Facultad de
Ciencias Sociales y
Tecnologías de la Información
Talavera de la Reina. UCLM

UNIVERSIDAD DE CASTILLA-LA MANCHA

TRABAJO FIN DE GRADO

BACHELOR IN COMPUTER ENGINEERING

Information sources reliability. Identifying disinformation sources through the complex networks paradigm.

Lucian Andrei Farcas

December, 2024

Facultad de
Ciencias Sociales y
Tecnologías de la Información
Talavera de la Reina. UCLM

UNIVERSIDAD DE CASTILLA-LA MANCHA

TRABAJO FIN DE GRADO

Department of Technologies and Information Systems

Specialization Information Systems

Information sources reliability. Identifying disinformation sources through the complex networks paradigm.

Author: Lucian Andrei Farcas

Supervisor: Alfonso Niño Ramos

December , 2024

*T*o my parents and
all the people that supported me.
Thank you

Statement of Authorship

I, Lucian Andrei Farcas with NIE X9862857C, declare that I am the sole author of the final degree project entitled “Information sources reliability. Identifying disinformation sources through the complex networks paradigm” and that the aforementioned work does not infringe the laws in force on intellectual property and that all non-original material contained in said work is appropriately attributed to its legitimate authors.

Talavera de la Reina, at …..

Signed:

Resumen

Estamos viviendo en una era dominada por la información, donde la política, el trabajo diario, las redes sociales y los pasatiempos están moldeados por el constante bombardeo de datos y narrativas. Sin embargo, la fiabilidad de la información se ha convertido en uno de los mayores desafíos de nuestro tiempo. La prevalencia de la desinformación y la información errónea ha socavado la confianza en los recursos en línea, dificultando cada vez más la distinción entre la verdad y la falsedad. Abordar este problema requiere herramientas y metodologías innovadoras.

Las redes complejas proporcionan un marco poderoso para comprender las dinámicas estructurales y relacionales de los ecosistemas de información. Al modelar cómo las narrativas se propagan e interconectan, las redes complejas permiten descubrir patrones ocultos, identificar fuentes influyentes y detectar clústeres de contenido relacionado. Este trabajo propone un enfoque para aprovechar estas herramientas: utilizar modelado de temas para extraer temas latentes del conjunto de datos de PolitiFact, construir una red basada en la similitud de coseno, simplificar la red a través de un Árbol de Recubrimiento Mínimo (MST, por sus siglas en inglés) y aplicar detección de comunidades para revelar ideas sobre la estructura y la propagación de la desinformación.

Esta metodología ofrece una perspectiva novedosa al integrar análisis semántico y estructural para abordar el problema crítico de la información poco confiable. No solo identifica temas prominentes de desinformación, sino que también resalta sus interrelaciones y mecanismos de propagación, allanando el camino para estrategias más efectivas en la lucha contra la difusión de narrativas falsas.

Abstract

We are living in an era dominated by information, where politics, daily work, social media, and hobbies are all shaped by the constant bombardment of data and narratives. However, the reliability of information has become one of the greatest challenges of our time. The prevalence of disinformation and misinformation has undermined trust in online resources, making it increasingly difficult to discern truth from falsehood. Addressing this issue requires innovative tools and methodologies.

Complex networks provide a powerful framework for understanding the structural and relational dynamics of information ecosystems. By modeling how narratives propagate and interconnect, complex network allows us to uncover hidden patterns, identify influential sources, and detect clusters of related content. This paper proposes an approach to leverage these tools: using topic modeling to extract latent themes from the PolitiFact dataset, constructing a network based on cosine similarity, simplifying the network through a Minimum Spanning Tree (MST), and applying community detection to reveal insights into the structure and spread of disinformation.

This methodology offers a fresh perspective by integrating semantic and structural analysis to tackle the critical problem of unreliable information. It not only identifies prominent disinformation themes but also highlights their interrelations and propagation mechanisms, paving the way for more effective strategies to combat the spread of false narratives.

Acknowledgements

I would like to express my deepest gratitude to all those who have supported and guided me throughout the journey of completing this project.

I am profoundly grateful to my family, whose love, encouragement and support have been my foundation throughout this progress.

I extend my heartfelt thanks to my tutor, Alfonso Niño Ramos, for his invaluable guidance, patience, and insightful advice.

Thank you all for being part of this journey.

Contents

List of Figures

List of Tables

Listings

Acronyms

NLP Natural language processing

ML Machine learning

JNI Java native interface

P Probability

LDA Latent dirichlet allocation

Chapter 1

Introduction

In today's fast-paced digital era, the way we consume and interact with information has transformed dramatically. The internet and social media platforms have become the primary sources for news, opinions, and narratives, shaping public discourse and influencing decision-making on both individual and societal levels. However, alongside these opportunities comes a pressing challenge: the proliferation of misinformation and disinformation, which erodes trust in information systems and has far-reaching consequences for politics, science, and everyday life.

This project aims to address this critical issue by exploring innovative methods to analyze and combat false information. By leveraging advanced analytical tools such as topic modeling and complex network analysis, this work proposes a systematic approach to uncovering patterns and insights within disinformation networks. The ultimate goal is to provide practical methodologies and actionable insights that can empower individuals, organizations, and policymakers to better navigate the complex and often misleading information landscape.

1.1 Motivation

In a world where the internet and social media dominate nearly every aspect of our lives, filtering the information we consume has become a fundamental task to maintain not only our sanity but also an informed perspective. The digital age has democratized access to information, making it easier than ever to share ideas, opinions, and news. However, this unprecedented flow of information comes at a significant cost: the overwhelming challenge of distinguishing truth from falsehood in a landscape saturated with biases, sensationalism, and outright fabrication.

This task is especially daunting in the digital space, where content creators, influencers, and even traditional news outlets are often incentivized to prioritize engagement over accuracy. Controversial and divisive content attracts more attention and drives higher traffic than balanced and neutral reporting. As a result, narratives tend to polarize, reinforcing the false dichotomy of "you are either with me or against me". This dynamic enhances the difficulty of finding common ground, leaving those who seek objectivity vulnerable to criticism from all sides. The rise of echo chambers and algorithmically driven content feeds further amplifies this issue by isolating users within their preferred biases, creating an environment where misinformation and disinformation thrive.

