Simulation of Virus Propagation and Acceptance of Socio-Sanitary Measures Through an Intelligent Model

Daniel J. García-Paz, Amparo Alonso-Betanzos, Bertha Guijarro-Berdiñas, and Alejandro Rodríguez-Arias

Universidade da Coruña, LIDIA, Faculty of Computer Science, 15071 A Coruña, Spain
Universidade da Coruña, Centro de Investigación CITIC, 15071 A Coruña, Spain
Correspondence: alejandro.rodriguez.arias@udc.es

DOI: *https://doi.org/10.17979/spudc.000024.43*

Abstract: During the most critical moments of the SARS-COV-2 pandemic, various containment measures were enacted to hinder the virus's spread and mitigate its impact. This work focuses on studying the impact of the population's adherence level to socio-sanitary measures on the virus's spread, aiming to better understand its relevance in crisis situations. To achieve this goal, we use an agent-based model (ABM) that incorporates a special type of agent that represents social networks, for example, twitter, to analyze the influence of social networks on the agents' decision-making.

Internally, our model relies on two models that allow for simulation development. On the one hand, an epidemiological model based on an adaptation of the SIR model allows us to simulate the spread of the virus. On the other hand, a decision-making model is responsible for analyzing the levels of acceptance of containment measures by citizens and allows simulation of interactions between agents. On this basis, Twitter has been incorporated as a critical node, which allows information to be extracted about the opinions of the agents and how these affect the population's adherence to socio-sanitary measures. This information is obtained thanks to the application of sentiment analysis techniques on a set of tweets related to COVID-19.
As a result, a useful tool was obtained for policy makers to simulate the psycho-social behaviour of citizens in the face of different restrictive measures in order to evaluate their effectiveness.

1 Introduction

The SARS-CoV-2 crisis is still seen as a period of uncertainty and significant social change, even three years later. It was strongly marked by restrictions and the imposition of socio-sanitary measures on the citizens in many countries, aiming to mitigate the spread of infections and consequently, the advancement of the virus causing the pandemic.

Over the past three years, various studies such as Saez et al. (2020) or Aleta and Moreno (2020) have indeed demonstrated that all these measures positively contributed to the fight against COVID-19. However, despite their widespread implementation, it is known that these measures have been extremely controversial due to social movements that denied their effectiveness or even the existence of the virus itself, and consequently, refused to adhere to them. This raises the possibility of questioning the impact that individual actions can have on the spread of a virus or, in other words, whether the fact that an individual or a minority refuses to adhere to the measures has a real effect on their effectiveness for the entire population.

In this context, social networks becomes particularly relevant as a source of numerous discussions in which both critics and supporters of socio-sanitary measures have used them to express their opinions and, at times, to try to persuade others of their stance. With this in mind, it's worth asking: Beyond the merit of socio-sanitary measures, what is the impact of these discussions on social networks on the population? Do they affect the effectiveness of the measures? Do they instill distrust in the measures among the public?

The search for answers to these questions indeed places us in a complex situation because the study of human behavior, the factors influencing it, and how all of this affects the evolution of an epidemic is not a trivial matter. This is where the use of computational models becomes particularly relevant, allowing us to study the behavior of a population in a controlled virtual environment. Agent-based models represent systems by simulating the actions of their individual components (agents, which act autonomously), and the interaction between them and their environment, allowing complex processes and systems to be solved. A key feature of ABM is that it allows us to tackle or study problems related to the concept of emergent phenomena. This means examining system dynamics that appear from the behavior and interaction of the individuals that make up the system, thus allowing the study of questions related to how the global behavior of society links to individual behavior. As a result, ABMs have been widely used in various fields, including epidemiology, Amouroux et al. (2010), and the social sciences, Groff et al. (2019).

This work proposes the application of an agent-based model to study how the effectiveness of socio-sanitary measures is influenced by the population's behaviour, which will be shaped by the information they receive through social networks. To achieve this, the city of A Coruña (Spain) will be used as the geographical virtual environment, based on Geographic Information System (GIS) data, and real socio-demographic data from the city will be used to populate the agents. These agents will exhibit certain behaviors that allow them to satisfy their individual needs, and these behaviors will be influenced by the information they receive through social networks. To accomplish this, a decision-making model (see section 2.3) will be implemented to simulate social behaviors, the acceptance of socio-sanitary measures, and communication among the agents. For modeling the virus's spread, an epidemiological model based on the SEIRD model will be implemented (see section 2.2).

