UNIVERSIDAD POLITÉCNICA DE MADRID

ESCUELA TÉCNICA SUPERIOR
DE INGENIEROS DE TELECOMUNICACIÓN

GRADO EN INGENIERÍA DE TECNOLOGÍAS Y SERVICIOS DE TELECOMUNICACIÓN

TRABAJO FIN DE GRADO

DEVELOPMENT OF AN AGENT-BASED MODEL FOR SIMULATING THE INFLUENCE OF JIHADIST TERRORIST ATTACKS ON THE 2004 SPANISH GENERAL ELECTION

GABRIEL JOSÉ FUNES GABALDÓN

JUNIO 2022

TRABAJO DE FIN DE GRADO

Título:	Desarrollo de un modelo basado en agentes para la simulación de la influencia de los ataques terroristas en las elecciones generales españolas en 2004
Título (inglés):	Development of an agent-based model for simulating the influence of jihadist terrorist attacks on the 2004 Spanish general election
Autor:	Gabriel José Funes Gabaldón
Tutor:	Carlos Ángel Iglesias Fernández
Departamento:	Departamento de Ingeniería de Sistemas Telemáticos

MIEMBROS DEL TRIBUNAL CALIFICADOR

Presidente:	—
Vocal:	—
Secretario:	—
Suplente:	—

FECHA DE LECTURA:

CALIFICACIÓN:

UNIVERSIDAD POLITÉCNICA DE MADRID

ESCUELA TÉCNICA SUPERIOR DE
INGENIEROS DE TELECOMUNICACIÓN

Departamento de Ingeniería de Sistemas Telemáticos
Grupo de Sistemas Inteligentes

TRABAJO FIN DE GRADO

DEVELOPMENT OF AN AGENT-BASED MODEL FOR SIMULATING THE INFLUENCE OF JIHADIST TERRORIST ATTACKS ON THE 2004 SPANISH GENERAL ELECTION

Gabriel José Funes Gabaldón

Junio 2022

Resumen

Hace 18 años, uno de los peores ataques terroristas que ha sufrido España saturó de información a los principales medios de comunicación. Muchos mensajes que estos medios transmitían se veían influenciados por diferentes referentes políticos, incluyendo el mismo gobierno, y, por lo general, debatían sobre la responsabilidad de los ataques señalando a dos posibles autores: ETA y Al Qaeda.

Esta controversia causó un giro inesperado en las elecciones generales que se celebraron tres días después de las explosiones, dando al PSOE la victoria en el Congreso cuando días antes las encuestas más fiables situaban al PP como el favorito a la canditatura.

Este proyecto se centra en el diseño e implementación de una simulación social basada en agentes mediante el desarrollo de un modelo artificial para todos los votantes españoles que participaron en las elecciones generales de 2004.

Este modelo desarrollado en Soil cuenta con la integración de dos entidades diferentes: un entorno exterior de difusión y un entorno interior de interacción. El entorno de difusión actúa como emisor de mensajes hacia la población, generando un efecto de influencia en los agentes al igual que lo harían los medios de comunicación. El entorno interior se forma mediante la generación de una red artificial de agentes conectados que interactúan entre sí para intercambiar variables y parámetros.

El objetivo principal de este proyecto es analizar y extraer información de los ficheros obtenidos tras ejecutar múltiples simulaciones en lo que respecta a la influencia de los atentados del 11M y su tratamiento por los medios de comunicación en los resultados electorales del 14M. Esto se consigue gracias a la filtración de dichos datos y análisis de sensibilidad, comparación de los distintos modelos generados y a la visualización de las variables más relantes y su evolución media a través de las simulaciones.

Palabras clave: Simulación social basada en agentes, atentados del 11M, votantes, medios de comunicación, Soil, red artificial

Abstract

18 years ago, one of the worst terrorist attacks that Spain has suffered overloaded with information the most relevant mass media channels. Most of this transmitted information was influenced by different political figures, including the government; and, in general, the messages received by the Spanish population discussed the responsibility of the attacks, pointing at two different perpetrators: ETA and Al-Qaeda.

This controversy caused an unexpected turn on the general elections that took place three days after the explosions were produced, which resulted in a victory for PSOE even when a few days earlier, the most reliable surveys showed PP as the favorite of the candidacy.

This project focuses on the design and implementation of an Agent-Based Social Simulation (ABSS) based on the development of an artificial model to simulate all Spanish voters who participated in the Spanish general elections of 2004.

This model has been developed in the Soil and integrates two different entities: an external diffusion environment and an internal interaction network. The environment acts like the emitter of all messages to the population, generating an influence effect on the agents, as the media would do in voters. The internal environment is composed of the generation of an artificial network where neighbor agents interact between them, interchanging variables and parameters.

The main goal of this project is to analyze and extract information related to the influence that mass media exercised in the population on the results of the general elections by talking about the terrorist attacks, all of which is extracted from the output files obtained after running the simulation several times. We achieve this by filtering the data and performing an analysis of sensibility, a comparison between the generated models, and visualization of the most important variables and their evolution through simulation.

Keywords: Agent-Based Social Simulation, 11M terrorist attacks, voters, mass media, Soil, artificial network

Agradecimientos

Voy a hacer buen provecho de estas líneas para agradecer profundamente a toda la gente que me ha ayudado a desarrollar, no sólo este proyecto, sino también mi vida personal y profesional durante estos cuatro años de carrera.

A mi familia, sobre todo a mis padres y a Laura, por el apoyo incondicional que me ha brindado incluso en los momentos más difíciles.

A todos los amigos, compañeros y profesores que me han acompañado de la mano durante mi recorrido por la ETSIT.

A todos los miembros del GSI, en particular, a Carlos y a Fernando por ofrecerme su ayuda siempre que la necesitaba y hacer un seguimiento continuo de todo mi trabajo y código. También a mis compañeros de laboratorio, con los que se ha hecho más amena la estancia de muchas mañanas programando.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

As humans, over the years, we have been developing many ways of communicating things to each other, different techniques that have been evolving due to the social implementation of new technologies. These techniques have changed from the first mass media methods that arrived to perpetuate knowledge towards future generations, to the infinite technological resources we can use today to continue building a society where everyone is aware of what is happening at all times. Because it is undeniable that in the 21^{st} century we are, the mass media has the powerful capacity to influence cognition and comprehension [Wei66]; in relationships and voting behavior. This kind of influence is such that we can find many articles that describe, for example, subconscious marketing communication techniques [BOL20] and how these techniques affect the way customers behave emotionally toward advertising on television and other types of mass media.

The title and content of this project were inspired by the article [MCSLC17]. We contacted some of the researchers who drafted the article to obtain all the data that they used about the messages and publications of mass media between the days this project aims to simulate and adapt those data to the technologies used in this project as well as some calibrated variables. In this chapter, we review one of the most terrible and controversial events of our century. An event that directly affected its victims in addition to the overwhelmingly Spanish population.

1.1 11M Terrorist attacks

Madrid, on March 11th 2004. It was early in the morning when it happened, between 7:36 and 7:39 AM. Ten explosions were simultaneously produced in four different areas of some Renfe Cencanías stations between Alcalá de Henares and Atocha. This caused the immediate death of 176 people, and the other 15 died in different hospitals after the terrorist attack. Collectively, these explosions caused a set of 191 deaths and more than 1800 injured citizens. This desolating incident was immediately communicated through the mass media to most of the Spanish population, causing great impact in Spanish March 14th general elections which were hold three days after the attacks. The political parties with the greatest impact, predominantly Partido Popular (PP) and Partido Socialista Obrero Español (PSOE), took advantage of the situation mentioned to tilt the scales in their own favor, manipulating the mass media.

This manipulation consisted mainly of the conviction of terrorist attacks against different terrorist groups and the generation of political propaganda and polarized messages to obscure the opposition. Overall, messages transmitted by the PP politicians maintained that the attacks were committed by the Spanish terrorist group ETA, while PSOE politicians argued that the current government (led by PP) was lying to the Spanish population, organized a big social protest, and blamed Al-Qaeda as responsible for the terrorist attack. Moreover, these messages gradually moved as the days passed from pointing to ETA to pointing to Al-Qaeda as a consequence of new investigation factors that little by little designated Al-Qaeda as the perpetrators of the terrorist attacks. Other media channels tried to hide those evidences to help whichever political segments. Despite this, the goals of this project do not include the debate of which of these two political parties was right and what their real intentions were.

1.2 Project goals

This project aims to find an accurate agent-based model that fits the activity on artificial social networks about this incident from 11th March to 14th March. In particular, our goal is to understand how political parties and media channels influenced Spanish society.

In an effort to achieve this objective, this project establishes four main goals to meet.

- Investigation of different variables for psychological influence and Agent-Based Social Simulation (ABSS).

- Simulation of controlled environments where the mass media and neighbours interact.
- Improvement of network interactions and isolation of processes for its study.
- Extraction of valuable data and conclusions from the obtained data.

1.3 Structure of this document

In this section, we provide a quick overview of all the chapters included in this document, which are structured as follows.

Chapter 1. Introduction introduces a general vision of the project, as well as its main goals and motivations.

Chapter 2. Enabling technologies provides a summary of the principal features and techniques that exist today to develop the proposed models.

Chapter 3. 11M model description. This chapter defines the evolution of simulation models designed to achieve the proposed objectives. Inside these models: agent behavior, environment parameters, and other factors.

Chapter 4. Results analyzes the results obtained from the models and explain their validation.

Chapter 5. Conclusions discusses the outcomes and conclusions and ponders future work.

Appendix A. Impact of this project reflects the possible impacts of this project.

Appendix B. Economic budget sets the details of a minimum budget to carry out the developed project.

CHAPTER 2

Enabling technologies

With a focus on fulfilling the goals of this project, it is necessary to study and look deeper into the enabling technologies that make it possible to simulate particular situations or social contexts. Today, multiple models can be found for the purpose of foretelling social simulations using agents.

In this chapter, we describe the most important parameters to take into account in the development of this project's model and the technologies through which these types of model work.

2.1 Psychological background

This section provides a definition of two important concepts for this work: resilience and word of mouth.

Resilience is a psychological concept studied by researchers in multiple disciplines that refers to the ability to maintain or regain mental health despite having experienced adversity [HSDG+11]. Several factors interact with the enhancement or reduction of resilience of a given individual with a certain amount of influence. For this project, resilience is a state variable that controls the current position of a network agent's vote (belonging either to the PP segment, the PSOE segment, or being abstainer). Its value can be changed throughout the simulation depending on the amount of external and internal influence received, both the interaction between a neighbor and the polarized messages of the mass media and the interaction between a neighbor and the Word Of Mouth (WOM) process with other neighbors.

Word Of Mouth (WOM) stands for the dissemination and interaction between individuals about specific terms for which these individuals have a positive, neutral or negative perspective. Some researchers [Ber14] suggest WOM to split into five different transmitter purposes, such as impression management, persuasion, and information acquisition, among others; all with their own effects on sharing. The one that stands out as regards transmission of polarized content of a certain topic is persuading others. The behavior of this process is shown in Fig. 2.1. This effect causes the spread from agents to other agents of emotionally polarized messages; and in this project, it is modeled in the way neighbors interact using WOM after they are or are not by the most representative channels of this project's mass media scenario on the subject of 11M terrorist attacks.