Complicating matters further is the reality that we are living in an era of digital and information warfare. Governments, corporations, and even individuals weaponize information to influence opinions, manipulate behaviors and destabilize societies. In this context, skepticism has become a necessary survival skill. The familiar phrase “take it with a grain of salt” has evolved into a mandatory mindset when consuming online content. Yet, the sheer volume of information we encounter daily can be paralyzing, leading many to disengage entirely or succumb to cynicism.

Despite these challenges, interaction with the digital space is inevitable. Social media platforms, news sites and other online resources have become integral to modern life, connecting us to essential information, communities and opportunities. Given this dependence, the solution cannot be to abandon the digital world but rather to develop robust processes and systems to help users evaluate the reliability of the information they consume. Tools and methodologies for identifying and combating false information are not just useful, they are essential for preserving the integrity of public discourse and empowering individuals to make informed decisions.

This project seeks to contribute to the fight against misinformation and disinformation by offering both a framework and practical insights into identifying false information. Specifically, it aims to:

- **Build a methodology for identifying disinformation**. By analyzing the characteristics and patterns of fake news, the project seeks to create a systematic approach that can be applied to new or existing sources of disinformation. This methodology leverages advanced analytical techniques, such as topic modeling and network analysis, to go beyond surface-level detection.
- **Identify patterns and similarities**. By examining connections and commonalities between disinformation and truthful statements, the project aims to uncover deeper insights into how false narratives are constructed and spread. These patterns can inform strategies for combating the dissemination of disinformation and fostering trust in credible information sources.

Ultimately, this work aspires to provide actionable solutions for navigating the complex and often hostile information landscape. By understanding the mechanisms that drive misinformation and disinformation, we can move closer to a digital space where truth and reliability are not casualties of the battle for attention. This effort is a small but crucial step toward empowering individuals and communities to reclaim agency in their relationship with information.

1.2 Document Structure

This document is structured into several key sections exposing the process and results of this project. Below is an overview of each section:

1. **Introduction**. The introduction sets the stage by discussing the motivation behind the study.
2. **Objectives**. This section outlines the specific goals of the research.
3. **State of the art**. A comprehensive review of existing literature is provided in this section. It covers key concepts of fake news, explores available datasets, discusses topic modeling, introduces the fundamentals of complex networks and examines how these tools have been used in studying disinformation.
4. **Tools and methodology**. This section describes the tools and methodologies used in this research.
5. **Results**. The results section presents the key findings and outcomes of the research.
6. **Conclusions**. This section evaluates the research outcomes, discussing the extent to which the objectives were achieved.

Chapter 2

Objectives

The general objective of this project is to design a methodology for identifying the disinformation potential of new or pre-existing sources on the web. The achievement of this objective will be possible by completing the following partial objectives:

- From sources of information of a known nature on “Politifacts.com” data, develop the necessary methodology to quantify the qualitative characteristics recorded of the statements.
- Determine the minimum set of representative characteristics for assessing similarities between statements using, and if necessary developing, dimensional reduction techniques.
- Simplify the initial similarity network by reducing complete networks to trees coating.
- Automate the analysis of communities in statement similarity networks considering the role of the resolution parameter in multi-resolution methods for determining communities.

Chapter 3
State of the art

In this chapter some key aspects for this paper that will be of relevance later on will be exposed, some of this aspects include the main focus of this work: disinformation and complex networks; the main tool used to make this possible: topic modelling; and how all of this concepts relate to each other including other research on this field.

3.1 Information, misinformation and disinformation

Information has always been a powerful tool, influencing various aspects of human life, from sharing a simple recipe to devising complex military strategies, it is present in every aspect of our lives and has an extreme influence in it, shaping our understanding, decisions and actions. How many historic-changing events could have been prevented or changed if the right individuals knew the right information at the right moment? Wars, financial crises, terrorist attacks or revolutions are often the result of what the individuals surrounding them did and didn't knew at crucial moments.

The role of information in society has transformed dramatically over the past decades, fueled by advancements in technology, the proliferation of the internet, and the ubiquitous presence of social media. Its influence has grown exponentially in recent years due to the sheer volume, speed, and accessibility of data in our digital age.

A decade ago, access to information was more constrained by the limitations of traditional media. People relied heavily on printed newspapers, television, and radio broadcasts as their primary sources of news and updates. These mediums inherently restricted the flow of information to scheduled intervals and limited formats. According to a study conducted by the International Telecommunications Union (ITU), global internet usage in 2012 was approximately 2.4 billion users, a big contrast to the 5.4 billion users recorded in 2023 [7] [12] . This increase in connectivity has fundamentally altered the dynamics of information dissemination and consumption.

In the current era, individuals are exposed to unprecedented amounts of information daily. A 2017 article from the Frontiers revealed that the average individual now consumes around

74 GB of data per day through digital platforms [42], a figure that has only increased with the COVID-19 pandemic an the more recent AI.

Another important factor is the amount of data being generated now days. According to a Statista report [2] [1] in 2023 global data generation stood at 0.337 zettabytes per day (337.000.000 terabytes) and its still rising day by day.

In today's world, where vast amounts of information are generated and consumed every day, ensuring its accuracy and safeguarding it from malicious intent is a complex and crucial task. This challenge highlights the importance of distinguishing between different types of information. These distinctions are essential for understanding the nature of the information we encounter and its potential impact:

- **Disinformation**: refers to "false information deliberately and often covertly spread (as by the planting of rumors) in order to influence public opinion or obscure the truth" [3]. The intent to harm or manipulate distinguishes disinformation from other forms of false information .
- **Misinformation**: means "incorrect or misleading information" [4]. It is the result of errors in comprehension, communication or misinterpretation and lacks the purposeful drive to deceive .
- **Malinformation**: on the other hand, is true information meant to cause harm [38].

With this, researches classify information based on intent, content or both, although for this classification, everything is found under the umbrella of "*fake news*", which is a much more known name for disinformation, misinformation and all types of false information, but one term with no agreed definition in the literature [17]. In this paper we will use it to encapsulate all types of false information including all the types listed previously.