With this foundation established, our aim in this work is to simulate the spread of information to agents through social networks, represented as special agents in the model. These agents will be fed by a set of tweets from the 6-month period to be simulated, between June and November 2021, which will be processed using sentiment analysis techniques to gain an understanding of the opinions that socio-sanitary measures elicit in the population and how this opinion impacts the behaviour of the agents and their acquaintances.

2 Proposed model

As mentioned before, the model relies on 4 pillars: the environments, the epidemiological model, the decision-making model, and the critical nodes representing social networks.

2.1 Enviroment

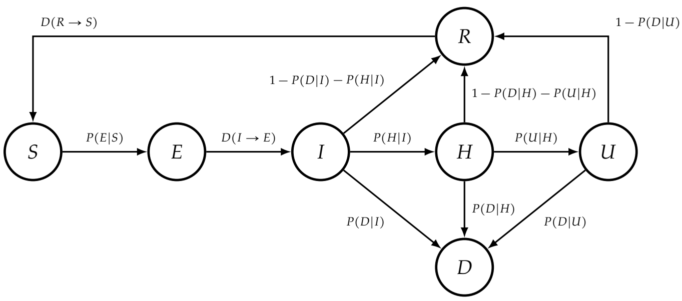
The environment (see figure 1) has been developed based on GIS data from the city of A Coruña, which has been used to implement both the city's structure based on its census zones and the location of leisure establishments, supermarkets, government buildings, or offices, among others.

Figure 1: Representation of the city of A Coruña in the proposed model.

Influence network

Another fundamental part of the environment in which agents interact is the existence of networks of agents with whom they can interact. For the creation of influence networks, two different subgroups of agent contacts have been established. These groups will consist of two networks, on one hand, a network of friends composed of a random set of agents in the environment who are within a similar age range (maximum of 5 years difference) as the agent, furthermore, there is the possibility of adding random friendship relationships between agents. On the other hand, the citizen will have a social circle, Hamill and Gilbert (2009), a set of individuals who will be part of their influence network due to geographic proximity.

2.2 Epidemiological model



where:

$P(Y|X)$: Probability that an individual in state X transitions to state Y.

$D(X \rightarrow Y)$: Days that an individual in state X takes to move to state Y.

Figure 2: SEIRDS adaptation

As an epidemiological model, an adaptation of the SEIRDS model has been chosen (see Figure 2), an evolution of the SIR model proposed by Kermack and McKendrick (1927). The SEIRDS model addresses the limitations of the SIR model, which only considers the existence of three possible states (susceptible, infected, and recovered), by adding a latency period, the possibility of reinfection, and the lethality of the disease.

Regarding this model, our adaptation takes into account two additional assumptions: the hospitalization of infected individuals and the admission to the ICU of hospitalized individuals, as the probability of death increases in each of these cases, reducing the probability of recovery.

2.3 HUMAT

Decision-making and agent interaction are crucial aspects in the development of this system. To achieve this, we have chosen to implement an HUMAT architecture, Antosz et al. (2019), which allows us not only to implement a decision-making model based on the individual needs of the agents but also to create networks of relationships among agents that enable the spread of information and the influence of some agents over others.

The HUMAT architecture addresses the problem of decision-making in a five-step cycle (see Figure 3) after initialization:

- **Evaluate:** The cycle starts with the evaluation of the two alternatives presented to the agent to satisfy their needs: 1) accepting the measures or 2) rejecting the measures.
- **Choice:** The agent selects its behavior based on its evaluation.
- **Action:** The agent acts accordingly, which will impact the probabilities of infection and communications with other agents in its circle of influence.
- **Experimentation:** The effects resulting from the agent's decision-making emerge.
- **Update:** The agent's needs are updated based on its experiences.

