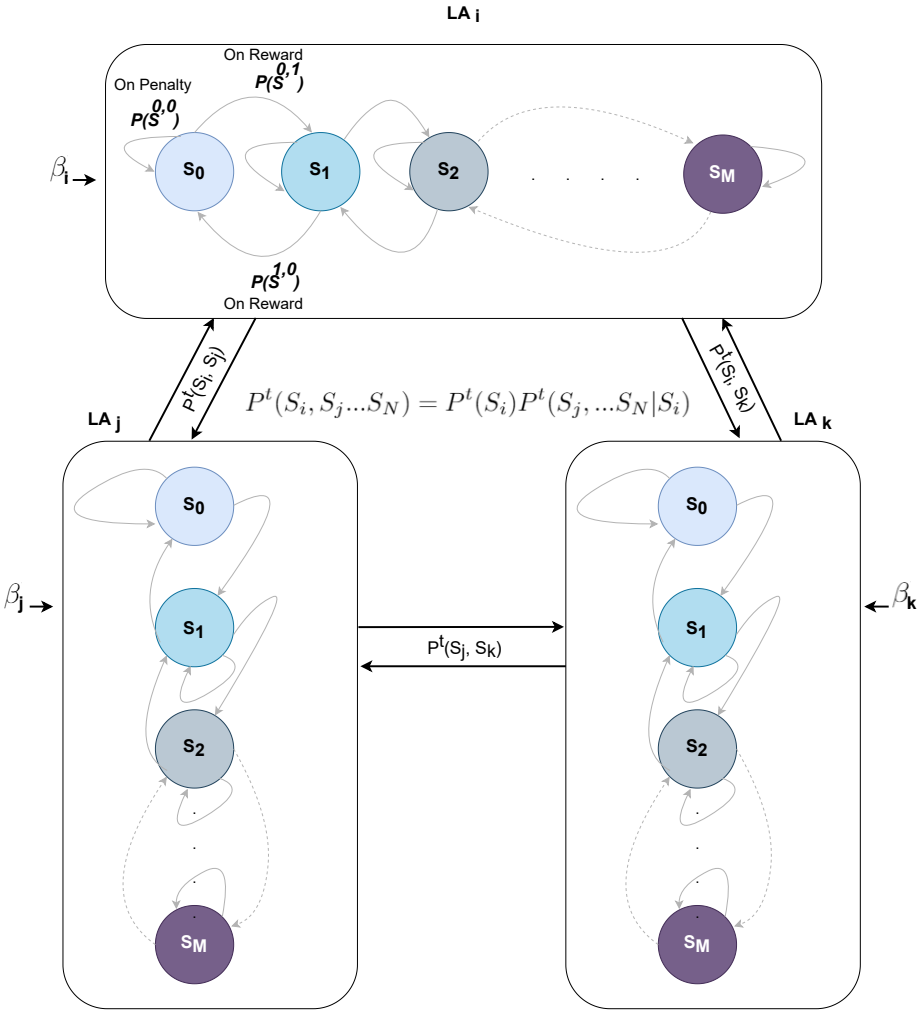


Figure H.2: A toy example of sampled network of three users and their associated three automata

Appendix I

Paper G Supplementary

I.1 MHP Simulations Setup

I.1.1 Bias-towards

In such Multivariate Hawkes Process (MHP), we simulated the temporal bias activity of all the 940 users. That means we trained an MHP with timestamps and event counts aggregated over the time realizations period of the samples labeled as *bias=1*. Hence, for each user and time realization (e.g., 180 minutes), all timestamps within the time realization window were structured accordingly. For instance, the 1^{st} time realization had only the timestamps for events where their Twitter creation times were within the first three hours of the 24^{th} of March 2018. Accordingly, the 2^{nd} time realization contained the bias-towards samples timestamps that occurred between 3:00 AM to 5:59 AM on the same day. Then, we kept shifting the time realizations and their associated timestamps the same way until day 27^{th} 8:59 PM, where the following three hours of the day were not part of the MHP training since they were left for testing the exact three hours predictions.