Content-based fake news are classified using false text, hyperlinks, embedded content, images, false videos, audios, etc. In general, they are classified by the multimedia they posses.

Intent-based fake news are more discussed and include clickbait, hoax, rumor, satire, propaganda, framing, conspiracy theories and others:

- Clickbait. Refers to misleading headlines and thumbnails of content on the web that tend to be fake stories with catchy headlines.
- Hoax. False or inaccurate stories used to masquerade the truth and is represented as factual.
- Rumor. Ambiguous claims that are disseminated with a lack of evidence to support them.
- Satire. Stories that contain a lot of irony and humor.
- Propaganda. News stories created by political entities to mislead people.
- Framing. Refers to employing some aspect of reality concealing the truth and misguiding people, often out of context.
- Conspiracy theories. The belief that an event is the result of secret plots generated by

powerful conspirators.

Although it is of interest to differentiate all of these types, it is important to highlight that there exist an overlap in the types listed before and thus is possible to have false information falling within multiple categories.

3.1.1 Risk and interest in fake news

Fake news and all of its categories poses profound risks across societal, political, and economic domains. These risks, amplified by the rapid evolution of communication technologies, make it a pressing challenge for modern societies:

1. **Erosion of public trust**. One of the most significant risks of disinformation is its ability to undermine trust in key institutions such as governments, media, and scientific organizations. When disinformation campaigns spread false narratives about elections, public health, or climate change, they create doubt in the legitimacy and credibility of these institutions. For instance, during the 2020 U.S. presidential election, coordinated disinformation campaigns targeted the integrity of voting systems, leading to widespread mistrust and civil unrest [33]. Such erosion of trust destabilizes democratic processes and makes societies more vulnerable to further manipulation.
2. **Threats to public health**. The global COVID-19 pandemic highlighted the dangerous intersection of disinformation and public health. False claims about the origins of the virus, vaccine efficacy, and alternative cures led to widespread confusion and vaccine hesitancy, prolonging the pandemic and increasing mortality rates [30]. For instance, rumors that the virus was a hoax or that certain unproven remedies like drinking bleach could cure it resulted in preventable deaths and widespread panic. Public health authorities struggled to counter the infodemic [9] that spread faster than official health advisories.
3. **Economic disruption**. Disinformation can disrupt markets and undermine economic stability. False reports about companies, stock values, or economic policies can cause fluctuations in financial markets, leading to losses for investors and businesses. For example, during the GameStop stock surge in early 2021, false narratives about market manipulation circulated widely, contributing to panic and confusion among retail investors [36]. Furthermore, businesses targeted by disinformation campaigns may suffer reputational damage, which can affect consumer trust and sales.
4. **National security threats**. Disinformation is increasingly being used as a tool in cyber warfare. Foreign actors deploy coordinated campaigns to destabilize nations, influence elections, and undermine public confidence in governments. For instance, Russian interference in the 2016 U.S. presidential election involved the dissemination of false information to manipulate voter behavior and sow discord among the electorate [20]. These activities represent a growing national security threat, as they weaken the social fabric and leave nations vulnerable to external manipulation.

5. **Psychological and cognitive impacts**. Persistent exposure to disinformation affects individuals' mental health and cognitive processing. Fake news, particularly during crises, can lead to heightened anxiety, fear, and confusion. Moreover, individuals who consume disinformation repeatedly may develop a distorted perception of reality, becoming more susceptible to conspiracy theories and less capable of critical thinking [25]. This can erode rational decision-making at both individual and societal levels.

All of this risks can be clearly identified in Thomas Rid's book "Active Measures: The Secret History of Disinformation and Political Warfare" [41] where he goes through some of the most relevant events in history and how many of them where surrounded by fake news, or as Thomas describes them: *active measures*.

But not everything is bad news, since increasing prevalence and impact of fake news have sparked significant interest among academics, policymakers, and international organizations. In this regard it is possible to differentiate multiple advances to combat fake news:

- **Academic research**. Fake news have become a focal point in disciplines such as sociology, psychology, and computer science. Researchers are examining the psychological mechanisms that make individuals susceptible to disinformation, including cognitive biases, confirmation bias, and emotional triggers [22]. Studies also explore how social media algorithms amplify disinformation and how misinformation spreads through network structures, using models from network science and data analytics [32].
- **Policy and regulation**. Policymakers worldwide are grappling with the challenge of regulating disinformation without infringing on free speech. Initiatives like the European Union's Code of Practice on Disinformation aim to hold social media platforms accountable for the spread of false information [13].
- **Technological interventions**. The development of artificial intelligence (AI) and machine learning tools has garnered interest for their potential to combat disinformation. These technologies analyze patterns in content dissemination, detect fake news, and identify bot-driven campaigns. However, their effectiveness is limited by the sophistication of disinformation tactics, including the use of AI-generated deepfakes [23].
- **Media literacy campaigns**. Public education initiatives are gaining momentum as an essential tool to counter disinformation. Media literacy programs aim to teach individuals how to critically evaluate information sources, recognize manipulation techniques, and verify facts. These campaigns are particularly effective in schools, where they can equip future generations with the skills needed to navigate the complex digital landscape [10].
- **Global collaborations**. The international nature of disinformation requires a coordinated global response. Organizations like the World Health Organization [5] and the United Nations nations_united_nodate are working to address disinformation during crises such as pandemics and armed conflicts. Collaborative frameworks like the Global Internet Forum to Counter Terrorism (GIFCT) bring together governments, tech companies, and civil society to combat the spread of harmful content [8].

- **Corporate responsibility**. Social media companies have come under increased scrutiny for their role in enabling the spread of disinformation. Platforms like Facebook, Twitter, and YouTube are investing in fact-checking partnerships, content moderation algorithms, and user awareness campaigns. However, critics argue that these efforts are often insufficient and that greater accountability is needed [43].