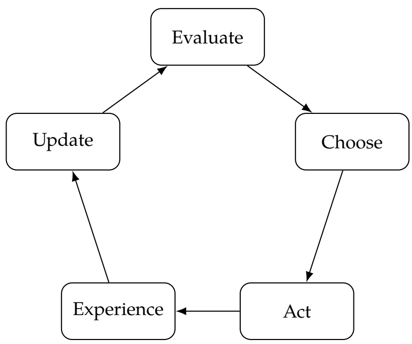


Figure 3: Decision-making strategy

2.4 Critical nodes

In our model, social network information is propagated through critical nodes, representing specific agents in the model capable of significantly altering agent behavior by establishing communication with them. In our case, these critical nodes represent social networks, specifically Twitter, which will influence agent behavior through the dissemination of tweets that have been previously processed using natural language processing techniques (see section 3). For the propagation of information, the critical node sends between 4 and 6 tweets daily to each of the users of the social network. These communications will involve 2 types of tweets (in favor or against the measures), which will modify the satisfaction of the individual needs of the agents.

3 Natural language processing

Sentiment analysis through Twitter data has been a topic of study since the beginning of the pandemic, being part of various studies such as Lopez and Gallemore (2021), Kaur et al. (2020), or Dubey (2020). In this line, the main objective of this study is to analyze whether this approach can be used to study the impact that information circulating on social networks has on the public opinion and, consequently, on the acceptance of socio-sanitary measures. For this purpose, it is necessary to obtain information from social networks, in our case, Twitter, and then process it to extract relevant information.

To obtain this information, the Twitter API has been used, which allows the collection of tweets through various filters, providing information related to a specific topic, user, or geographic location. In our case, we have chosen to collect tweets about Covid-19 within the Spanish-speaking community, as it is generally assumed that this community will post tweets that will reach our population. The result has been a set of 4,164 tweets spread over 24 months, from March 1, 2020, to February 28, 2022, of which we will use a subset corresponding to the 6-month period from June to November 2021, which coincides with a surge in COVID-19 infections in the city, according to data from the Galician Health Service (SERGAS).

The processing of tweets is carried out in two stages. First, a preprocessing step is performed because tweets, by their nature, often contain elements that do not provide information or cannot be processed by the language model. In this step, hashtags and URLs have been removed from the tweets, and emojis have been replaced with explanatory texts for each one.

On this basis, for the tweet processing, ”pysentimiento”, Pérez et al. (2021), has been used, a natural language processing framework designed for Twitter and based on RoBERTa, Liu et al. (2019), a pre-trained language model based on transformers designed for natural language processing tasks like sentiment analysis. This model allows us to obtain a series of values from a tweet that indicate its polarity, i.e., whether the tweet is positive, negative, or neutral (see figure 4). These polarity values are used for the propagation of information through the critical node, altering the satisfaction of individual agent needs, and consequently, their behavior. This allows us to study the relevance of these types of messages, both for Twitter users themselves and for the rest of the population.

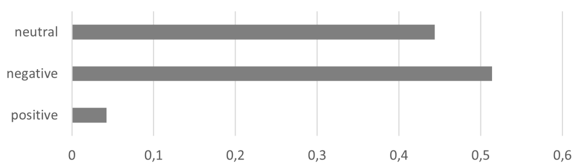


Figure 4: Percentages of tweets polarization.

4 Simulations

The model executions have been carried out with a total population of 215,000 agents, corresponding to the adult population of the city of A Coruña, of which 300 are infected at the beginning of the simulation And it will take place between June and November 2021. These agents are distributed across 187 census zones and move through approximately 4,000 different locations, including leisure establishments, supermarkets, government buildings, offices, among others.

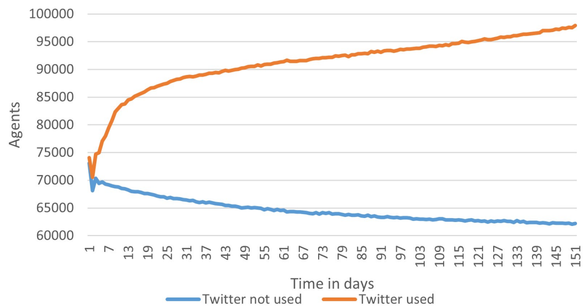
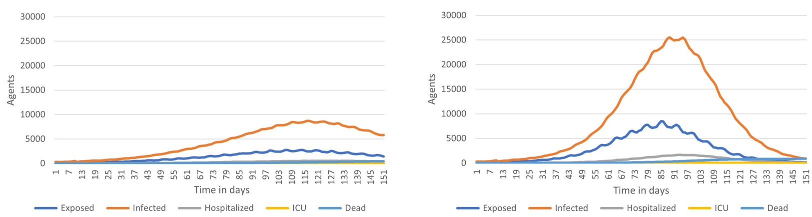


Figure 5: Reject behaviour evolution comparation

With the aim of observing the impact of information circulating on social networks on the public opinion, two types of model executions have been conducted, one in which critical nodes were not included and another in which they were included (see Figure 5).