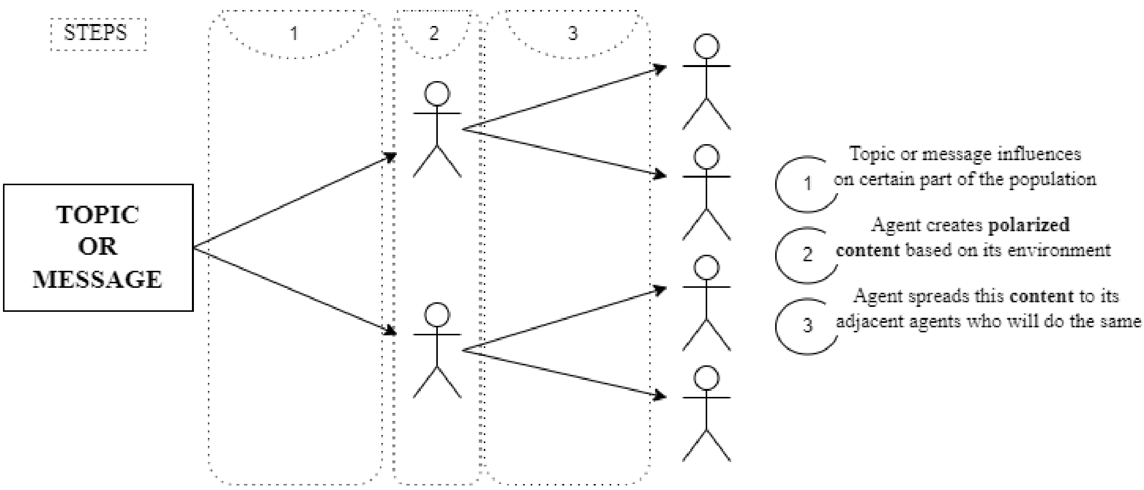

Figure 2.1: How WOM persuasion works

2.2 Agent-Based Models

Modeling is the establishment of how certain things work, data obtained by observational learning, to obtain a representation of the whole system and different ways to analyze and validate it. In our increasingly complex world, agent-based modeling arises due to the need to compute large-scale micro-simulation models [MN09].

2.2.1 Agents

There is no proper theoretical definition for the agents in an agent-based model. In a practical context, within a single simulation, we can find multiple types of agent that behave differently according to their attributes, behavior rules, or even interactions with the environment. Thus, agents are considered to have specific properties that must be defined. Some of the most important are described in Table 2.1.

Agent's property	Description
Autonomy	Self-reliance from the environment and other agents
Interaction	Mechanisms to interchange direct information with its neighboring agents (or other agents indirectly) and with the environment
Situation	Some behaviours rely on its possible interactions, current state and data transmitted or received
Adaption	Approach to suit their attributes better with their environment as the time increases and specific trials are simulated

Table 2.1: Most relevant properties of modeled agents

2.2.2 Simulation

In an ABSS, a simulation is the process in which the model is executed in different ways to obtain the expected or desired states for all agents.

There may be many variables that describe a single simulation. However, some of the most relevant are the following.

- Model of all agents - external and internal to the network - and, optionally, their initial state and parameters
- Environment class and parameters that interact externally with the agents
- Network generator algorithm with its respective parameters or network model with its agents established

- Additional variables such as time steps, events and agents' update interval among others

These variables are often gathered in a single file that is executed with a range of possibilities such as the format you want to store the results or the number of trials to be simulated, among others.

2.2.3 Networks

Aiming at the creation of a more realistic ABSS, the need to build realistic synthetic populations is especially important [ABBSGG15]. From this article, we also extract Figure 2.2. These social networks are basically made up of a number of nodes or agents connected by links that can, but do not have to, be oriented, weighted, or labeled. The formation of these networks can be classified as follows.

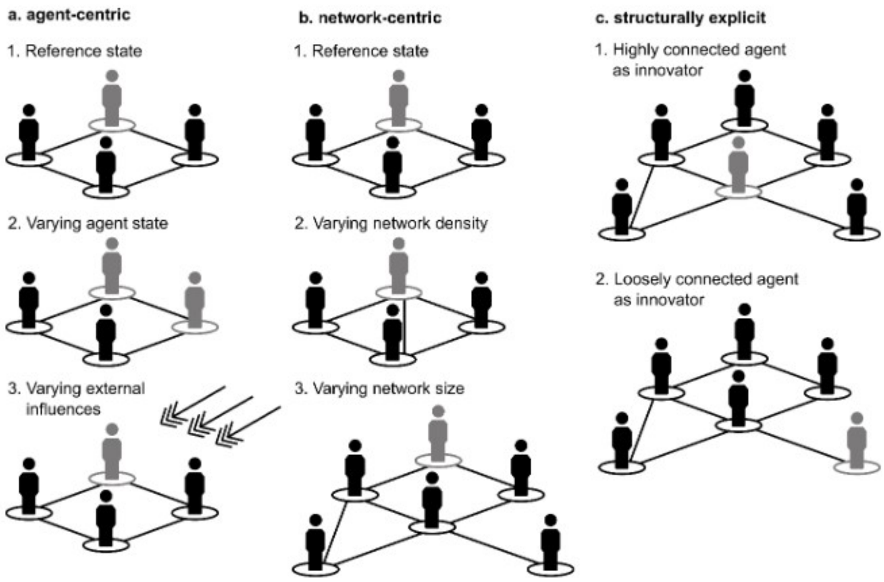

Figure 2.2: Network formation classification [WGFM20]

In this project, the generation of networks is performed in a static agent-centric way, where agents vary their state and internal variables by external influence. Additionally, social network models can be classified according to the type of network that the intended owner wants. Some of the most relevant are shown in Table 2.2. For the project, the Barabási-Albert model is used as an algorithm to generate the simulation network [BA99]. This algorithm employs a stochastic process called preferential attachment, in which new nodes are more likely to join those already linked.

Network type	Popular model	Description
Random network	Erdös-Rényi model	Isolated nodes connect to each other with a given probability
Small world network	β -model (Watts 1998)	Nodes estimates number of adjacent nodes to establish a network with a particular degree
Scale-free network	Barabási and Albert model (2001)	Preferential attachment algorithm generates connections between nodes
Spatial network	-	Nodes are located in a space where they link most likely to close nodes

Table 2.2: Network generator models

2.3 Agent-Based Technologies

This specific section discusses the main concepts of technologies, languages, and libraries that have allowed this project to be carried out. The agent-based model for social simulation is written in Python, over the infrastructure of Soil [SISR17], a specific simulator focused on artificial social networks.

2.3.1 Agent technologies

2.3.1.1 Python

Open source software has become one of the booming technological movements of the 21$^{\text{th}}$ century [CPDRBG14]. Because of that, Python becomes relevant due to its ease of use and its optimized set of tools. This programming language was created in 1991 by Guido van Rossum as a result of working with the Amoeba OS. It is not until 1999 that Python's creator proposes to the Defense Advanced Research Projects Agency (DARPA) the development of a computing curriculum suitable for creating more effective and simpler tools for analysis and program development [NG19]. This is why today Python is much less complicated than other environments and improves the way computational science and artificial intelligence algorithms are done. In point of fact, Python is one of the most popular programming languages globally, highlighting it in Data Science, Internet of Things and Machine Learning applications where ABSS are included, as shown in Figure 2.3.

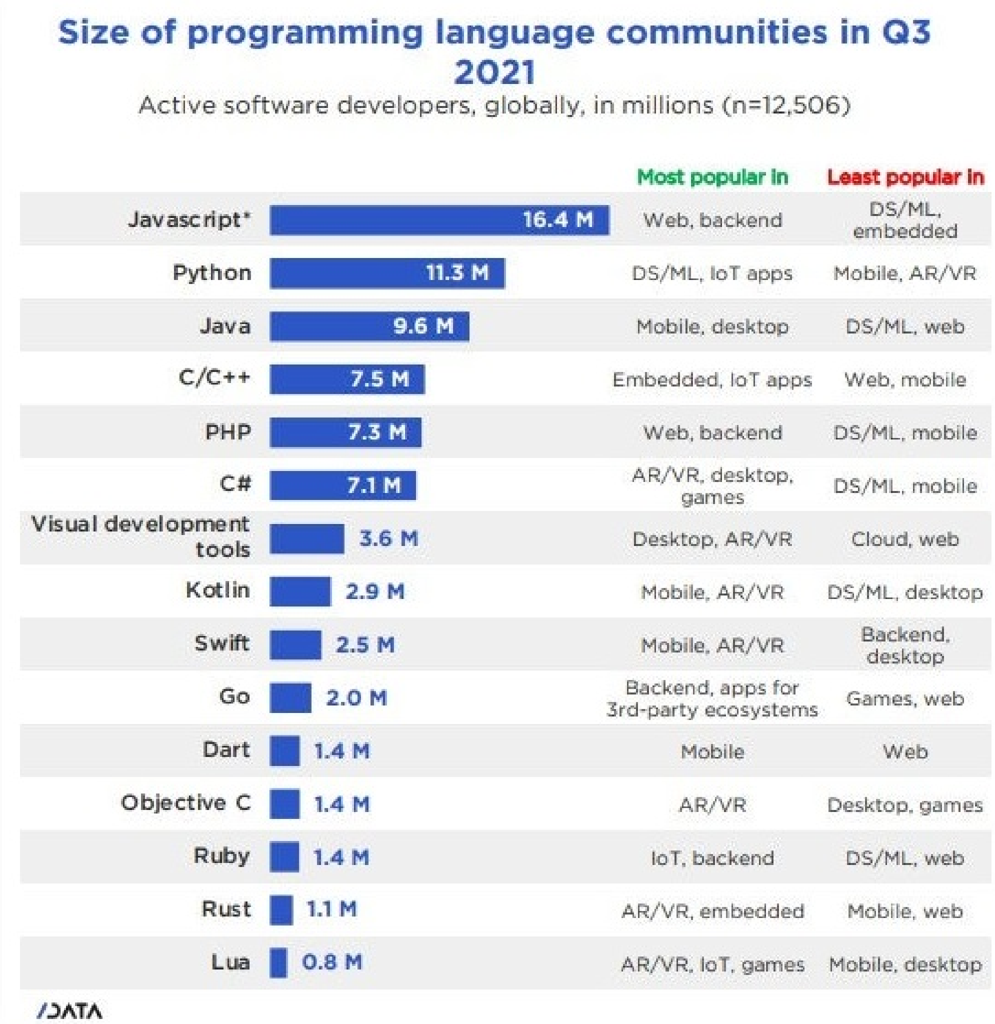

Figure 2.3: Top software development programming languages [RG21]

2.3.1.2 Soil

Soil is a Python framework for ABSS [SISR17] created by GSI that focuses on modeling and simulation of specific social scenarios. These simulations are configured either programmatically or by their corresponding YAML files [BKEI09], which include all the features to be taken into account. Moreover, Soil makes use of NetworkX's built-in methods to initiate and modify multiple types of complex networks that provide graph-geometric generator models. Specifically, NetworkX [HSSC08] is a Python package used for complex network analysis and structure study.

Soil offers insight into the construction of a ABSS divided into two main elements: agent classes and simulation configuration. Agent classes define the behavior of one or more agents throughout the simulation process. There are two agent roles: network and environment agents. On the one hand, network agents are assigned to a particular node in the topology, and they can interact between them and with the environment. On the other hand, environment agents are not part of the generated artificial network, which means that they are freely assigned to the environment. In a soil simulation, it is not mandatory to program them, as their only ability is to control environmental variables and functions.

The simulation configuration must include all the necessary parameters to run the simulation. The most relevant are the following.

- The total number of agents to be simulated, their type, which means what behavior are they following through the simulation, and their respective weights. In this project, this last parameter has been used in the model to establish a 42% of initial PP voters, a 35% of initial PSOE voters, and a 23% of initial abstainers.
- Network topology and parameters which may include a network generator algorithm, number of nodes (must be equal to number of network agents), and other variables such as a network degree.
- Initial values of all environment parameters that are going to be used.
- The number of trials that illustrate how many times the simulation is going to be executed and the maximum number of steps that can be performed in a single simulation.

As will be shown in Fig. 2.4, Soil simulates the agent network based on the different configuration files. Once the configuration is loaded, the Agent-Based Social Simulation can be performed step by step. All the status and parameters of all agents are stored when the simulation ends in four different files. A pickle of the simulation. A YAML or JSON file that shows how the simulation has been performed and can be used to rerun it. One or more CSV or SQLITE files, depending on the number of trials we had defined, that store all the information for all steps, agents and variables. A GEXF file that can be visualized to see the network topology used. This cluster of archives has been used in this project both for the development of an agent-based model and for the multiple-time running of the entire social simulation.