3.1.2 Disinformation data sets

One of the biggest bottle-necks when researching fake news is data. It may be contradictory after knowing the quantity of information generated and consumed worldwide, but most of the time this data is in a raw state and not appropriate for research. Fake news data sets are defined as curated collections of digital content specifically compiled for studying, detecting, and analyzing. These data sets typically include:

- **Fake and real news articles**: to train and test algorithms in distinguishing genuine content from fake.
- **Social media posts**: often annotated to identify patterns in disinformation propagation.
- **Fact-checking databases**: verified true or false claims, often sourced from professional fact-checking organizations. If a data sets includes this kind of information

Another important factor when talking about disinformation data sets are the main use that they are applied to since this will dramatically influence which data sets are useful and which aren't for a specific use case. Generally, there are four main goals when working with this data sets:

1. **Training ML models**. These data sets provide the foundation for training algorithms to classify content as fake or real. By offering labeled examples, they enable models to learn patterns and features associated with disinformation, such as linguistic markers, propagation patterns, and metadata anomalies.
2. **Understanding disinformation dynamics**. Researchers use these data sets to study how disinformation spreads across digital platforms, identifying critical nodes, influences, and pathways of dissemination. Insights from these studies inform counter-strategies to disrupt false information flows.
3. **Benchmarking detection techniques**. Disinformation data sets offer standardized benchmarks for evaluating the effectiveness of detection algorithms. By comparing performance across different data sets, researchers refine methods to improve accuracy and scalability [26].
4. **Evaluating fact-checking tools**. Fact-checking initiatives rely on annotated disinformation data sets to verify claims and enhance their algorithms. These data sets help identify gaps in automated tools and improve their reliability in real-world scenarios.

Some of the most relevant data sets that fall under some of this characteristics are:

- **LIAR Data Set**. This data set comprises 12,836 short statements sourced from *Politifact.com*, categorized into multiple truthfulness levels. It serves as a benchmark for NLP tasks aimed at detecting false claims. Studies like Wang [48] utilized LIAR to train algorithms that classify statement veracity accurately. Another usage for this data set is from Bahad et al. [18] where it was used to train bidirectional LSTM (Long Short-Term Memory) models, achieving high accuracy in distinguishing fake from real statements.
- **FakeNewsNet**. Combining content and social context, this data set includes news articles and user interactions from fact-checking organizations. Researchers leverage it to study how user engagement contributes to the spread of fake news. Shu et al. [45] used it to develop propagation-based disinformation detection models.
- **2016 X election data**. Formerly known as Twitter, this data set includes millions of tweets tied to the known disinformation campaign from the 2016 U.S. presidential election. It enables the study of bot-driven campaigns and their influence on public discourse. Bessi and Ferrara analyzed this data to reveal the critical role of automated accounts in spreading fake news [20]. Another study made in 2019 [24] revealed even more information about this data.
- **Hyperpartisan sources**. Comprising news articles labeled as fake, real, or satirical, this data set supports content-based disinformation detection. Pennycook and Rand made it to explore public susceptibility to false narratives and public perception of different sources [40].
- **COVID-19 fake news**. With the rise of the COVID-19 “infodemic” various data sets focusing on pandemic-related disinformation emerged. These include annotated posts, articles, and videos with credibility labels. Micallef et al. [35] used these data sets to study misinformation’s impact on public health and real-time mitigation strategies.

As previously mentioned, real-world data isn’t suitable for direct research on fake news. After examining several datasets, it becomes clear that the primary challenge lies in labeling. Determining whether information is true or false, or whether it is intentionally misleading, is a complex and challenging task that can only be solved using mixed approaches like crowdsourcing and ML.

3.2 Topic Modelling

Topic modeling is a statistical and computational technique used to identify hidden topics within large collections of unstructured text data. It is particularly useful in NLP and text mining applications for organizing, understanding, and summarizing vast amounts of textual information [28] [47] [16].

This section, though inherently complex, aims to provide an overview of the general principles and some of the most commonly employed techniques in the field. However, our focus will be to emphasize the concepts that are directly relevant to the development of this paper.

The primary goal of topic modeling is to group words in a corpus into clusters or "topics," where each topic is a distribution over words, and each document is a distribution over topics. This probabilistic representation allows researchers to uncover the thematic structure of a text corpus, even when the documents are not explicitly labeled with categories or topics.

3.2.1 Key concepts

As stated before topic modeling is a machine learning technique that uncovers latent structures in large collections of unstructured text data. This large collections are called *corpus* and are composed of multiple documents. This documents can be as large as research papers or as short as a few lines of text.

The primary goal of topic modeling is to group words in a corpus into groups or "*topics*", where each *topic* represents a collection of words that frequently occur together in a meaningful context. For instance, a topic on "sports" might include words like "team", "game", "score" and "player".

In this context a document inside the corpus is represented as a distribution over topics. For example, an article about a sporting event might be 70% "sports" and 30% "entertainment".

To accomplish this, topic modelling uses different mathematical foundation, but the most relevant is BoW (Bag-of-Words) model, where text is transformed into a matrix of word frequencies or term-document occurrences in which word order is ignored.

3.2.2 Common Techniques

To extract topics from a corpus, there exists multiple techniques, each one more appropriate and suitable for certain use cases, but with the a similar final result.

In first place there is the **Latent Dirichlet Allocation** or LDA, which was introduced in ML by Blei et al. [21] [14] and is one of the most popular algorithms for topic modeling. It assumes:

- Each document is a mixture of fixed number of topics.
- Topics are probability distributions over a vocabulary.

Conceptually LDA assumes documents have been generated through random sampling of pre-document topics and it attempts to reverse engineer this sampling. Mathematically, LDA defines the following process:

1. Assigning each document a multinomial distribution over topics, parameterized by a Dirichlet variable α.
2. Assigning each topic a multinomial distribution over words, with another Dirichlet variable β.
3. Generating documents by sampling words based on their topic distributions.

LDA employs inference algorithms like Gibbs Sampling [21] to estimate this variables mentioned before:

$$P\left(z_i = t \mid z^{-i}, w\right) = \frac{n_{m,t}^{-i} + \alpha}{\sum_{t'=1}^{T}\left(n_{m,t'}^{-i} + \alpha\right)} \times \frac{n_{t,w_i}^{-i} + \beta}{\sum_{v'=1}^{V}\left(n_{t,v'}^{-i} + \beta\right)}$$

However to understand this algorithms requires foundational knowledge in statistics and Monte Carlo techniques [15] which are outside of the scope of this paper. A schematic representation LDA can be seen in Figure 3.1 to better understand the process.