Regarding the influence of critical nodes, we can observe that indeed, the information circulating on social networks has a significant impact on the public opinion. This translates into an increase in the number of agents rejecting socio-sanitary measures and, consequently, an increase in the number of infected agents (See figure 6). This is because, as can be seen in Figure 4, the information obtained on Twitter has a significant bias towards negativity, with positive polarization being around 5%. This leads to an increase in the rejection of socio-sanitary measures.



(a) SEIRD evolution without Twitter

(b) SEIRD evolution with Twitter

Figure 6: SEIRD evolution comparation

5 Conclusions

Based on the results of the simulations, it can be concluded that the proposed model is capable of analyzing the variation in the levels of acceptance of socio-sanitary measures by the population and how this variation affects the spread of the virus. It is also evident that the addition of critical nodes to the model correctly alters the behavior of the agents, allowing us to conclude that the information circulating in the model is having the intended impact.

Furthermore, it can be observed that the information distributed by the critical nodes has had a significant impact on the decision-making of the agents. Indeed, there is a direct correlation between the number of agents rejecting socio-sanitary measures and the number of infected agents. Therefore, we can conclude that both the decision-making model, the epidemiological model, and the critical nodes exhibit the expected behavior.

However, it has become evident that the sentiment analysis technique used is not capable of correctly classifying tweets. This is because the negative polarization does not directly imply a rejection of socio-sanitary measures. It has been observed that the technique used cannot differentiate between tweets expressing a negative opinion about the measures and tweets expressing a negative opinion about other questions. Therefore, it would be interesting to consider alternative techniques that allow us to address this problem and achieve accurate classification between these two types of tweets.

Acknowledgments

CITIC is funded by the Xunta de Galicia through the collaboration agreement between the Consellería de Cultura, Educación, Formación Profesional e Universidades and the Galician universities for the reinforcement of the research centres of the Galician University System (CIGUS).

Bibliography

A. Aleta and Y. Moreno. Evaluation of the potential incidence of covid-19 and effectiveness of containment measures in spain: a data-driven approach. *BMC medicine*, 18(1):1–12, 2020.

E. Amouroux, P. Taillandier, and A. Drogoul. Complex environment representation in epidemiology abm: application on h5n1 propagation. *Journal of science and technology*, pages 13–25, 2010.

P. Antosz, W. Jager, G. Polhill, D. Salt, A. Alonso-Betanzos, N. Sánchez-Maroño, B. Guijarro-Berdiñas, and A. Rodríguez. Simulation model implementing different relevant layers of social innovation, human choice behaviour and habitual structures. *SMARTEES Deliverable*, 2019.

A. D. Dubey. Twitter sentiment analysis during covid-19 outbreak. *Available at SSRN 3572023*, 2020.

E. R. Groff, S. D. Johnson, and A. Thornton. State of the art in agent-based modeling of urban crime: An overview. *Journal of Quantitative Criminology*, 35:155–193, 2019.

L. Hamill and G. Gilbert. Social circles: A simple structure for agent-based social network models. *Journal of Artificial Societies and Social Simulation*, 12(2), 2009.

S. Kaur, P. Kaul, and P. M. Zadeh. Monitoring the dynamics of emotions during covid-19 using twitter data. *Procedia Computer Science*, 177:423–430, 2020.

W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.

Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*, 2019.

C. E. Lopez and C. Gallemore. An augmented multilingual twitter dataset for studying the covid-19 infodemic. *Social Network Analysis and Mining*, 11(1):102, 2021.

J. M. Pérez, J. C. Giudici, and F. Luque. pysentimiento: A python toolkit for sentiment analysis and socialnlp tasks. *arXiv preprint arXiv:2106.09462*, 2021.

M. Saez, A. Tobias, D. Varga, and M. A. Barceló. Effectiveness of the measures to flatten the epidemic curve of covid-19. the case of spain. *Science of The Total Environment*, 727:138761, 2020. URL *https://www.sciencedirect.com/science/article/pii/S0048969720322786*.