We set the decay factor for this MHP to 0.6. The primary purpose of such MHP was to predict all users' bias-towards activity to calculate the initial probability of being biased toward the election. The probabilities for each user were calculated according to the frequency of having an associated event belonging to the label *bias=1* in training and predicted data. This process was not incentivized and was only created to calculate such probabilities for the optimization loss function domain (see Equation G.4, Equation G.9, and Equation G.10).

I.1.2 Bias-against

The same concept of the Bias-towards-MHP training also applies to this process. Therefore, we established it to predict all the 940 users' activity for the bias-against event category to calculate the initial probabilities of users being biased against the election (see Equation G.4, Equation G.9, and G.10).

We set the decay factor to 0.7. It is essential to highlight that this process was not incentivized. Alternatively, we established the same MHP event category as

discussed below but only for a sampled network (100 users), where that process was incentivized.

I.1.3 Bias-against-sampled

This process had the same training timestamps concept as the above two MHPs. The difference between a sampled bias-against-MHP and the 940 bias-against-MHP is that the latter was used to estimate the initial probabilities of being against the election. At the same time, the former was essential to predict these probabilities after intervention and assigning the incentives. That means some of these probabilities would change, indicating how good the incentives were for some users for mitigating the bias towards the manipulating election campaign and optimizing the loss function (see Equation G.4, Equation G.9, and Equation G.10). Therefore, this process was incentivized with the amounts of state values from the converged LAs. It had only 100 users since it was part of the simulated MHPs during the Monte Carlo simulation. Hence, we intervened with the sampled network and evaluated each user inside it for the associated LA state value because our proposed non-propaganda incentivization also presented a bias-against concept.

We set the decay factor to 0.9 in this process. Since this is a sampled network of users, this process was repeated with different samples for both training and prediction.

I.1.4 Propaganda-sampled

In this process, we have followed the same structure for training the timestamps but only for the sampled 100 users. The process was repeated with a different sample for training and prediction each time. This MHP was not incentivized since we did not wish to intensify the propagation of political propaganda and was only simulated to obtain the predicted counts for users. The obtained counts were used in the ratio parameter for the optimization loss function (see Equation G.4 and Equation G.6). We set the decay factor for this process to 0.9.

I.1.5 Non-propaganda-sampled

Similar to the propaganda-sampled MHP, we established a repeated MHP for the non-propaganda event category where a random sample represented the timestamps for training and predicting the activity of the sampled users (100 users). This process was incentivized with the incentive amounts from the current evaluated user's associated LA state value.

This process had a decay factor of 0.6. The predictions in this process were the direct outcome of the intervention procedure and assigning of an incentive for the current examined user during the Monte Carlo simulation. Therefore, the event counts were evaluated as part of the ratio parameter in the optimization loss function (see Equation G.4 and Equation G.6).

I.1.5.1 Societal Circle A

To predict all users' activity on the network on how they engaged with the societal circle A, we established this MHP on all the 940 users to predict their future generated events for that circle. The same training timestamps structure was adopted for the MHP with a decay factor of 0.9. Since this is not a sampled network MHP, we ran it only once to be able to predict the initial probability of engagement with circle A (see Equation G.4 and Equation G.9) for each user without any incentivization.

I.1.6 Societal Circle B

The exact purpose of the societal circle A-MHP was adopted in this process since all societal circles on all the 940 users must be predicted to calculate the initial probabilities of being in a specific societal circle. Since this process was used to calculate the initial probabilities, we did not apply any incentivization. We assigned the decay factor for this process with 0.9.

I.1.7 Societal Circle B-sampled

The only main difference between the sampled societal circle-B-MHP and the non-sampled circle-B-MHP is the repetition and incentivization in the former. This MHP was incentivized since we wanted to break other circles by intensifying the engagement with it in addition to agreeing. If the latter had occurred, we would increase the probability of engaging with circle B and agreeing with what it represents, which optimizes our loss function (see Equation G.4 and Equation G.9). This process followed the exact configurations in the non-sampled societal circle-B-MHP.

I.1.8 Societal Circles C, E, F

These three MHPs are identical to the non-sampled circle-A and non-sampled circle-B MHPs. Since we had to consider, all circles predicted counts to calculate the initial probabilities of engaging with a circle. The only difference was the decay factors used as we assigned them with 0.75, 0.9, and 0.9, respectively.