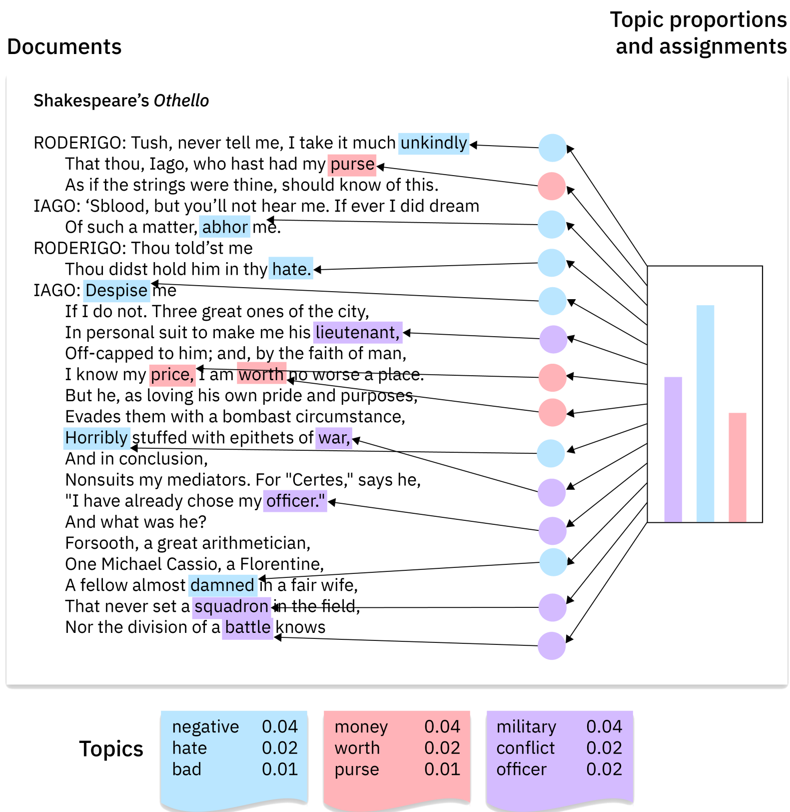

Figure 3.1: Schematic representation of LDA [21]

Other techniques include [47]:

- **Latent Semantic Analysis (LSA)**. An algebraic approach that uses Singular Value Decomposition (SVD) to identify latent patterns. Although fast, it is sensitive to noise and lacks probabilistic interpretability.
- **Probabilistic Latent Semantic Analysis (PLSA)**. A precursor to LDA that models topics as probability distributions but struggles with scalability and overfitting.
- **Correlated Topic Models (CTM)**. Builds on LDA by capturing correlations between topics using logistic normal distributions.

- **Pachinko Allocation Model (PAM)**. Introduces topic hierarchies, enabling the modeling of more complex relationships.
- **Dynamic Topic Models (DTM)**. An extension of LDA to capture temporal evolution in topics across time.

3.2.3 Evaluation Metrics

Evaluating the quality of topics is crucial in ensuring that the inferred topics are meaningful and useful. Several metrics are available for this purpose:

One of the most relevant is **perplexity**. Perplexity measures how well a probabilistic model predicts a set of unseen data. In ML it represents the number of words a model has to chose from when generating content. Perplexity values go from 1 to infinity, 1 being the ideal score. For example, if a ML model has a perplexity of 23.12 it means the model had to chose between 23 different words when generating content. The 0.12 decimal indicates slight deviations of the model and the closer they are to 0 the better [28] [47] [16] [44].

In the context of topic modeling, perplexity evaluates the likelihood of the test data given the model. Mathematically, it is defined as:

$$\text{Perplexity } = \text{exp} \left(-\frac{1}{N} \sum_{d=1}^{D} \text{log } P(\text{ d } d) \right)$$

Where $P(d)$ is the accumulated likelihood or probability of document d and N is the number of words in the corpus. Going deeper, to calculate $P(d)$ is necessary to accumulate all the probabilities of each word w given the document d. This is defined as:

$$P(w \mid d) = \sum_{t} P(w \mid t) \cdot P(t \mid d)$$

Where $P(w \mid t)$ is the likelihood of word w given the topic t and $P(t \mid d)$ is the likelihood of topic t being present in document d.

However, perplexity has limitations, as improvements do not always correlate with more coherent and better topics, but a less confused model that *can subsequently imply* better topics.

Other useful metrics are: - The **coherence score**, which evaluates the semantic similarity of words within a topic. For instance, a topic with high coherence would have words that fall under the same theme (football, ball, player), while one which low coherence will have words hardly related to each other (tree, glass, numbers) - Topic diversity and **exclusivity** are also relevant since they indicate the overlap between words across topics. - Surprisingly, **human interpretation** is a key aspect when evaluating topics

Thanks to all of this metrics, its possible to overcome some challenges when using topic modelling like choosing the number of topics to generate or properly preprocessing the corpus.

3.3 Complex networks

Complex networks are mathematical representations of systems composed of nodes (or vertices) and edges (or links) that describe relationships or interactions between them. These networks emerge in diverse fields, such as biology, sociology, computer science, physics, etc. reflecting the interconnected nature of real-world phenoms. Examples include social networks (e.g., Facebook or X), transportation networks (e.g., airline routes), biological systems (e.g., protein interaction networks), and the World Wide Web [34] [19] [11] [49] [31].

The field of complex network science is rooted in graph theory, which dates back to Euler's solution to the *Seven Bridges of Königsberg* problem in 1736 [6]. In Figure 3.2 an abstract representation of the bridges is shown. With this representation, Euler made the following observation: if there is a path crossing all bridges, but never the same bridge twice, then nodes with odd number of links must be either the starting or the end point of this path. Therefore, this path doesn't exist since all the nodes in the seven bridges of Königsberg have odd number of links.