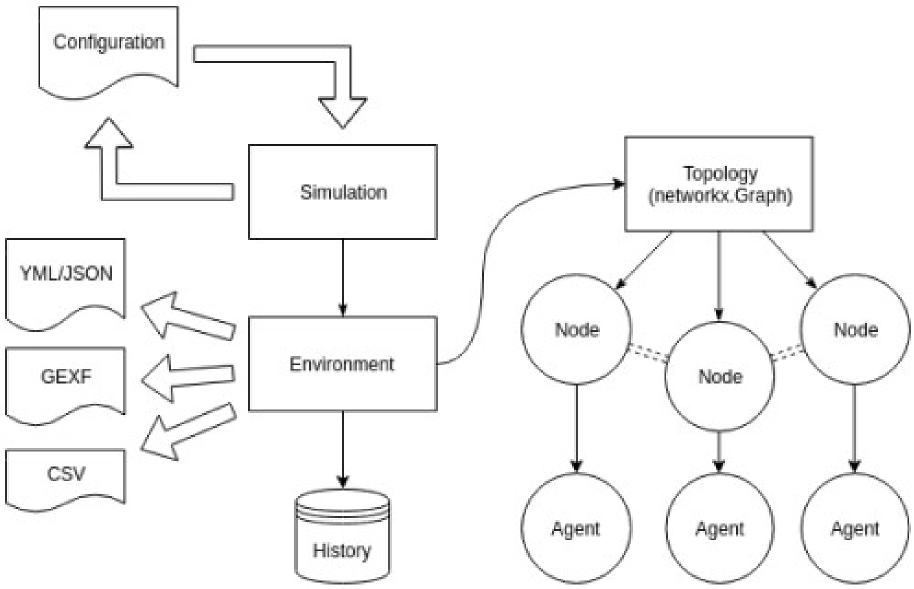

Figure 2.4: Soil relations for Agent-Based Social Simulation

2.3.2 Data analysis technologies

2.3.2.1 Pandas

Pandas [MtPDT22] is an open-source Python library that provides easy-to-use high-performance data structures and data analysis tools as defined in its official documentation. The library provides a solid foundation on which a huge data analysis ecosystem can be established [M+11]. The Panda library has been under development since 2008 by Wes McKinney and the Pandas Development Team, whose last release, Release 1.4.1, was published in February 2022. In this project, Pandas has been used in the analysis and visualization components to read and process DataFrames in CSV format and plot the results obtained.

2.3.2.2 Jupyter

Jupyter [Jup22b] is a web-based free software environment where multiple languages code, data and notes coexist in the same notebook. Since this project's model has been developed in Python with Soil, Jupyter has been used as an organizer document where you can import libraries such as *matplotlib.pyplot* and *pandas*.This is achieved thanks to Jupyter's IPython

Kernel, the responsible of running our analysis and visualization code. More specifically, GSI's JupyterHub has been used, which allows remote work from any computer without the necessity of installing any dependencies. Unfortunately, despite all this, many of the results obtained need to be extrapolated to a real world number of voters. For this reason, we use another tool that allows plotting large amounts of data stored in CSV files.

2.3.2.3 Datawrapper

Datawrapper is a free-to-use browser tool that is used to generate all types of graphs, maps, and tables from multiple types of input data [Dat22]. We have chosen this application to adapt and match some of the results obtained in the graphs of its counterparts shown in the article [MCSLC17]. Another reason is that the mentioned tool offers simplicity and clarity in showing rather confusing information in a very clean way.

CHAPTER 3

11M model description

In this chapter, the ABSS developed for this project will be described in detail. From a global perspective, this project builds four different models.

- Model 1. An empty non-interacting simulation where agents remain with their initial values until the end. In other words, the results predicted in pre-election surveys before attacks occurred (Sect. 3.2.1).

- Model 2. A simulation in which neighbor-to-neighbor WOM interactions are erased and only the mass media influences the agents (Sect. 3.2.2).

- Model 3. Unlike the previous, a simulation where mass media influence is ignored and only neighbors interact between them (Sect. 3.2.3).

- Model 4. A complete simulation where all interactions are activated (Sect. 3.2.4).

To achieve it, the first thing to take into account is the elaboration of a realistic model based on what the mass media transmitted in 2004 and how it did so.

3.1 Penetration of mass media in 2004

In order to determine how to approach this project's model to what actually happened, one needs to study the effect and penetration of mass media in the Spanish population when 11M attacks occurred. For this purpose, this section collects relevant data from AIMC general overview (February - November 2004) [plIdMdC04] and AIMC general framework of Spanish mass media (2005) [plIdMdC10].

Both statistical studies reflect the Spanish society talking about its consumption of the media. Data are also specified for different types of media where the press, radio, and television play the most important roles. As in the article, the selection of mass media channels follows the criterion in which the most relevant channels that emitted at a national scale and included polarized messages must be modeled, and they are shown in Table 3.1.

CHANNEL NAME	El País	El Mundo	ABC	Cadena SER	TVE	Antena 3	Telecinco
MASS MEDIA TYPE	Press			Radio	Televisión		

Table 3.1: Selected mass media channels

First, Table 3.2 illustrates an essential summary of the audiences that are related to the three types of media that are modeled.

	MASS MEDIA	**CONSUMERS PER DAY**	**RELATIVE PENETRATION**	**TOTAL CONSUMERS**	**TOTAL PENETRATION**
PRESS	El País	2155	14'39%	14980	41'1%
	El Mundo	1314	8'77%		
	ABC	903	6'03%		
RADIO	SER	5180	45'18%	11466	31'5%
TELEVISION	Telecinco	3276	10'04%	32632	89'6%
	TVE	3139	9'62%		
	Antena 3	3113	9'54%		

Table 3.2: Mass media audience

Second, since these mass media channels generate different messages through the steps of ABSS, the audiences of their messages must be distinguished depending on the time they are emitted. Table 3.3 shows a general overview of the share of television screens at different time slots.

	MORNING (4 AM - 1PM)	**AFTERNOON (1PM - 5PM)**	**EVENING (5PM - 8PM)**	**NIGHT (8PM - 4AM)**
PERCENTAGE PER TIME SLOT	6%	26%	15%	53%

Table 3.3: Television audience

When it comes to radio, Table 3.4 presents, in the first row, data on the time of day that corresponds to the consumption of general radio in the second row and the audience of the radio station SER in the last row.

6:00 AM	**7:00 AM**	**8:00 AM**	**9:00 AM**	**10:00 AM**	**11:00 AM**	**12:00 PM**	**1:00 PM**
4.0%	11.0%	16.9%	18.6%	18.0%	15.7%	12.4%	9.2%
0.2135%	0.59%	0.90%	0.99%	0.96%	0.84%	0.66%	0.49%

2:00 PM	**3:00 PM**	**4:00 PM**	**5:00 PM**	**6:00 PM**	**7:00 PM**	**8:00 PM**	**9:00 PM**
5.9%	4.1%	6.0%	7.7%	7.6%	6.5%	5.9%	5.5%
0.31%	0.22%	0.32%	0.41%	0.41%	0.35%	0.31%	0.29%

10:00 PM	**11:00 PM**	**0:00 AM**	**1:00 AM**	**2:00 AM**	**3:00 AM**	**4:00 AM**	**5:00 AM**
6.9%	9.3%	11.4%	6.3%	3.2%	2.0%	1.3%	1.2%
0.37%	0.50%	0.61%	0.34%	0.17%	0.11%	0.07%	0.06%

Table 3.4: Radio audience

Making a brief digression about these time slots, it must be considered that the way press works is different from the other two types. While radio and television are broadcast according to their programs, the press generates one newspaper per day. With the goal of simplifying these daily papers, they are divided in Table 3.5 into three main sections with different scopes: front-page, news, and opinion pieces.

	TOTAL AUDIENCE	**FRONT PAGE**	**NEWS**	**OPINION PIECES**
El País	5.91%	5.91%	1.48%	0.59%
El Mundo	3.60%	3.60%	0.90%	0.36%
ABC	2.48%	2.48%	0.62%	0.25%

Table 3.5: Press audience

Finally, and taking into account that in this project's ABSS all messages can be received by any agent, the previous tables are the basis of the generation and adjustment of polarized messages' audiences.

3.2 Models description

Breaking down the whole code, the following items represent the main elements of all the models, identifying which is used in which model.

- **Environment.** Consists on a large group of messages that are released in a specific step of the simulation. They are described in Sect. 3.3.1. All models have an environment, but in Models 1 and 3 this environment does not diffuse information to the agents, so there is not an agent-environment interaction.

- **Neighbor.** The only type of agent that represents the behavior of over 18 voters with access to mass media. All models have a linked agent network, but in Models 1 and 2, agents do not interact between them.

- **Mass media channel.** Not an agent itself, but a set of parameters with the ability to modify the environment and change some agents' variables. These parameters are only defined for Models 2 and 4, where mass media have influence in the population.

It is needed to take into account that the simulation runs in different steps in which the period between 11M 8:00 AM and 14M 8:00 AM is simulated with the correspondence of one step for each hour; generating a number of steps calculated according to Eq. 3.1.

$$\begin{gathered}
Number\ of\ steps = 3[simulated\ days] \cdot 24[steps/day] = 72 \\
11M- > 1^{\text{st}}step - 16^{\text{th}}step \\
12M- > 17^{\text{th}}step - 40^{\text{th}}step \\
13M- > 41^{\text{st}}step - 64^{\text{th}}step \\
14M- > 65^{\text{th}}step - 72^{\text{nd}}step
\end{gathered} \tag{3.1}$$

In these models, only one type of agent can be found: the neighbor. The class diagram represented in Fig. 3.1 shows all of the agent's parameters that appear in the Neighbor class that inherits from the soil.agents. Finite-State Machine (FSM) class in order to eliminate the need of specifying a step method and to enable controlling the simulation's flow with the transitions between states.

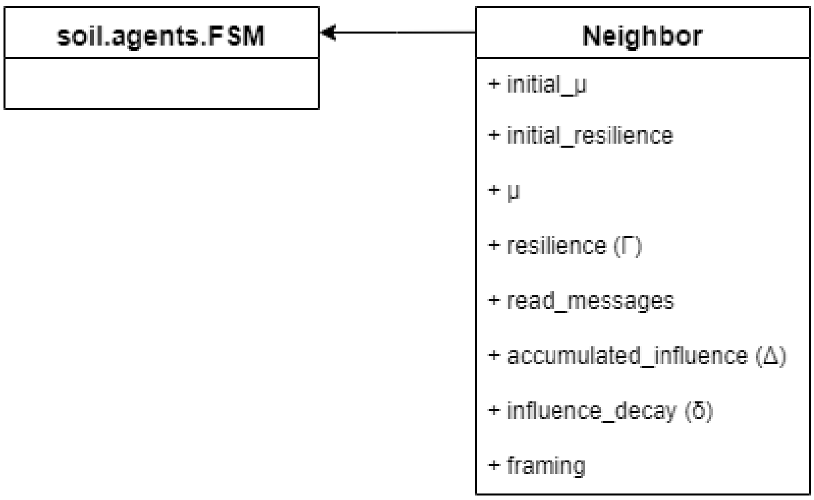

Figure 3.1: All agent variables

It has already been described how the polarized content of every message works. Every agent manages this effect by operating a status variable defined as resilience (Γ). This variable is modeled by another positive real variable represented as μ and is confined to the interval [0,10). As introduced in Sect. 2.1, resilience represents the current political position of an agent with respect to its three possibilities. The resilience value will change for each agent through the simulation, depending on the amount of information it receives from the media and other agents. As shown in Equation 3.2, this value is a positive integer in the range [0,2], where each of the three possible values represents its potential final votes. The "0" result is for the agents who vote PSOE and the "2" result for those who vote PP. The function returns "1" for agents who decide to abstain from voting.

$$\Gamma_i(\mu_i(t)) = \begin{cases} 0 & , \text{if } \mu_i(t) \in [0, 3.333), \\ 1 & , \text{if } \mu_i(t) \in [3.333, 6.667), \\ 2 & , \text{if } \mu_i(t) \in [6.667, 10). \end{cases} \tag{3.2}$$

As the simulation runs at each step, resilience is the last variable to be computed because it is directly dependent on μ, the variable that really changes for every interaction. And once resilience has been defined, we can evaluate in the simulation results how it changes over the simulated days (11M - 14M). However, it is interesting to have a specific parameter showing the changes in resilience from the beginning to the end of ABSS, which only obtains a value at step = 72. We define it as the framing effect, whose possible outputs are illustrated in Table 3.6, and it can change its result for each agent to analyze the population's vote tendency. In this project, we do not deepen our understanding of the framing theory of how a message can be perceived and treated by an individual.