Eventually, we ignored simulating circle D since it had only two users in the original PEGYPT dataset, which was not enough to train a MHP. However, our results were reliable since they had only two users.

Appendix J

Paper G Supplementary

J.1 Simulation Results

Figure J.1 gives an example of some simulations with 100 users' real versus predicted event counts.

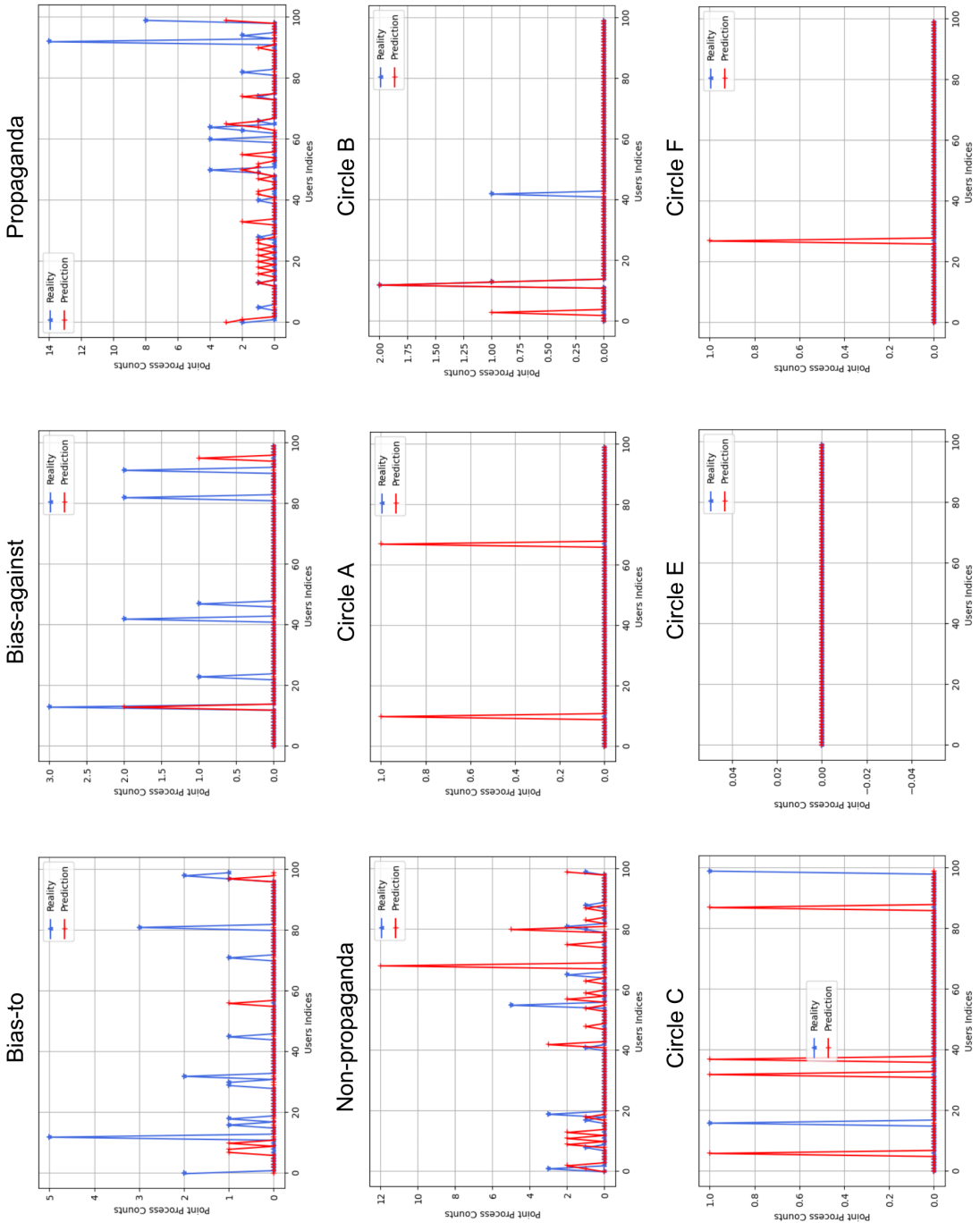


Figure J.1: An example of some simulations with 100 users' real versus predicted event counts