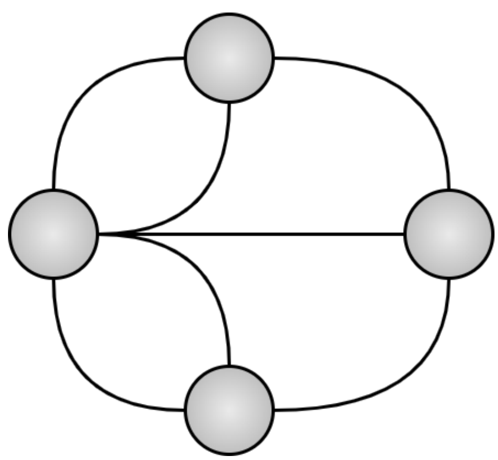

Figure 3.2: Graph of the seven Königsberg bridges.

As stated before, networks are often composed by nodes and links; nodes are normally used for representing entities or concepts, while the links represent the relationship or interactions between those entities or concepts. With this almost anything can be represented with networks since it's up to us to decide what are the "*nodes*" and what are the "*links*".

Another important feature of networks is that they can be *undirected* or *directed*. In a directed network, links go only one way: from the origin to the destination; while in a undirected network links are bi-directional and the order of the two nodes in a link does not matter.

Unlike regular graphs (e.g., grids or trees), complex networks exhibit irregular and heterogeneous patterns in their topology. They often possess characteristics such as **small-world** effects (short path lengths between nodes), **high clustering**, and **scale-free** properties (degree distributions following a power law). These features allow them to capture the complexity and diversity of real-world systems [34] [49].

The study of complex networks is a multidisciplinary one, combining graph theory, statistical physics, data science and more. By analyzing their structure and dynamics, researchers can uncover underlying principles governing the behavior of interconnected and real-world systems, such as the resilience of networks to failures, the spread of information, or the emergence of collective behavior.

3.3.1 Basic metrics

Understanding the structure of complex networks requires the computation of various metrics, which quantify aspects like connectivity, centrality, and clustering. These metrics serve as tools to analyze and compare networks systematically [19] [11].

The **degree** of a node is the number of edges connected to it. In directed networks, nodes have both an in-degree (incoming edges) and out-degree (outgoing edges). Formally, for a node i, the degree is:

$$k_i = \sum_j A_{ij}$$

Where A_{ij} is the adjacency matrix of the network, and $A_{ij} = 1$ if nodes i and j are connected.

Another related concept is **degree distribution**, which refers to the probability distribution of node degrees across the network. Many real-world networks exhibit a heavy-tailed degree distribution, where most nodes have a small degree, but a few hubs have a very high degree. In scale-free networks, this distribution follows a power law like the following:

$$P(k) \propto k^{-\gamma}$$

Scale-free networks, characterized by this distribution, are resilient to random failures but vulnerable to targeted attacks on hubs.

The **average path length** is the mean or average of the shortest paths between all pairs of nodes in the network. Networks with a small average path length exhibit the small-world phenomenon, where most nodes can be reached from any other node in a few steps. it is represented as:

$$L = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij}$$

Where N represents the nodes of the network and d_{ij} is the shortest path between nodes i and j.

The **clustering coefficient** measures the tendency of nodes to form tightly-knit groups. It is calculated for each node as the ratio of the number of existing links between its neighbors to the maximum possible links between them. For a node i, the clustering coefficient would be:

$$C_i = \frac{2T_i}{k_i\left(k_i - 1\right)}$$

Where T_i is the number of groups involving node i, and k_i is its degree. The average clustering coefficient is the mean of C_i across all nodes, indicating the network's overall tendency to form tightly-knit groups [19] [11].

3.3.2 Centrality measures

Centrality measures help identify the most relevant or influential nodes in a network. Different centrality metrics capture various aspects of "importance", such as connectivity, intermediary roles, or influence [19] [11].

Degree centrality is the simplest centrality measure, based on the number of edges a node has. This measure is particularly useful in undirected networks, where the number of connections directly correlates with influence.

Closeness centrality quantifies how close a node is to all other nodes in the network. Nodes with shorter distances to others have higher closeness centrality. For a node i, it is calculated as:

$$C_C(i) = \frac{1}{\sum_j d(i,j)}$$

Where $d(i,j)$ is the shortest path between nodes i and j.

The is also the **betweenness centrality** that measures the extent to which a node lies on the shortest paths between other nodes. Nodes with high betweenness centrality often serve as bridges or bottlenecks. This can be critical when undestanding control points or vulnerabilities in networks. For a node i, it is calculated as:

$$C_B(i) = \sum_{s \neq i \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}$$

Where σ_{st} is the total number of shortest paths from node s to node t, and $\sigma_{st}(i)$ is the number of those paths going through i.

Lastly, the **eigenvector centrality** assigns importance to nodes based on the importance of their neighbors. This "importance" can be calculated through the connectivity or centrality of a given node.

3.3.3 Community detection

A *community* is a group of nodes within a network that have stronger interactions among themselves than with the rest of the network. For example, in social networks, communities often represent groups with shared interests or activities [34] [31].

Community detection algorithms aim to identify densely connected groups of nodes within a network and can lead to unrevealing hidden information in complex networks such as data patterns or unique behaviors.

Regardless of the algorithm chosen, the quality of the detected communities must be evaluated using metrics such as:

- **Modularity**. Higher values indicate better-defined communities.
- **Normalized Mutual Information (NMI)**. Quantifies agreement between detected and ground-truth communities.
- **Conductance**. Evaluates the fraction of edges that leave a community relative to the total number of edges within and leaving the community.

Some of the most used algorithms for community detection are: - The **Louvain** method. This is one of the most popular and efficient algorithms for community detection. It is based on the optimization of **modularity**, a metric that quantifies the density of edges within communities relative to edges between communities. Although really good, it has its limitations, like producing disconnected communities in networks with low modularity and its processing being order dependent. The algorithm procedure cam be defined in 3 steps: - Initially, each node is assigned to its own community. - Nodes are iteratively reassigned to neighboring communities if doing so increases modularity. - Once no further improvement is possible, communities are aggregated into "super-nodes," creating a new network. The process is repeated on the aggregated network until modularity converges. - Related to the previous, the **Leiden** method builds upon the Louvain algorithm, addressing some of its limitations, particularly the issue of disconnected or poorly connected communities. The key difference lies in a refinement step in its procedure that ensures communities remain well-connected by splitting and merging disconnected components before proceeding to the next iteration.