Just after 11M attacks **Step = 0**	**Just before the elections** **Step = 72**	**Framing effect** **on agent**
PP [2]	PP [2]	Reinforcing
PSOE [0]	PSOE [0]	Reinforcing
Abstainer [1]	Abstainer [1]	Reinforcing
PP [2]	Abstainer [1]	Deactivating
PSOE [0]	Abstainer [1]	Deactivating
Abstainer [1]	PP [2]	Activating
Abstainer [1]	PSOE [0]	Activating
PP [2]	PSOE [0]	Converting
PSOE [0]	PP [2]	Converting

Table 3.6: Possible values of the framing effects [MCSLC17]

Despite still taking about environment's messages, every agent stores all the messages it has already received in a list called read_messages. Therefore, Agent-Based Models 2 and 4 detect when an agent has been influenced with a specific message to avoid repeating the process.

Thus far, we have not talked about how agents are able to scatter its perceptions towards the rest of agents inside the social network. This WOM interaction that takes place in models 3 and 4 and which we talked about in Sect. 2.1 represents a one-way conversation where an agent talks to its neighbors and they listen. This "listening" is modeled by a random variable, independent of every conversation, that controls the influence change value of an agent in its neighbors.

Furthermore, each agent has a talking probability $p_t \in (0, 1]$, an environmental common variable for neighbors in the same political segment. When a WOM interaction passes through this talking probability, the way in which influence is exercised over others follows Equation 3.3. In this interaction, *inf* represents a random number between "0" and "1".

$$\mu_i(t) = \mu_i(t) + [\mu_j(t) - \mu_i(t)] * \ inf \tag{3.3}$$

Therefore, given a certain step t, an agent i is influenced by another agent j with random *inf* if they are directly connected through the simulation network, which means they are neighbors. It will be seen in Sect. 3.3.2 how this influence is mitigated as the simulation time goes by.

In the next sections, we will differentiate how a Neighbor agent behaves through the simulation depending on the model we are running.

3.2.1 Model 1: No interaction model

This model builds the basis for the rest of the models. Figure 3.2 establishes the basic start and end of the workflow of each agent. In all models, the real FSM states are represented with capital letters, whereas the rest are intermediate states, where different functions or calculations are computed.

First of all, Soil initializes every agent's variables (illustrated in state "setting variables") just before an agent is placed in the artificial network. The initialization ("start") function of the agent's class assigns their initial variables to obtain their parameter of the framing effect at the end of the simulation. After checking that this has been done ("initial μ assigned"), agents do not change their variables at all. "Next step" stands for the preparation of the simulation to proceed to the next step. Once the simulation is almost over (current_step = 71 represented as "14^{th} March") every agent jumps into a voter state (always in the last step) where their individual framing effect is computed, and they die. Otherwise, the agents are still getting infected or, in this first model, agents do not interact at all.

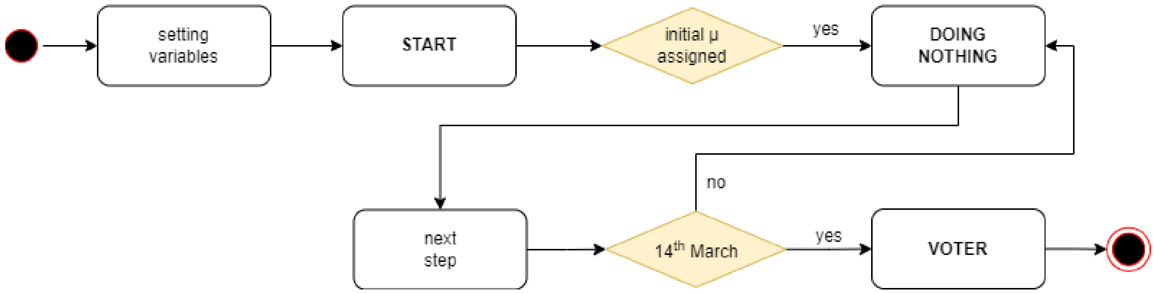

Figure 3.2: Agent's behavior through the simulation in Model 1

3.2.2 Model 2: No WOM model

From this model on, the new features implemented with respect to their previous models are shown in green. Model 2 in Figure 3.3 mass media as the first interaction with network agents, so, instead of doing nothing, the agents are now infected with mass media messages at every step except the last. In this project, infection stands for the action of receiving influence from any source.

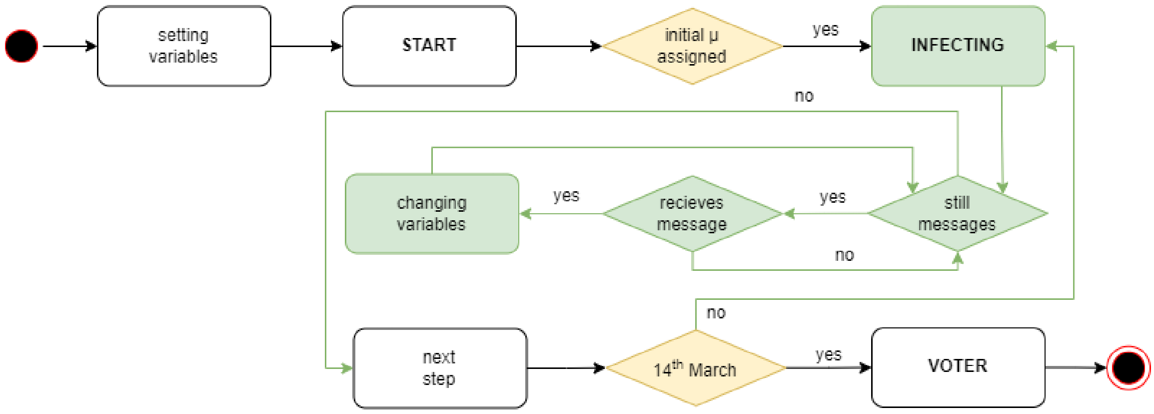

Figure 3.3: Agent's behavior through the simulation in Model 2

The program passes through a loop with all the available messages given a certain step. If they are messages that are still to be transmitted to the population ("still messages"), then the audience parameter of each message acts like a barrier. For example, if a message has an audience value of "0.1", a function called *prob* returns True for approximately 10%

of the total population. In this way, "receives message" checks if a particular agent has been influenced by that message and changes its μ value. If there are no messages left, the "next step" computes $\Delta, \mu and \Gamma$ (Section 3.3.2) and performs the previously explained model test, following the same behavior as in Model 1 in the last step.

3.2.3 Model 3: No media model

This model erases media diffusion, introducing WOM interactions between neighbors. Figure 3.4 represents in green how the "infecting neighbors" check searches for all the neighbors of a particular agent making use of a Soil function called *get_neighboring_agents*. In a way that recalls Model 2, "still neighbors" test if a particular agent has tried to talk with all its neighbors with a probability of ρ and they are still neighbors to establish a conversation with. When this check fails, the model computes the Γ value for each agent and follows the same behavior as seen in previous models.

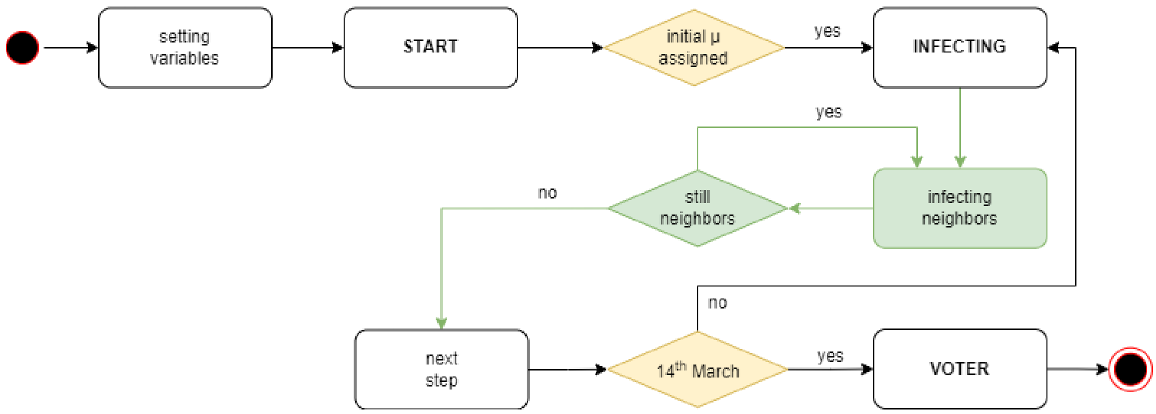

Figure 3.4: Agent's behavior through the simulation in Model 3

3.2.4 Model 4: Complete model

Lastly, Figure 3.5 shows a new connection between models 2 and 3, establishing the states relationship of the complete model that will be analyzed. When an agent has finished being influenced by mass media, it begins infecting its neighbors. All the interactions are treated as in the previous models with the extra of combining them all. The result of this combination was indeed the second model to be designed after Model 1. To lead to better understanding of this project scenario, Model 4 had to be split into Models 2 and 3. Those two models were more flexible and easier to calibrate some of their variables without changing the workflow of the rest of the simulation.

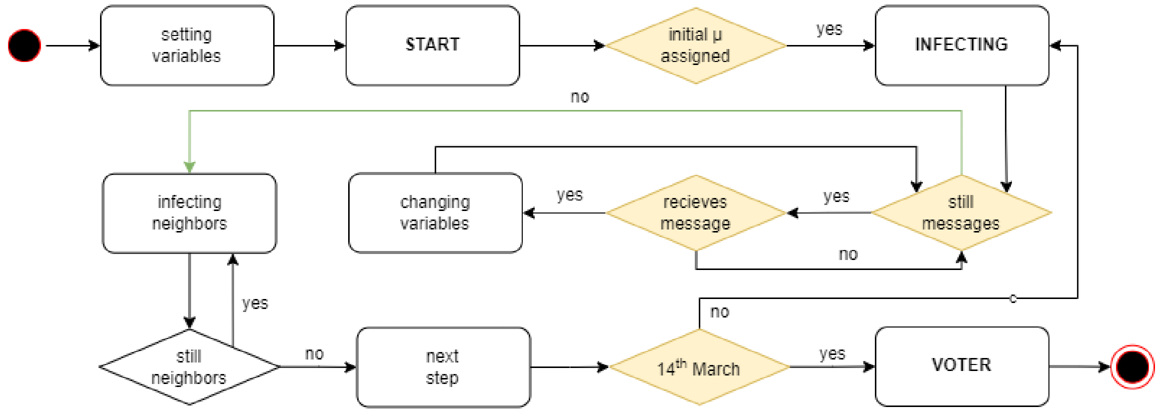

Figure 3.5: Agent's behavior through the simulation in Model 4

3.3 Other Model Elements

We have already seen the most important element that always has a presence in all models: the neighbor agent. In this section, we expand on the rest of the elements presented in the first list of Sect. 3.2 where we already talked about the elements that act in each model.

3.3.1 Simulation environment and message properties

As mentioned previously, the simulated environment consists of a heterogeneous group of messages with specific parameters. All messages are stored as a list in a CSV file imported by the Python model for its reading. It is important to note that the messages have a polarization value modeled as an integer number that represents the information provided about the perpetrators of the attacks as displayed in Table 3.7.

The polarization variable is assigned a 2 or -2 value, depending on whether the message strongly points to ETA or Al-Qaeda, respectively. If the message blames in a weaker way, a 1 or -1 is allocated. For lack of attribution of attacks' authority, polarization takes value 0.

For this and the rest of the parameters, Table 3.8 shows how each message works as a Python dictionary with multiple keys. Audience is computed individually for each message, checking the tables in Sect. 3.1. It is required to sum the appropriate audiences' percentage if a message is transmitted for at least two hours (two different steps). Normal messages are diffused in less than one hour, but as the minimum measurement in a Soil simulation is of one step, a 10 minute message, for example, is diffused among the same hour on equal terms as a 50 minute message.

POLARIZATION	DESCRIPTION
2	Message strongly points ETA as perpetrators.
1	Message weekly points ETA as perpetrators.
0	Message doesn't point specific terrorist groups.
-1	Message weekly points Al-Qaeda as perpetrators.
-2	Message strongly points Al-Qaeda as perpetrators.

Table 3.7: Possible values for messages' polarization

KEYS	DESCRIPTION	POSSIBLE VALUES
name	Mass media channel that transmits the message	El País, El Mundo, ABC, SER, TVE, Antena3 or Telecinco
start	Message start hour (step)	[0,72)
end	Message end hour (step)	(0,72]
type	Mass media channel type that transmits the message	Television, radio or press
message polarization	Message's level of polarization	[-2,2]
audience	Percentage of the population with the potential of being influenced by the message	(0,1]
segment	Message transmitted by a person affiliated to a specific political party	PP, PSOE, both or none

Table 3.8: All keys and their possible values of a message

To achieve the desired environment of this simulation, once all the messages are defined, two functions need to control which messages are transmitted to the population given a specific step.

- Function **filter_messages**. This function receives all messages stored in the environment (transmitted between 11M and 14M) and returns those that are generated within the current step. For example, if Telecinco emits a message between 8 AM and 11 AM the first day, this message must be available to be received by the agents from the environment at steps 1, 2 and 3. The programmatic condition to achieve this is the following.

$$message['start'] <= current\_step \ and \ message['end'] > current\_step$$

- Function **update_messages**. This function ensures that the selected messages from the previous function are updated in the environment before agents begin to be influenced by them for a specific step.

The rest of the environmental variables are discussed in Sect. 3.2 and Sect. 3.3.2.

3.3.2 Mass media channels

In this section, we discuss how all the channels of the mass media that are disaggregated in Sect. 3.1 participate in the environment of this project ABSS.

In order to simplify the model, the mass media behaves as an entity that does not really exist inside the simulation but provides the necessary tools to extract information from the messages received by the population. This is achieved by including particular parameters to examine and analyze the response of agents to specific messages. These parameters the following.

- **Inside messages:** Name and type of the mass media channel, its polarization, and whether the content of the message belongs or has been transmitted by a member of any political segment. In addition, the audience of each message depends on the hours and period of time in which it is transmitted.

- **Environment variables:** To analyze the results obtained after simulating the model, we use different variables to assess the average and maximum influence of the mass media in the 14^{th} March general elections. Some of them are the influence of the press, the influence of radio and the influence of television.

The way in which mass media messages influence and infect the population follows a behavior similar to the interaction between neighbors, and is described in Equation 3.4, where *inf* is again a random variable between values “0” and “1”.

$$\mu_i(t) = \mu_i(t) + message['polarization'] * inf \tag{3.4}$$

One of the most important functions of the model regarding mass media is the ability of an agent to maintain the influence received from a message for the rest of the simulation and how this influence is mitigated through the simulation. This effect is produced by the combination of two agent parameters previously shown in 3.1: accumulated influence and influence decay. Firstly, the accumulated influence (Equation 3.5) measures the difference between two consecutive steps μA and multiplies that difference with the influence decay, accumulating this value each step for a particular agent i.

$$Accumulated\ influence_i = \Delta_i = \sum_{t=1}^{t} [\mu_i(t) - \mu_i(t-1)] * \delta_i \tag{3.5}$$

Lastly, the result obtained indirectly modifies the resilience at the end of the step by altering the value of μ as shown in Equation 3.6.

$$\mu_i = \mu_i - \Delta_i \tag{3.6}$$

3.3.3 Artificial network

As additional information, we have developed a small program in Python inside Jupyter’s interface to export our 764 linked agents network to a PNG image. Despite this, Figure 3.6 represents only the northwest sector of our complete artificial network, because of the huge volume of bytes that a network of this kind needs to be visualized, where each agent is drawn as a blue node and is linked to its neighbors with black straight lines.

Figure 3.6: Artificial network northwest sector

3.4 Setting up the simulation and calibration of variables

Table 3.9 samples a summary of the specifications that must be set up before the simulation starts to run. All the following parameters are included in the YAML file where ABSS is configured.

Specification	Description	Value(s)
Number of trials	How many times we run the simulation	1
Max time	Maximum number of steps of the simulation	72
Network generator	Network topology used	Barabasi and Albert model (scale free network)
Number of network nodes	Number of agents connected in the network's model	764
Network degree	Node degree correlation	1,2,3,4,5,6
Probability of neighbor spread	Talking probability to neighbors	1%,2%,3%, 4%,5%,6%
Influence decay	The total decay that influence suffers every step	0.0078, 0.0039, 0.002
Agent distribution	Distribution of agents in accordance with their initial political preferences	35% of initial PSOE voters 23% of initial abstainers 42% of initial PP voters

Table 3.9: Initial simulation specifications and their possible values

In this section, we must introduce the purpose of calibrating this project's model. Making a lot of agents and variables work together to behave as their are expected to do so, enabling us to extract valid results from their corresponding trials, is a tedious task. For that reason, a trial-and-error integration process is required to be performed, in which some environmental parameters are tested at a time to shape the model correctly.

The main purpose of this process is to adjust not only the results, but also the procedure that all agents follow to act as similar to real-world people as possible. The parameters tested for this purpose were:

- **Network parameters.** Because of the gigantic amount of resources required to simulate a model with miles or even millions of agents, it is necessary to take a smaller group to sample it after for all Spanish beyond the 18 population. Knowing that this ABSS real target is composed of the total sum of 29,238,662 abstainers, PSOE and PP voters, we can obtain all the prime numbers of that total: 2, 191, 382, 76541, 153082 and 14619331. First, we took the decision to simulate a total of 382 agents; but, doing a brief sensibility analysis, where for a larger amount of agents, a greater precision for the model was obtained; it was determined to use twice the agents. That is, a total of 764 agents connected with a network degree of m = 2, leading to an agent / voter map of 1: 38270.5.

- **Influence decay.** This decay is a special model variable that models how quickly an agent forgets the previous received influence. After many tests, the optimal value assigned to each type of voter is shown in Eq. 3.7, taking into account that the mass media transmitted messages during 64 steps (64 hours).

$$\begin{aligned} Influence\ decay_{PP} = \delta_{pp} = \frac{1}{64} \ * \ 0.5 = 0.0078 \\ Influence\ decay_{PSOE} = \delta_{psoe} = \frac{1}{64} \ * \ 0.25 = 0.0039 \\ Influence\ decay_{Abstainer} = \delta_{abstainer} = \frac{1}{64} \ * \ 0.125 = 0.002 \end{aligned} \tag{3.7}$$

- **Talking probability.** This parameter measures the probability with which an agent starts a conversation with its adjacent agents. After many tests, the optimal value assigned for each type of voter is shown in Eq. 3.8, taking into account that PSOE had more impact in voters' events such as the 13M protest in front of PP's site.

$$\begin{aligned} Talking\ probability_{PP} = \rho_{pp}(t) = 0.02 \\ Talking\ probability_{PSOE} = \rho_{psoe}(t) = 0.04 \\ Talking\ probability_{Abstainer} = \rho_{abstainer}(t) = 0.01 \end{aligned} \tag{3.8}$$

CHAPTER 4

Analysis of the results

In this section, the most relevant results obtained as the outcome of this project ABSS are presented, contrasted, and analyzed. The first table shown 4.1 represents the raw information obtained that can be extracted directly from the files obtained after performing several simulations. Note that these data are specified in the number of agents. If we want to extrapolate to the real world, we need to increase its value (now in voters) by a factor of 1:38,270.5 as mentioned in Sect. 3.4.

To make an accurate comparison between our ABSS and what actually happened in 2004, we computed the absolute deviation between the percentage of votes in each model and in the real world, as shown in Eq. 4.1.

$$Absolute\ deviation = \sum_{i=0}^{2}(|percentage_i(real) - percentage_i(model)|) \tag{4.1}$$

Table 4.1 shows the results of the 2004 Spanish General Election. These data are directly extracted from the Votes and Seats Distribution posted by Constitución Española [Esp04] and the previous voting intentions are taken from Centro de Investigaciones Sociologicas (CIS) [CIS04], specifically from a national survey with interviewers of both sexes.

	Real world	**Models**			
Political parties	**2004 Election results**	**Complete**	**No media**	**No WOM**	**No interactions**
PP	9,763,144	255	272	308	321
PSOE	11,026,163	288	281	270	267
Abstainers	8,449,355	221	211	186	176
Total votes percentage					
PP	33.39%	33.38%	35.60%	40.31%	42.02%
PSOE	37.71%	37.70%	36.78%	34.34%	34.95%
Abstainers	28.90%	28.93%	27.62%	24.35%	23.04%
Absolute deviation		0.05%	4.42%	14.84%	17.25%

Table 4.1: Direct results and comparison with reality

4.1 Resilience results

Henceforth, all the figures and result graphs can not be observed at first sight from the output of the model, but are extracted from filtering and correlating variables and values.

First, we will show multiple figures of how resilience evolves through simulation, depending on the model with which we are working. Starting with the complete model and as seen in Figure 4.1, the functions represent the ups and downs of the average resilience of all agents belonging to the same segment in a given step of the model simulation. In its complementary Figure 4.5, it is shown how this evolution of resilience varies through simulation. When the function hits values greater than zero, the resilience suffers an increment with respect to its previous step value; on the other hand, if it has values below zero, then the resilience decreases. The evolution of average resilience in Model 4 should the most similar behavior to what happened in reality, the combination of mass media diffusion and interactions between neighbors.

We can also see that the three curves in the previous referenced figures have very similar behaviors caused by the interactions occurring during the simulation, which affect all agents in their measure, depending on their resilience value. This observation leads to the inference that the three political segments are more correlated than we expected them to be, and they try to compensate their deviations when neighbors from different political segments interact. For example, when a high-polarization message hits the population, even if a segment of the population is not very affected by it, the rest of the infected agents will balance their neighbors by spreading their new ideas.

The no-WOM model samples a behavior very similar to that of the complete model. The average resilience varies through the simulation in Figure 4.3 almost remains the same compared to the first figure we saw. This is mainly due to the interaction between mass media channels and agents; and how the most abrupt changes are direct consequences of this interaction. Figure 4.7 shows exactly this effect, with a behavior similar to the previous one and slightly higher polarization values.

Despite the fact that the no-media model outputs results closer to the complete model results than the other two models, its resilience change behavior looks nothing like it. Figure 4.2 shows clearly how the average resilience value for PP voters decreases over the simulation, as well as the abstainers' one. However, abstainers' resilience declines slowly because of the framing effect of PP voters deactivating while decreasing their resilience little by little. PSOE voters' average resilience does increase through the simulation due to the framing effect of abstainers activating and to the fact that there is no external influence (mass media in our case) that affects directly to them.

Moreover, this special model does not follow a specific pattern every time it is simulated. The lack of a series of events that model the environment leaves all network agents with the randomness of behaving differently each time they are launched in a new trial. This effect can be seen in Figure 4.6 where the average resilience deviation takes random and very low values with the main characteristic of remaining most of the time below zero for the PP segment and abstainers and above zero for the PSOE segment.