3.4 Disinformation and complex networks

Once reached this point it is necessary to make an intersection between disinformation and complex networks and explore how researchers have employed various methodologies to analyze the structure, behavior, and impact of fake within complex networks:

- **Investigating Fake and Reliable News Sources Using Complex Networks**: This study aimed to identify disinformation spreaders by analyzing the relationships among websites using audience overlap metrics, commonly employed in SEO. A network of approximately 12,200 nodes was constructed, combining sub-networks based on these metrics. Complex network analysis was applied to visualize and analyze website relationships, highlighting the potential of SEO metrics and network analysis to identify communities of websites, including those disseminating fake news [32].
- **Modeling Disinformation Networks on Twitter: Structure, Behavior, and Impact**. This research investigated the structural and behavioral differences between disinformation sources and legitimate accounts on Twitter. Using interaction-based network structures, it examined the number of nodes, edges, clustering and other key metrics. Additionally, it explored dynamic aspects such as the speed and patterns of disinformation spread. The findings underscored the role of network topology in understanding and combating disinformation [37].
- **Disinformation Detection Using Graph Neural Networks: A Survey**. This survey reviewed approaches for automatic detection of disinformation, focusing on Graph Neural Networks (GNNs). It highlighted the strengths of GNNs in processing graph data and their suitability for social networks. The study emphasized that GNNs effectively model complex relationships and interactions, making them powerful tools for detecting disinformation on platforms like Twitter and Facebook [29].
- **A Unified Graph-Based Approach to Disinformation Detection Using Contextual and Semantic Relations**. This study proposed a graph-based framework that combines contextual and semantic information to improve disinformation detection. Using a meta-graph that integrates relational, semantic, and topical data, the approach was applied to Twitter datasets from the 2016 US election. The framework consistently improved accuracy in disinformation detection, demonstrating the value of combining multiple dimensions of graph data [39].
- **Automatic Detection of Influential Actors in Disinformation Networks**. Focusing on identifying key actors in disinformation campaigns, this study integrated natural language processing, machine learning, and graph analytics with a novel causal inference approach. Applied to hostile information campaigns, including datasets from the 2017 French elections, it achieved high precision in detecting disinformation accounts, mapped critical communities, and identified high-impact actors [46].
- **Behavioral Forensics in Social Networks: Identifying Misinformation, Disinforma-**

tion, and Refutation Spreaders Using Machine Learning. This research proposed a method to identify actors spreading misinformation, disinformation, and its refutation based on user behavior. By extracting network features through deep learning-based graph embedding models and training a machine learning classifier, the study demonstrated effective detection of malicious actors with high precision, highlighting the relevance of behavioral analysis in combating disinformation [27].

The approach of this paper although not at the same level as others previously mentioned, integrates multiple techniques such as topic modeling, complex networks, and community detection to explore disinformation in a new way.

A particularly novel aspect is the use of the **Minimum Spanning Tree (MST)** to distill the network into its core structure, retaining only the most critical relationships. This simplification enhances interpretability, which is often lacking in dense network analyses. Few studies employ MSTs in disinformation research, making this a standout feature.

Community detection on the MST adds another layer of insight. By identifying clusters of thematically and structurally similar documents, this approach uncovers **narrative communities** that highlight how disinformation themes propagate and interlink. Unlike approaches that stop at topic prevalence, this method integrates **content and structure**, offering a **system-wide perspective**.

Applying these techniques and topic modelling to PolitiFact.com (a dataset traditionally studied for classification accuracy) shifts the focus to structural relationships within fake news narratives.

Chapter 4

Tools and methodology

This chapter will expose all the material that made this project possible, explaining the what, how and when.

4.1 Methology

Chapter 5

Results

This chapter shows all the results obtained after using the tools from the previous chapter and applying multiple workflows that will be described in this chapter.

5.1 Proposed workflow

How i ended up on this topic about information source reliability and complex networks, why i chose this data,

Chapter 6

Conclusions

Discuss about the results obtained and what can we say about them. If topics has anything to do with reliability on information, if other factors count when searching reliability in information, etc.

6.1 Objective retrospective

Have we achieved the goals we defined? If so explain how, what and why

Mention future tasks to do, like applying more than one algorithm, comparative studies, apply workflows on other data, etc.

Bibliography

[1] Breaking Down The Numbers: How Much Data Does The World Create Daily in 2024? | Edge Delta.

[2] Data growth worldwide 2010-2028. *Statista*.

[3] 2024. Definition of DISINFORMATION.

[4] 2024. Definition of MISINFORMATION.

[5] Disinformation and public health.

[6] Euler: Koenigsberg bridges < Maze texts.

[7] Facts and Figures 2023 - Internet use.

[8] GIFCT about. *GIFCT*.

[9] Infodemic.

[10] Journalism, 'Fake News' and Disinformation: A Handbook for Journalism Education and Training.

[11] *Network Science by Albert-László Barabási*.

[12] Statistics. *ITU*.

[13] 2024. The 2022 Code of Practice on Disinformation | Shaping Europe's digital future.

[14] 2024. What is Latent Dirichlet allocation | IBM.

[15] 2021. What Is Monte Carlo Simulation? | IBM.

[16] 2024. What is topic modeling? | IBM.

[17] Aimeur, E. et al. 2023. Fake news, disinformation and misinformation in social media: A review. *SOCIAL NETWORK ANALYSIS AND MINING*. 13, 1 (Feb. 2023), 30. DOI:https://doi.org/10.1007/s13278-023-01028-5.

[18] Bahad, P. et al. 2019. Fake News Detection using Bi-directional LSTM-Recurrent Neural Network. *Procedia Computer Science*. 165, (Jan. 2019), 74–82. DOI:https://doi.org/10.1016/j.procs.2020.01.072.