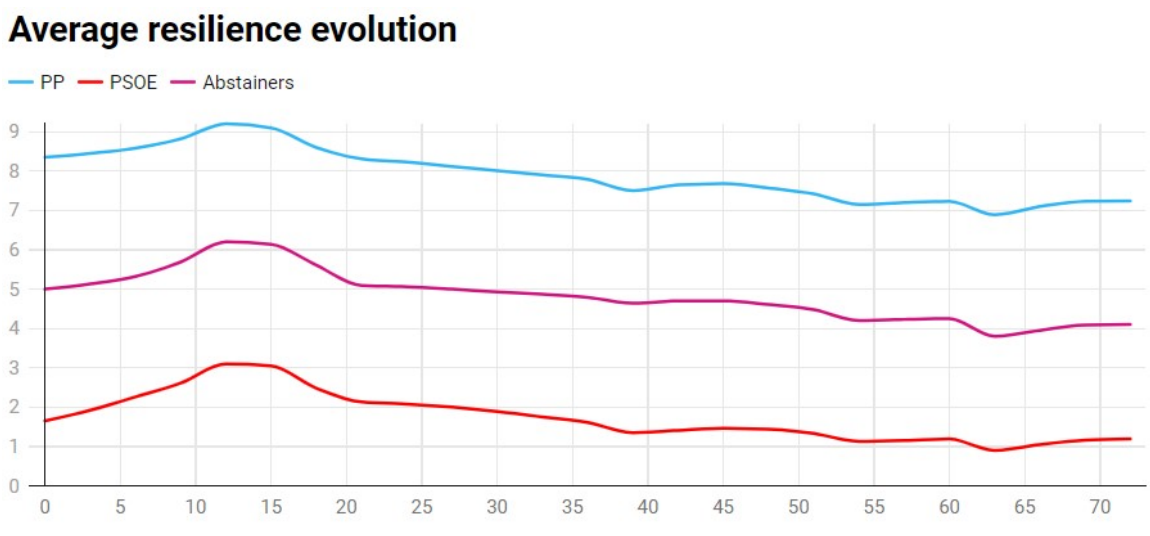

Figure 4.1: Average resilience evolution (Model 4)

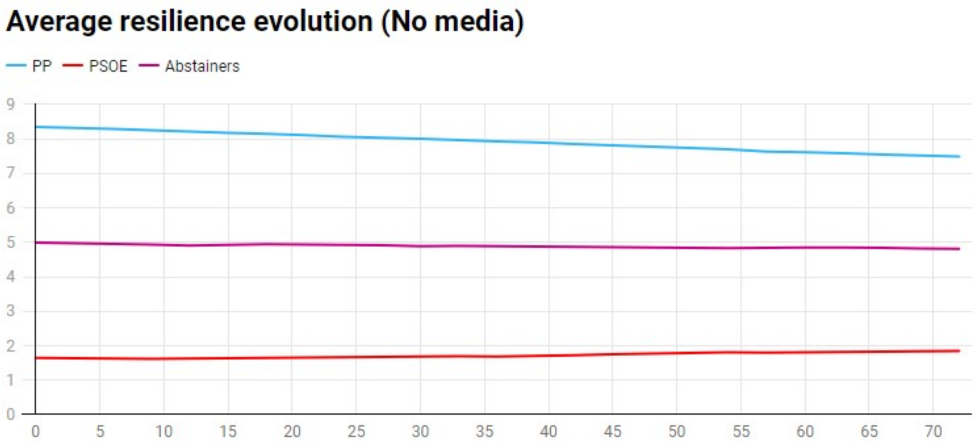

Figure 4.2: Average resilience evolution (Model 3)

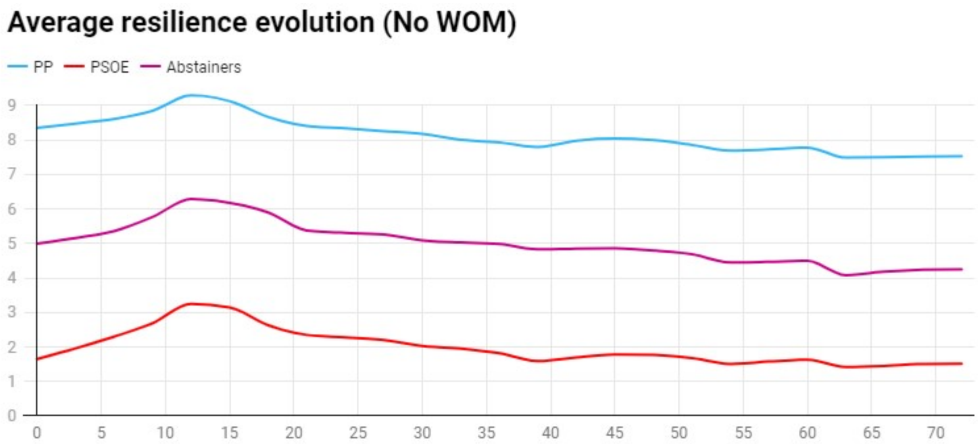

Figure 4.3: Average resilience evolution (Model 2)

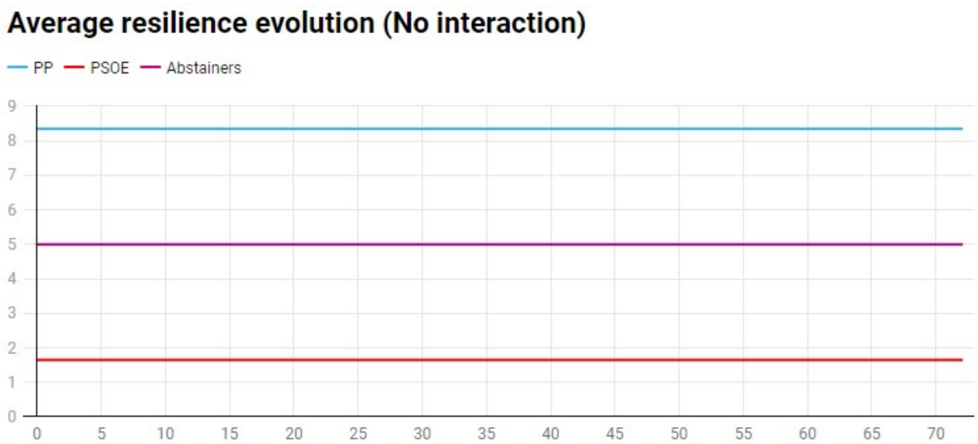

Figure 4.4: Average resilience evolution (Model 1)

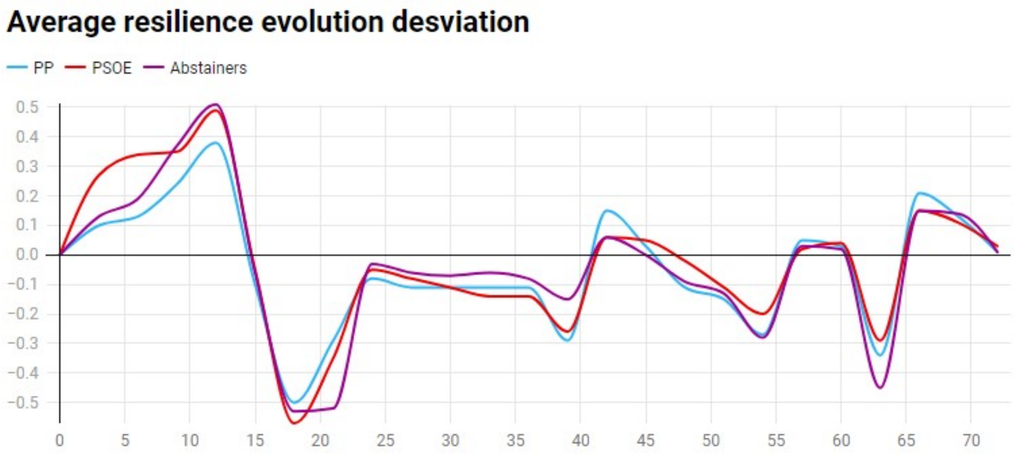

Figure 4.5: Average resilience evolution deviation (Model 4)

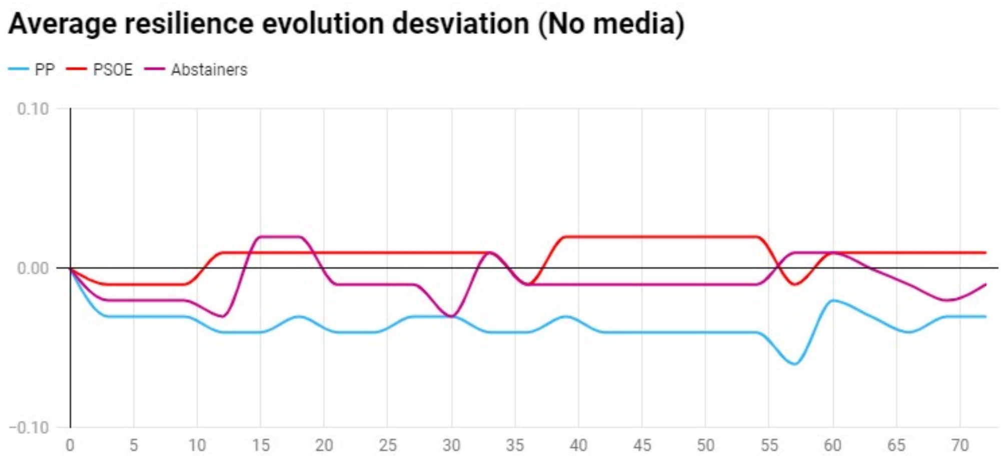

Figure 4.6: Average resilience evolution deviation (Model 3)

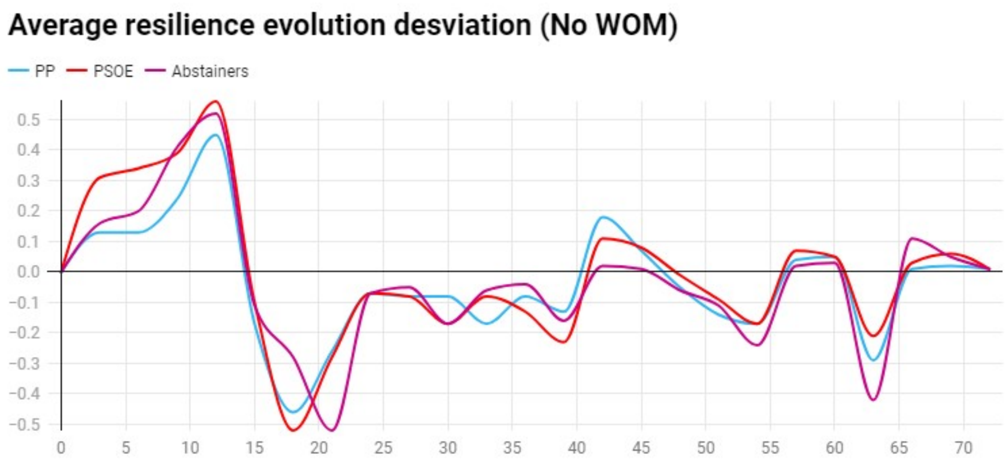

Figure 4.7: Average resilience evolution deviation (Model 2)

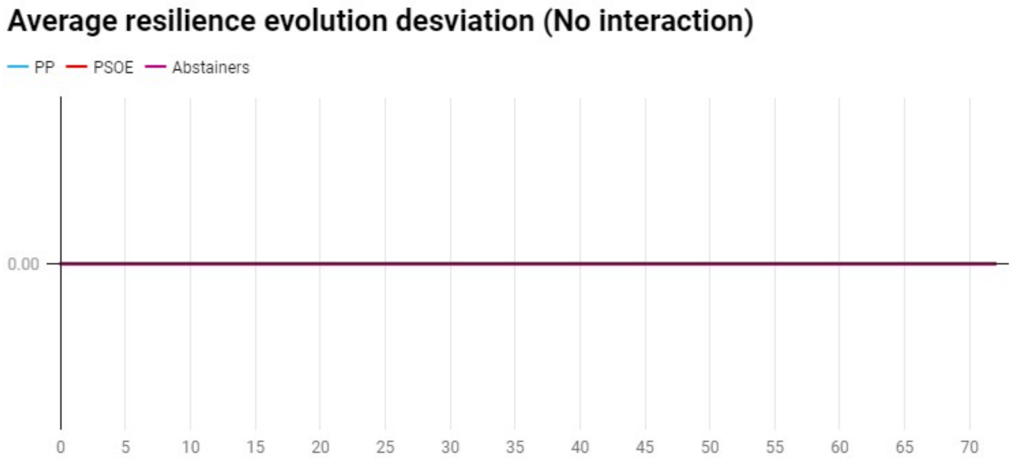

Figure 4.8: Average resilience evolution deviation (Model 1)

Furthermore, it is known from Sect. 3.4 that voters who belong to the PP segment have a higher influence decay than the others. We can deduce then, even without knowing the messages that were transmitted, that in the steps where we can find a bigger alteration of resilience or high resilience deviation values, polarized and high-audience messages were emitted. This effect is reflected in Table 4.2 where positive messages account for ETA pointing and negative ones for Al-Qaeda pointing. Broadly speaking, during the first 15 hours of the simulation, the mass media channels are more angled to transmit messages about the participation of the terrorist group ETA in the attacks. There was also a try of some of these channels to obscure some of the evidence that placed Al-Qaeda as the perpetrators, evidence that was brought to light on 12th March, the day where the first messages blaming Al-Qaeda appeared.

	STEPS
Medium impact positive messages	0-8, 65-68
High impact positive messages	8-12
Medium impact negative messages	37-38, 52-54
High impact negative messages	62-64

Table 4.2: Steps in which high-audience polarized messages were emitted

Besides, a uniform tendency to decrease the resilience value in non-forwarding hours can be observed. This is a direct consequence of WOM interactions between neighbors, where agents belonging to the PSOE segment are more likely to talk and influence the resilience of their neighbors.

4.2 Sensitivity analysis

The purpose of this section is to explore in more depth different scenarios in which some variables take different values from their expected values.

4.2.1 WOM influence and network generation

The first sensibility analysis was performed using the No-media model with the purpose of strengthening Sect. 3.4 talking about the neighbor talking probability and the network degree variables. In order to achieve this objective, that model's ABSS has been executed multiple times changing the parameters already mentioned and only recording interactions between agents of the same segment. These changes are made by increasing low initial values with the purpose of progressively accelerating the diffusion of information process.

All results are simplified in Table 4.3 after all combinations of these two variables were tested. For every segment, the first column represents the numerical values used for the variables. The next two columns illustrate if at least half of the agents have established a conversation and, if they had, in which step this event has occurred.

When we discuss again this project model, our goals for the WOM interaction topic are.

- Most agents involved in a political segment need to interact with other agents involved in the same or opposite segment.
- Abstainers do not have the imperative need to spread their thoughts because they currently do not belong to any political segment.
- To make a similar approach to reality, highly connected agents should not establish a conversation with all its neighbors in a short period of time.

Analyzing the results, the optimal values for voters inside political segments are those in which at least half of the population has established a conversation but these conversations are taken place as slowly as possible. Talking about abstainers, we are not interested in them establishing a lot of conversations, so we can take the minimum value possible for the model to work correctly.

Therefore, we can confirm that the best combination of parameters is the first use of m = 2 as the degree of the network because the voters of PP and PSOE establish the minimum number of conversations required and the abstainers reached their goals and the second application of $\rho(t)$ values of 0.02 (for the PP voters), 0.04 (for the PSOE voters) and 0.01 (for the abstainers).

Going further, we can evaluate the behavior of these variations to obtain more conclusions. It is easy to see that when increasing the values of both diffusion parameters, the three types of agents tend to build new conversations more quickly. However, these conversations lean to achieve some saturation value as the parameters are amplified more and more because, in general, agents cannot transmit unlimited information across so many different network links.

		Value	Half talked	Step			Value	Half talked	Step			Value	Half talked	Step
PP segment	Talking probability	0.01	False	-	PSOE segment	Talking probability	0.01	False	-	Abstainers	Talking probability	0.01	False	-
		0.02	True	38			0.02	False	-			0.02	False	-
		0.03	True	26			0.03	False	-			0.03	False	-
		0.04	True	20			0.04	True	47			0.04	False	-
		0.05	True	17			0.05	True	38			0.05	False	-
		0.06	True	15			0.06	True	35			0.06	True	50
	Network degree	1	False	-		Network degree	1	False	-		Network degree	1	False	-
		2	True	38			2	True	47			2	False	-
		3	True	16			3	True	22			3	True	59
		4	True	8			4	True	11			4	True	28
		5	True	6			5	True	8			5	True	14
		6	True	5			6	True	6			6	True	10

Table 4.3: WOM sensitivity analysis results

4.2.2 Influence of mass media

In this case, we vary the polarization of the messages for each type of mass media and look for changes in the political segments of the PP and PSOE of voters separately. The precondition of this sensitivity analysis is that for the results of all agents obtained in the complete model, we take the reference of "0" polarization. Thus, we can see what would have happened, since their overall polarization of their messages is near zero (most of the messages have null polarization), if both political segments had opted for a more aggressive campaign.

In our first Figure 4.9, we can see how television polarization finds the highest changes, followed by radio and the press. The function does not follow a linear behavior because, around the mean value ±0.5, all types of media saturate. It is also important to emphasize that below zero values, voters decrease faster than they grow over zero values. This illustrates how PP voters are more influenced by Al-Qaeda messages, having a natural tendency to slightly decrease their resilience value.

In the second Figure 4.10, we can observe how television polarization yet finds the highest changes beating radio and press; but this time, radio achieves a greater impact than in the PP segment, because Cadena SER has a stronger audience between PSOE voters. As before, all types of media saturate around the mean value ±0.5; but there is an important difference between these two figures, that is, the way PSOE votes decrease when influenced by ETA messages. In this case, they decrease more than PP votes did with Al-Qaeda messages and that happens because agents from PSOE segment have a lower influence decay value, which leads to suffer a strongest influence from the environment.

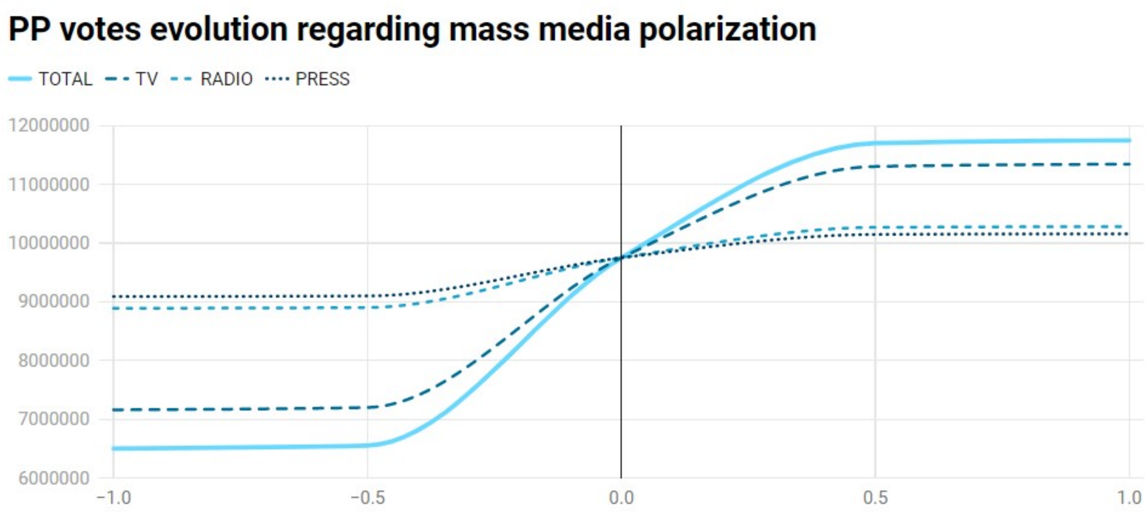

Figure 4.9: PP votes as a result of variations in messages' polarization

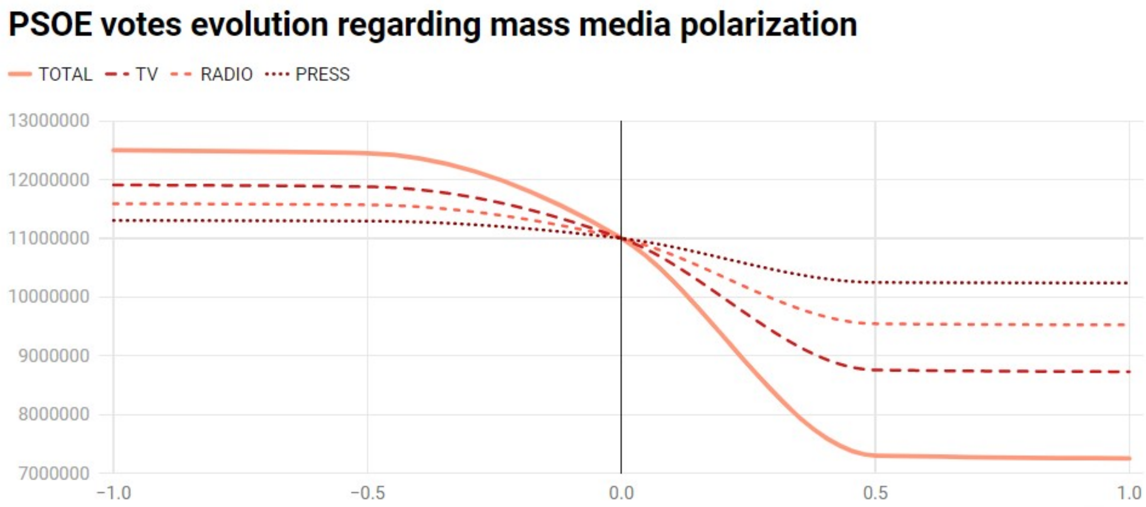

Figure 4.10: PSOE votes as a result of variations in messages' polarization

4.3 Framing effects

We have previously talked about framing effects in Sect. 3.2 and how they are taken into account in the development of the simulation. Figure 4.11 represents exactly the average values that agents take just after the complete model simulation ends to show whether they stayed in their first segment or moved to another.

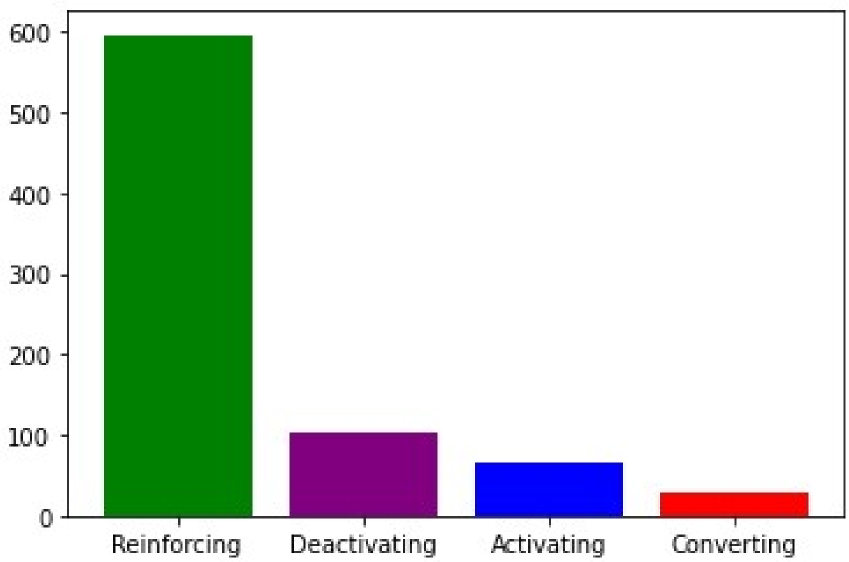

Figure 4.11: Average framing effect results

CHAPTER 5

Conclusions and future work

In this chapter we will describe the conclusions extracted from this project in Sect. 5.1, achievements of the project in Sect. 5.2 and lastly, the thoughts about future work in Sect. 5.3.

5.1 Conclusions

To conclude this project, first, we are going to recapitulate and summarize all the concepts previously described in the rest of chapters. 11^{th} March 2004 terrorist attacks marked a before and after in the 14^{th} March Spanish general elections result. How this incident was treated and transmitted by the most relevant mass media channels and how all citizens and politicians interacted between them influenced in the population by changing their political thoughts with the judgement of two different terrorist groups as perpetrators of the attacks.

With the goal of simulating the influence mentioned above, we have developed a complete ABSS with an absolute deviation of just 0.05% with respect to real election results to mimic the behavior of the population and extract solid conclusions about some of the reasons why Partido Socialista Obrero Español (PSOE) won an absolute majority in the Spanish Congress when days before the disaster, the most relevant surveys tagged Partido

Popular (PP) as favorites to win the elections.

The extracted results from the simulations suggest, talking about the framing effects in Sect. 4.3, a strong movement from PP voters towards political deactivation and an also important mobilization of abstainers to the PSOE segment. Other conclusions obtained from the sensitivity analysis indicate the importance of mass media in the impact caused by diffusion mechanisms, followed by the influence of WOM processes where PSOE voters exercise more political leverage than the rest of the population.

Moreover, Sect. 4.2.1 illustrates the importance of choosing network parameters because these directly alter the exposing of an artificial population to polarized information. Lastly, Sect. 4.2.2 shows how the aggregation of all mass media channels when they transmit the same type of messages have a stronger impact in the population than each media type by its own, despite television was clearly the most important of them in the early 2000s.

To sum up, in a well connected world like the ones we can artificially generate with a computer and the one that we live in, a well organized emission of information can change the course of history in barely three days.

5.2 Achieved goals

Network generation and study of artificial social networks

Before the development of any design, it has been required to collect information about artificial networks and how to approach these networks to a social scenario. It was also needed to compare different network generator algorithms to find the one that fits better with our model and see how it behaves during the simulation.