[19] Barabási, A.-L. and PÃ3sfai, M. 2016. *Network Science*. Cambridge University Press.

[20] Bessi, A. and Ferrara, E. 2016. Social bots distort the 2016 U.S. Presidential election online discussion. *First Monday*. (Nov. 2016). DOI:https://doi.org/10.5210/fm.v21i11.7090.

[21] Blei, D.M. Latent Dirichlet Allocation.

[22] Britt, M.A. et al. 2019. A Reasoned Approach to Dealing With Fake News. *Policy Insights from the Behavioral and Brain Sciences*. 6, 1 (Mar. 2019), 94–101. DOI:https://doi.org/10.1177/2372732218814855.

[23] Garimella, K. and Chauchard, S. 2024. How prevalent is AI misinformation? What our studies in India show so far. *Nature*. 630, 8015 (Jun. 2024), 32–34. DOI:https://doi.org/10.1038/d41586-024-01588-2.

[24] Grinberg, N. et al. 2019. Fake news on Twitter during the 2016 U.S. Presidential election. *Science*. 363, 6425 (Jan. 2019), 374–378. DOI:https://doi.org/10.1126/science.aau2706.

[25] Gwiaździński, P. et al. 2023. Psychological interventions countering misinformation in social media: A scoping review. *Frontiers in Psychiatry*. 13, (Jan. 2023). DOI:https://doi.org/10.3389/fpsyt.2022.974782.

[26] Islam, M.R. et al. 2020. Deep learning for misinformation detection on online social networks: A survey and new perspectives. *Social Network Analysis and Mining*. 10, 1 (Sep. 2020), 82. DOI:https://doi.org/10.1007/s13278-020-00696-x.

[27] Khan, E.M. et al. 2023. Behavioral Forensics in Social Networks: Identifying Misinformation, Disinformation and Refutation Spreaders Using Machine Learning. arXiv.

[28] Kherwa, P. and Bansal, P. 2018. Topic Modeling: A Comprehensive Review. *ICST Transactions on Scalable Information Systems*. 7, (Jul. 2018), 159623. DOI:https://doi.org/10.4108/eai.13-7-2018.159623.

[29] Lakzaei, B. et al. 2024. Disinformation detection using graph neural networks: A survey. *Artificial Intelligence Review*. 57, 3 (Feb. 2024), 52. DOI:https://doi.org/10.1007/s10462-024-10702-9.

[30] Linden, S. van der et al. 2020. Inoculating Against Fake News About COVID-19. *Frontiers in Psychology*. 11, (Oct. 2020). DOI:https://doi.org/10.3389/fpsyg.2020.566790.

[31] Mata, A.S. da 2020. Complex Networks: A Mini-review. *Brazilian Journal of Physics*. 50, 5 (Oct. 2020), 658–672. DOI:https://doi.org/10.1007/s13538-020-00772-9.

[32] Mazzeo, V. and Rapisarda, A. 2022. Investigating Fake and Reliable News Sources Using Complex Networks Analysis. *Frontiers in Physics*. 10, (Jun. 2022). DOI:https://doi.org/10.3389/fphy.2022.886544.

[33] McKay, S. and Tenove, C. 2021. Disinformation as a Threat to Deliberative Democracy. *Political Research Quarterly*. 74, 3 (Sep. 2021), 703–717. DOI:https://doi.org/10.1177/1065912920938143.

[34] Menczer, F. et al. 2020. *A First Course in Network Science*. Cambridge University Press.

[35] Micallef, N. et al. 2020. The Role of the Crowd in Countering Misinformation: A Case Study of the COVID-19 Infodemic. arXiv.

[36] Michaels, D. and Osipovich, A. 2021. GameStop Stock Surge Tests Scope of SEC's Manipulation Rules. *Wall Street Journal*. (Jan. 2021).

[37] Muñoz, P. et al. 2024. Modeling disinformation networks on Twitter: Structure, behavior, and impact. *Applied Network Science*. 9, 1 (Dec. 2024), 1–35. DOI:https://doi.org/10.1007/s41109-024-00610-w.

[38] Oxford English Dictionary 2024. Misinformation, n. Oxford University Press.

[39] Paraschiv, M. et al. 2021. A Unified Graph-Based Approach to Disinformation Detection using Contextual and Semantic Relations. arXiv.

[40] Pennycook, G. and Rand, D.G. 2019. Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences*. 116, 7 (Feb. 2019), 2521–2526. DOI:https://doi.org/10.1073/pnas.1806781116.

[41] Rid, T. *Active Measures: The Secret History of Disinformation and Political Warfare*.

[42] S, H. and A, K. Too Much Information, Too Little Time: How the Brain Separates Important from Unimportant Things in Our Fast-Paced Media World. Front. Young Minds. DOI:https://doi.org/10.3389/frym.2017.00023.

[43] Safieddine, F. et al. 2016. Corporate Responsibility in Combating Online Misinformation. *International Journal of Advanced Computer Science and Applications (IJACSA)*. 7, 2 (Jan. 2016). DOI:https://doi.org/10.14569/IJACSA.2016.070217.

[44] Saxena, S. 2024. Understanding Perplexity in Language Models: A Detailed Exploration. *Medium*.

[45] Shu, K. et al. 2019. FakeNewsNet: A Data Repository with News Content, Social Context and Spatialtemporal Information for Studying Fake News on Social Media. arXiv.

[46] Smith, S.T. et al. 2021. Automatic Detection of Influential Actors in Disinformation Networks. arXiv.

[47] Vayansky, I. and Kumar, S.A.P. 2020. A review of topic modeling methods. *Information Systems*. 94, (Dec. 2020), 101582. DOI:https://doi.org/10.1016/j.is.2020.101582.

[48] Wang, W.Y. 2017. ”Liar, Liar Pants on Fire”: A New Benchmark Dataset for Fake News Detection. arXiv.

[49] Wang, X. and Chen, G. 2003. Complex networks: Small-world, scale-free and beyond. *Circuits and Systems Magazine, IEEE*. 3, (Feb. 2003), 6–20. DOI:https://doi.org/10.1109/MCAS.2003.1228503.