Design and implementation of a 11M-14M model

For this area, the necessary variables and the relations between them were studied to develop a substantial model. Then we used that information to write and program all the agents, the environment, and the simulation configuration file. Finally, we run the simulation several times for each phase of the model, storing the results in CSV files.

Extraction and filtering of data for visualization

To continue the work on the subject, it is essential to understand all the values and results obtained in order to manipulate them toward the selection of relevant information. This includes the control and filtering of DataFrames in Jupyter and the extraction of those segmented DataFrames for their subsequent plotting.

Plotting the results and compare them to the ones in the article

Finally, use the necessary tools to plot the desired information in a simple way. Once all figures are defined, compare them to those illustrated in [MCSLC17] to find their similarities and differences; and draw the appropriate conclusions about them.

5.3 Future work

Improve WOM interaction between agents Modeling agent's behavior, there is one social aspect that has not been taken into account when coding neighbor's interactions. We could model a more attractive simulation where talking probability values change depending on the current hour of the day simulation is running in order to avoid a potential same number of interactions at 2 AM and at 7 PM for example.

Study alternative strategies for modeling

In this project, we have developed an agent-based model using Soil by defining Python classes and functions and structuring a configuration YAML file. It would be interesting to study other ways and programs to define these types of models to maybe find others whose characteristics fit better in the generation of social simulations.

Political behavior among agents

We already saw in Sect. 2.2.1 that agents are very versatile entities with the opportunity to be modeled as the system developer desires. In our ABSS, mainly only one variable of each agent controls its political behavior. We could set inside future goals a deeper search for political influence in specific populations and how to transfer it into several modifiable parameters within an artificial network.

Go deeper into network generation

Finally, the possibility of deepening the generation of artificial social networks has been considered. This topic is one of the most interesting pillars of the project because all interactions between neighboring agents depend on how they are connected. We could explore the possibility of changing network generation algorithms to compare them and even visualize each type for larger populations.

APPENDIX A

Impact of this project

This appendix reflects quantitatively or qualitatively on the possible impacts, whether positive or negative, of this project. These impacts are classified into social impact (Section A.1), economic impact (Section A.2), environment impact (Section A.3) and ethical implications (Section A.4).

A.1 Social impact

In recent years, political terrorism has been a relevant topic in many articles and research works. Most of these projects are related to the collection and contrast of real data for a later analysis of that information with public or private purposes. Moreover, many recent studies review the influence of mass media, or other types of mass communication channels such as social networks, on different political and social behaviors. If we combine these two topics with ABSS techniques, we obtain the essence of this project.

11^{th} March terrorist attacks marked a before and after all around the world, together with other well-known terrorist attacks at the beginning of the 21^{st} century. This caused worldwide concern about the ability of governments to guarantee the security of citizens, especially in critical infrastructures such as public transport. Because these incidents have such an impact on society that people want answers even before researchers start to find evidence or perpetrators, which irretrievably leads to the transmission of rumors or fake news. This transmission effect arises when an important political event like general elections is near to be held, giving us the idea that the messages emitted about the attacks may hide political propaganda towards certain segments of the population.

In addition, other related experiments [MS06] suggest mass media polarized messages to exert influence through direct opinion exchanges with a certain agent or even indirectly through the interaction with its neighbors. This project can be used for social purposes in terms of trying to avoid the influence of the media in a context with the absence of proven facts.

A.2 Economic impact

On the one hand, 11^{th} March terrorist attacks caused extreme economic damage in Spain, affecting a few hundredths of the total GDP of the nation [dE$^+$04]. More than 200 million euros as a sum of different costs such as infrastructure and train repairs, compensations, and serious drops in consumption inside different sectors, highlighting the touristic and railway ones.

On the other hand, political propaganda has been adapting to current technologies, preparing each year more sophisticated and more expensive ways of transmitting messages through the media or the Internet. In 2004 the most influential type of mass media was television as we saw in Sect. 5.1. Simulations like the one we developed in this project can estimate a minimum budget to influence a complete population by knowing the most influential channels and types of information transmission.

A.3 Environment impact

The environmental impact of this project involves three main footprints. The first one covers all the environmental impact of terrorist attacks which, as new technologies evolve, it also evolves to the use of more dangerous weapons with the ability to overwhelm whole countries. The second and third are directly related to the consumption of non-renewable energy for both transmission of information and the expensive development and implementation of Artificial Intelligence algorithms and models such as this project Agent-Based Social Simulation.

A.4 Ethical implications

The most important ethical implication that we can find in this project is the research and collection of personal data to model social and political behaviors. This kind of information must be obtained and used ethically, where sensitive data must remain completely anonymous and should not be used for other purposes than academic or research.

APPENDIX B

Economic budget

This appendix details an adequate budget to bring about the project developed satisfying all the economic costs. These expenses are segregated into physical resources (Section B.1), human resources (Section B.2), licenses (Section B.3) and taxes (Section B.4).

B.1 Physical resources

This project offers multiple models that can be simulated and visualized with just one computer, be it a tower or a laptop. Its specifications can vary according to how fast we want to run the simulation, but the minimum requirements to launch it are the following.

- **RAM:** 8 GB
- **Hard disk:** 10 GB SSD free space
- **CPU:** 4 to 8 CPU cores at 3.00+GHz

B.2 Human resources

In this section, we quantify human resources considering the total number of hours employed developing the software, as well as the average salary of an engineer.

Considering that a month has around twenty-two working days and four hours per working day (part-time schedule), the time cost of this project is estimated to be around 484 hours (five and a half months). The average salary provided to an engineer developing this kind of software is expected to achieve a total of 450 € per month. Thus, the total cost dedicated to the development of the software results in an approximate price of 2,475 €.

B.3 Licenses

This section discusses the costs of the software used in this project. As all the software that has been used is open source, the costs of all the licenses increase to 0 €.

B.4 Taxes

This last section manages the possible activity of selling the software of this project, once it has been implemented, to another company. If that occurred, taxes would derive from the sale of the previous mentioned software.

According to software taxation [dlTD10], it must be considered a tax of 15% of the total price of the software, supported and regulated by the Statue 4/2008 of Spanish law. Only in the case that a foreign company was interested in the sale would this scenario be analyzed.

Bibliography

[ABBSGG15] Frédéric Amblard, Audren Bouadjio-Boulic, Carlos Sureda Gutiérrez, and Benoit Gaudou. Which models are used in social simulation to generate social networks? a review of 17 years of publications in jasss. In *2015 Winter Simulation Conference (WSC)*, pages 4021–4032, 2015.

[BA99] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

[Ber14] Jonah Berger. Word of mouth and interpersonal communication: A review and directions for future research. *Journal of consumer psychology*, 24(4):586–607, 2014.

[BKEI09] Oren Ben-Kiki, Clark Evans, and Brian Ingerson. Yaml ain't markup language (yaml™) version 1.1. *Working Draft 2008*, 5:11, 2009.

[BOL20] Gerry Brierley, Wilson Ozuem, and Geoff Lancaster. Subconscious marketing communication techniques and legal implications. *Journal of Decision Systems*, 29(2):69–78, 2020.

[CIS04] CIS. Preelectoral elecciones generales 2004. `https://www.cis.es/cis/opencm/EN/2_bancodatos/estudios/listaMuestras.jsp?orden=1&desc=DESC&estudio=3714&cuestionario=3885&pagina=1`, 2004. Accessed 10/05/2022.

[CPDRBG14] Ivet Challenger-Pérez, Yanet Díaz-Ricardo, and Roberto Antonio Becerra-García. El lenguaje de programación python. *Ciencias Holguín*, 20(2):1–13, 2014.

[Dat22] Datawrapper. Datawrapper. enrich your stories with charts, maps, and tables. `https://www.datawrapper.de/`, 2022. Accessed 13/03/2022.

[dE$^+$04] Banco de España et al. Los efectos de los atentados del 11 de marzo de 2004, según los indicadores coyunturales. *Boletín Económico*, (MAY):1–8, 2004.

[dlTD10] Francisco de la Torre Días. La tributación del software en el irnr. algunos aspecttos conflictivos. `https://www.ief.es/docs/destacados/publicaciones/revistas/cf/10_14.pdf`, 2010. Accessed 20/06/2022.

[Esp04] Constitución Española. Elecciones generales de 14 de marzo de 2004. distribución de votos y escaños en el congreso. `https://app.congreso.es/consti/elecciones/generales/resultados.jsp?fecha=14/03/2004`, 2004. Accessed 10/05/2022.

[HSDG+11] Helen Herrman, Donna E Stewart, Natalia Diaz-Granados, Elena L Berger, Beth Jackson, and Tracy Yuen. What is resilience? *The Canadian Journal of Psychiatry*, 56(5):258–265, 2011.

[HSSC08] Aric Hagberg, Pieter Swart, and Daniel S Chult. Exploring network structure, dynamics, and function using networkx. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008.

[Jup22a] Jupyter. Architecture. `https://docs.jupyter.org/en/latest/projects/architecture/content-architecture.html`, 2022. Accessed 02/05/2022.

[Jup22b] Jupyter. Jupyter project documentation. `https://docs.jupyter.org/en/latest/`, 2022. Accessed 02/05/2022.

[M+11] Wes McKinney et al. pandas: a foundational python library for data analysis and statistics. *Python for high performance and scientific computing*, 14(9):1–9, 2011.

[MCSLC17] Ignacio Moya, Manuel Chica, Jose L Saez-Lozano, and Oscar Cordon. An agent-based model for understanding the influence of the 11-m terrorist attacks on the 2004 spanish elections. *Knowledge-Based Systems*, 123:200–216, 2017.

[MN09] Charles M Macal and Michael J North. Agent-based modeling and simulation. In *Proceedings of the 2009 winter simulation conference (WSC)*, pages 86–98. IEEE, 2009.

[MS06] Gary Mckeown and Noel Sheehy. Mass media and polarisation processes in the bounded confidence model of opinion dynamics. *Journal of Artificial Societies and Social Simulation*, 9(1), 2006.

[MtPDT22] Wes McKinney and the Pandas Development Team. pandas: powerful python data analysis toolkit release 1.4.2. `https://pandas.pydata.org/docs/pandas.pdf`, 2022. Accessed 02/05/2022.

[Nac19] Audiencia Nacional. Hechos probados. `https://ayuda11m.org/wp-content/uploads/2019/07/hechos-probados.pdf`, 2019. Accessed 26/04/2022.

[NG19] Abhinav Nagpal and Goldie Gabrani. Python for data analytics, scientific and technical applications. In *2019 Amity international conference on artificial intelligence (AICAI)*, pages 140–145. IEEE, 2019.

[plIdMdC04] Asociación para la Investigación de Medios de Comunicación. Resumen general de la aimc desde febrero hasta noviembre de 2004. `https://www.aimc.es/a1mc-c0nt3nt/uploads/2010/09/resumegm304.pdf`, 2004. Accessed 26/04/2022.

[plIdMdC10] Asociación para la Investigación de Medios de Comunicación. Marco general de los medios en españa 2005. `https://www.aimc.es/a1mc-c0nt3nt/uploads/2010/11/marco05.pdf`, 2010. Accessed 26/04/2022.

[RG21] Jeffrey Ramos González. Estos son los 10 lenguajes de programación más populares en 2021. `https://www.enter.co/especiales/dev/herramientas-dev/estos-son-los-10-lenguajes-de-programacion-mas-populares-en-2021/`, November 2021. Accessed 26/04/2022.

[SISR17] Jesús M. Sánchez, Carlos A. Iglesias, and J. Fernando Sánchez-Rada. Soil: An Agent-Based Social Simulator in Python for Modelling and Simulation of Social Networks. In Demazeau Y., Davidsson P., Bajo J., and Vale Z., editors, *Advances in Practical Applications of Cyber-Physical Multi-Agent Systems: The PAAMS Collection*, volume 10349 of *LNAI*, pages 234–245. Springer Verlag, June 2017.

[Wei66] Walter Weiss. *Effects of the mass media of communication.* ERIC, 1966.

[WGFM20] Meike Will, Jürgen Groeneveld, Karin Frank, and Birgit Müller. Combining social network analysis and agent-based modelling to explore dynamics of human interaction: A review. *Socio-Environmental Systems Modelling*, 2:16325–16325, 2020.